

Positive selection of primate *TRIM5 α* identifies a critical species-specific retroviral restriction domain

Sara L. Sawyer*, Lily I. Wu[†], Michael Emerman*[†], and Harmit S. Malik*[‡]

Divisions of *Basic Sciences and [†]Human Biology, Fred Hutchinson Cancer Research Center, Seattle, WA 98109

Communicated by Mark T. Groudine, Fred Hutchinson Cancer Research Center, Seattle, WA, December 29, 2004 (received for review December 8, 2004)

Primate genomes encode a variety of innate immune strategies to defend themselves against retroviruses. One of these, *TRIM5 α* , can restrict diverse retroviruses in a species-specific manner. Thus, whereas rhesus *TRIM5 α* can strongly restrict HIV-1, human *TRIM5 α* only has weak HIV-1 restriction. The biology of *TRIM5 α* restriction suggests that it is locked in an antagonistic conflict with the proteins encoding the viral capsid. Such antagonistic interactions frequently result in rapid amino acid replacements at the protein-protein interface, as each genetic entity vies for evolutionary dominance. By analyzing its evolutionary history, we find strong evidence for ancient positive selection in the primate *TRIM5 α* gene. This selection is strikingly variable with some of the strongest selection occurring in the human lineage. This history suggests that *TRIM5 α* evolution has been driven by antagonistic interactions with a wide variety of viruses and endogenous retroviruses that predate the origin of primate lentiviruses. A 13-aa "patch" in the SPRY protein domain bears a dense concentration of positively selected residues, potentially implicating it as an antiviral interface. By using functional studies of chimeric *TRIM5 α* genes, we show that this patch is generally essential for retroviral restriction and is responsible for most of the species-specific antiretroviral restriction activity. Our study highlights the power of evolutionary analyses, in which positive selection identifies not only the age of genetic conflict but also the interaction interface where this conflict plays out.

capsid | human endogenous retroviruses | HIV type 1 | SPRY

Innate immune strategies that defend primates against retroviruses are of considerable medical and evolutionary importance. Two distinct cellular antiviral defense systems, APOBEC3G/F (1–3) and *TRIM5 α* (4), that limit HIV infection have been described recently. APOBEC3G and APOBEC3F are cytidine deaminases that can cause hypermutation in the viral genome, but a viral accessory protein, Vif, can counteract their antiviral effect. *TRIM5 α* is a postentry restriction factor that accounts for the resistance to HIV type 1 (HIV-1) observed in rhesus monkey cells. It is not yet known how *TRIM5 α* mediates viral restriction, although a shorter, alternate transcript of the *TRIM5* gene has been shown to be an ubiquitin ligase (5). *TRIM5 α* restriction depends on the viral capsid, and its effect is saturable (4), although direct physical interaction between *TRIM5 α* and capsid has not been demonstrated. *TRIM5 α* from human and nonhuman primates also can restrict other lentiviruses and some strains of murine leukemia virus, a distantly related gammaretrovirus (6–9).

Although host genomes benefit from *TRIM5 α* 's recognition of viruses, it is in the best interest of the virus to evade recognition. Such antagonistic interactions have been formalized as the "Red Queen" hypothesis (10) and lead to the rapid fixation of amino acid replacements (positive selection), most likely at the interaction interface. The history of positive selection is thus informative for determining how long genes have been participants in genetic conflict, for identifying the likely sources of this conflict, and even for defining interaction domains involved. We previously have performed such an analysis on the *APOBEC* genes to show that APOBEC3G's role in

genome defense predates the origin of primate lentiviruses (11, 12) and that many other *APOBEC* cytidine deaminase genes likely participate in defending the primate genome against retroviruses.

Here, we show that the *TRIM5 α* restriction factor has undergone multiple episodes of positive selection that predate the origin of primate lentiviruses. Selection pressures on *TRIM5 α* vary widely among primate lineages, suggesting that distinct episodes of retroviral infection have dominated *TRIM5 α* evolution. The positive selection in *TRIM5 α* appears to be concentrated largely in the α -isoform-specific SPRY domain. We use this concentration of positively selected residues to propose and validate the presence of a SPRY "patch" that is essential for *TRIM5 α* 's general and species-specific restriction of retroviruses.

Materials and Methods

Primate Genomic DNA Sources. Genomic DNA samples obtained from Coriell Cell Repositories (Camden, NJ) are as follows (including Coriell repository numbers): *Pan troglodytes* (chimpanzee, NAO3448A), *Gorilla gorilla* (gorilla, NG05251B), *Pongo pygmaeus* (orangutan, NAO4272), *Erythrocebus patas* (patas monkey, NG06254), *Lagothrix lagotricha* (common woolly monkey, NG05356), *Ateles geoffroyi* (black-handed spider monkey, NGO5352), and *Saguinus labiatus* (red-chested mustached tamarin, NG05308). Genomic DNA from *Hylobates syndactylus* (island siamang, KB11539), *Colobus guereza kikuyuensis* (kikuyu colobus, OR160), *Pygathrix nemaeus* (douc langur, OR259), *Callithrix pygmaea* (pygmy marmoset, OR690), *Saimiri sciureus sciureus* (squirrel monkey, KB4544), *Callicebus donacophilus donacophilus* (Bolivian gray titi, OR1522), *Pithecia pithecia pithecia* (white-faced saki, KB5932), and *Alouatta sara* (Bolivian red howler, OR749) was obtained from the Center for Reproduction of Endangered Species FrozenZoo Project (San Diego Zoo, San Diego). *Papio anubis* (baboon) DNA was a personal gift from Trent Colbert (Fred Hutchinson Cancer Research Center). *Cercopithecus aethiops* (African green monkey) genomic DNA was prepared from Cos-7 cells with the QIAamp-DNA kit (Qiagen, Valencia, CA).

Sequencing of *TRIM5 α* Exons from Primate Genomic DNA. *TRIM5 α* was amplified and sequenced exon by exon from genomic DNA with PCR Supermix High Fidelity (Invitrogen) by using the PCR and sequencing primers shown in Table 2, which is published as supporting information on the PNAS web site. PCR products were sequenced directly, except in a few cases where they were first cloned into the TOPO TA cloning vector (Invitrogen), followed by sequencing of three independent clones. Exon reads

Abbreviations: dS, synonymous changes per site; dN, replacement changes per site; HIV-1, HIV type 1; NWM, New World monkey; OWM, Old World monkey.

Data deposition: All sequences reported in this paper have been deposited in the GenBank database (accession nos. AY843504–AY843520).

[†]To whom correspondence should be addressed at: Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, A1-162, Seattle, WA 98109. E-mail: hsmalik@fhcrc.org.

© 2005 by The National Academy of Sciences of the USA

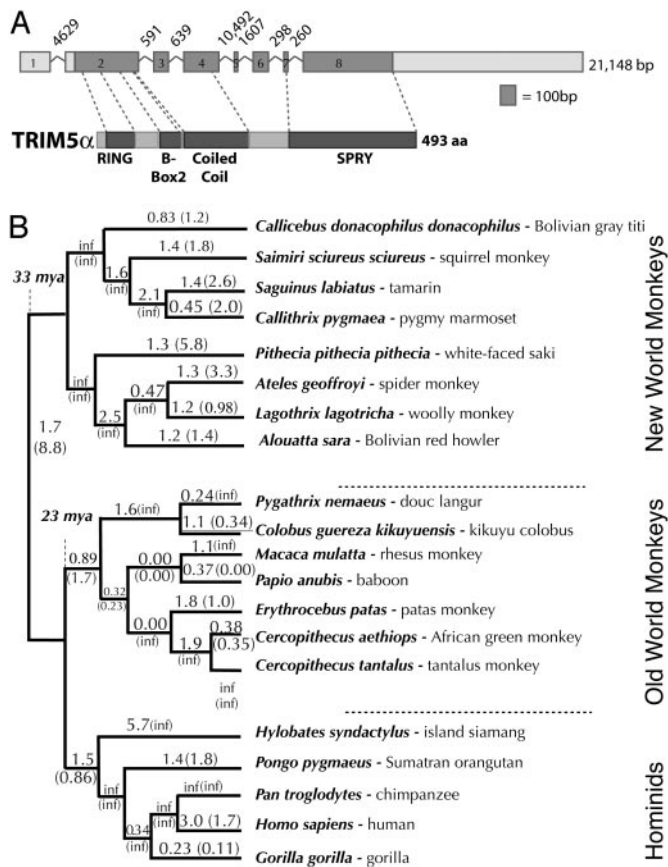


Fig. 1. Ancient and variable positive selection has shaped *TRIM5α* evolution. (A) *TRIM5α* is the longest of six reported transcripts of the eight-exon human *TRIM5* gene. It encodes a 493-aa protein consisting of a RING finger, a B-Box2, and a coiled-coil domain (signature domains of the TRIM family; ref. 24), as well as an α -isoform-specific SPRY domain. (B) *TRIM5α* was sequenced from a panel of primates representing 33 million years of evolutionary divergence. Values of dN/dS along each branch were calculated by using the free-ratio model of PAML for either the whole gene or for the SPRY domain alone (in parentheses), as shown on a cladogram of the accepted primate phylogeny (38). A dN/dS value of >1 suggests that positive selection has acted along that lineage. inf refers to cases where dS = 0.

were spliced together to create virtual transcripts for each primate, and have been entered into the GenBank database under the accession nos. AY843504–AY843520. The *TRIM5α* cDNA sequences for human (AY625000), *Macaca mulatta* (Rhesus macaque, AY523632.1), and *Cercopithecus tantalus* (tantalus monkey, AY593973.2) were obtained from the GenBank database. A phylogeny constructed by using the isolated *TRIM5α* genes is in good agreement with the accepted primate phylogeny, indicating that all sequences isolated by our PCR strategy are truly orthologous.

Confirmation of *TRIM5* Transcripts by Sequencing from RNA. The following monkey fibroblast cell lines were obtained from Coriell: *Ateles geoffroyi* (AG05352), *Lagothrix lagotricha* (AG05356), *Callicebus moloch* (AG06115A), and *Saguinus labiatus* (AG05308A). RNA was prepared from ≈ 1 million cells with the RNeasy kit (Qiagen). RT-PCR of the *TRIM5α* SPRY domain was performed with the primers shown in Table 2, by using the SuperScript One-Step kit (Invitrogen). PCR products were sequenced directly.

Sequence Analysis. DNA sequences were aligned by using CLUSTAL X (13), with hand alignment of small indels based on

Table 1. Positive selection in the *TRIM5α* gene

Data set	$-2(\ln \lambda)$	df	<i>P</i> value	dN/dS	Proportion of sites, %
All primates					
M0 vs. M3 (<i>k</i> = 3)	217.94	4	<i>P</i> < 0.0001	2.4, 7.2	33, 4
M1 vs. M2	109.88	2	<i>P</i> < 0.0001	4.5	16
M7 vs. M8	109.50	2	<i>P</i> < 0.0001	4.2	18
Hominids + OWMs					
M0 vs. M3 (<i>k</i> = 3)	57.42	4	<i>P</i> < 0.0001	1.8, 17	57, 2
M1 vs. M2	39.16	2	<i>P</i> < 0.0001	10	4
M7 vs. M8	40.21	2	<i>P</i> < 0.0001	10	4

We used likelihood ratio tests to determine whether any codon positions were associated with dN/dS significantly >1 and hence possibly subject to positive Darwinian selection. Neutral models (M0, M1, and M7) were compared to selection models (M2 and M8), which allow a proportion of codons for which dN/dS exceeds 1, or models for heterogeneity of dN/dS among sites (M3). As indicated by the *P* values, all analyses find very strong evidence for the selection model. Note that in M3, two classes of sites are permitted to have dN/dS >1 . Also indicated are the proportion of codons that were found to have dN/dS >1 , with the associated dN/dS values shown. Analyses using the f61 model of codon frequencies are shown, but similar results were obtained by using the f3 \times 4 frequency model (full PAML results are available in Appendix 1).

amino acid sequence. Maximum likelihood analysis was performed with CODEML in the PAML 3.14 software package (14). Global synonymous changes per site (dS)/replacement changes per site (dN) ratios for the tree (Fig. 1) were calculated by a free-ratio model, which allows dN/dS to vary along different branches. We tested whether dN/dS values were >1 in two lineages, those leading to gibbon and human, by using two methods as described in ref. 15 (see Appendix 1, which is published as supporting information on the PNAS web site). Briefly, likelihoods were compared when the lineage was fixed at dN/dS = 1 relative to when the lineage was allowed to have a dN/dS >1 (16). In the second method, we calculated the dN/dS ratios relative to the reconstructed ancestor and tested for significant deviations from dN/dS = 1 (16).

To detect selection in *TRIM5α*, multiple alignments were fitted to either the F3 \times 4 or F61 models of codon frequencies. Likelihood ratio tests of the data were performed by using different sets of site-specific (NS sites) models as follows: M0 (one-ratio) to M3 (discrete); M1 (two-state, neutral, dN/dS >1 disallowed) to M2 (selection, similar to model 1 but dN/dS >1 allowed); and M7 (fit to a beta distribution, dN/dS >1 disallowed) to M8 (similar to model 7 but dN/dS >1 allowed). In all cases, permitting sites to evolve under positive selection gave a much better fit to the data (Table 1). These analyses also identified certain amino acid residues with high posterior probabilities (>0.95) of having evolved under positive selection (17, 18) (Fig. 2; see also Appendix 1).

Secondary structure predictions and associated confidence values for the entire human *TRIM5α* protein were made by using the PSIPRED (19) Protein Structure Prediction Server (<http://bioinf.cs.ucl.ac.uk/psipred>).

Viral Infection Assays. CRFK (feline renal fibroblast) cells were obtained from the American Type Culture Collection and grown in DMEM/10% FCS. The human and rhesus *TRIM5α* alleles with a hemagglutinin epitope tag (4) in the LPCX retroviral vector were obtained from the National Institutes of Health AIDS Reagent Program (donated by Joseph Sodroski). Site-directed mutagenesis of the rhesus or human genes was performed by the QuikChange kit (Stratagene) in a single step (oligonucleotide sequences available on request). Retroviral vectors were produced in 293T cells, and supernatants were used to infect CRFK cells. Twenty-four hours after infection, 3 mg/ml

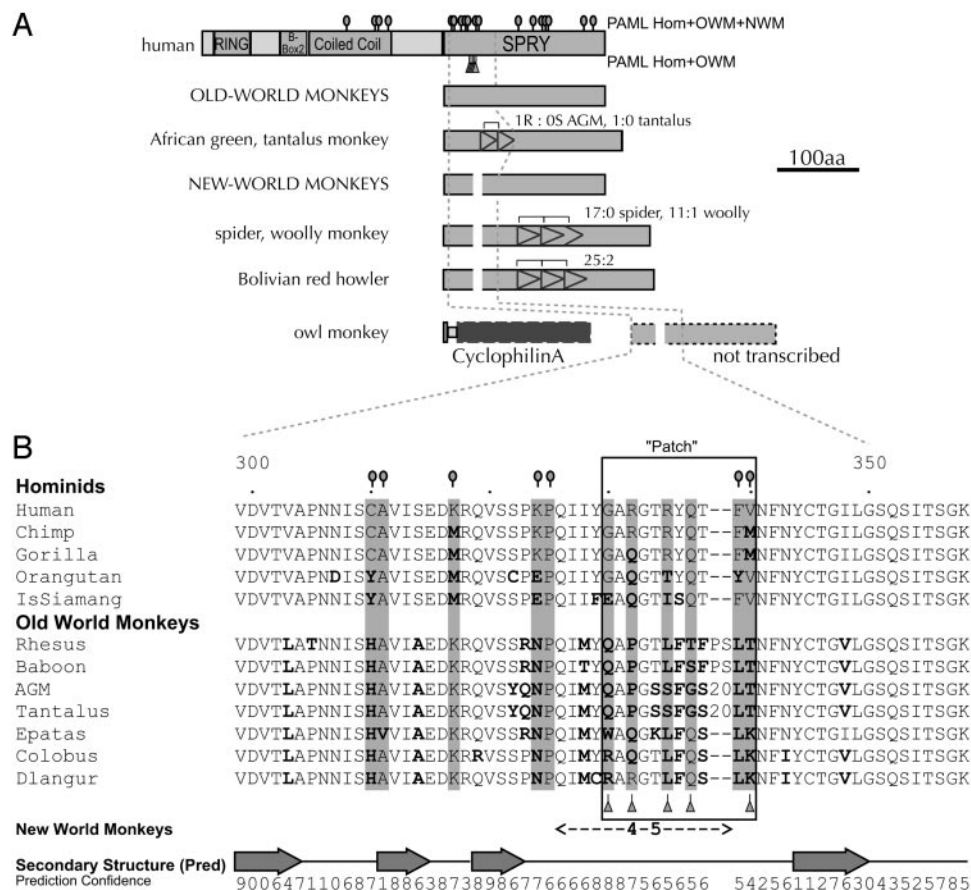


Fig. 2. The SPRY domain is a hot spot for insertions/deletions and positive selection. (A) Codons highlighted by PAML as being subject to positive selection with high posterior probabilities ($P > 0.95$) in an analysis of all primates are indicated as circles on stems above the TRIM5 α protein schematic. In a more limited analysis of just the hominids and Old World monkeys (OWMs), only five residues in the entire protein were identified as evolving under positive selection (triangles on stems below the TRIM5 α schematic). Representative SPRY domains from OWMs and New World monkeys are schematized, along with notable exceptions that have internal duplications (indicated by tandem arrowheads). These duplications appear to have accumulated more replacement (R) changes than synonymous (S) changes (R:S ratios indicated above). In owl monkeys, a CyclophilinA insertion (*TRIM-Cyp*) occurred between exons 7 and 8 (26), and it is believed that the 8th exon of *TRIM5* containing the SPRY domain is not transcribed. (B) Codons identified as being under positive selection are indicated in gray background and by using the same symbols as in A. Changes relative to human are indicated in bold. NWMs have a large deletion in the area of the patch (see Data Set 1, which is published as supporting information on the PNAS web site). Secondary structure predictions and confidence values (0, low; 9, high) were made by using the PSIPRED server (19). Arrows indicate β -strands, and lines indicate coils. According to this prediction, most of the TRIM5 α protein is predicted to be comprised of α -helices, whereas the entire SPRY domain is predicted with high confidence to be β -strands and coils with no α -helices.

puromycin was added to the medium to select for transduced cells to obtain pools of cells that stably expressed exogenous *TRIM5 α* genes. Expression was verified by Western blotting using a hemagglutinin antibody (Babco, Richmond, CA).

Single-cycle assays for HIV-1 were performed essentially as described in ref. 20. HIV-1 provirus was deleted for *env*, and GFP or luciferase was inserted into the *nef* region (20). SIVag-mTAN virus (21) was a kind gift of Ned Landau (The Salk Institute, San Diego). Virus was produced in 293T cells by cotransfection of the provirus with the VSV-G gene, titrated on cells without restriction factors, and frozen. Infection of CRFK cells with HIV-1 (GFP) was performed in 24-well plates with 3×10^4 cells per well by using 5-fold dilutions of virus that had been titrated previously to give between 1% and 80% infection. Two days after infection, the cells were fixed with 2% paraformaldehyde and analyzed by flow cytometry for GFP expression. Luciferase assays were performed with virus stocks that had been titrated previously to be in the linear range of the assay (between 10,000 and 1 million relative light units). Infections were performed in triplicate in 24-well plates, assayed with the luciferase assay kit (Promega) in a 96-well plate, and read on a luminometer.

Results

We sequenced the *TRIM5 α* gene ($\approx 1,482$ bp of coding sequence) from 17 primate genomes that represent 33 million years of evolution (Fig. 1). Comparison of the rates of nonsynonymous (that alter the encoded amino acid) and synonymous DNA changes between species can be used to assess the types of selective pressures that have acted on a gene (22). For most protein-coding genes, dS exceeds dN because amino acid replacements are generally detrimental to protein function and therefore are culled out of the population (purifying selection). We find that many branches of the primate phylogeny, including internal branches, show evidence for *TRIM5 α* evolution under positive selection (defined as $dN/dS > 1.0$; Fig. 1B). Thus, *TRIM5 α* has been subject to positive selection for at least 33 million years. This selection has been strikingly variable because dN/dS ratios along different branches are significantly different from each other ($P < 0.02$; Appendix 1). In contrast, the positive selection on *APOBEC3G* was found to be more constant (11, 12) because ω values along each branch did not differ significantly from each other ($P > 0.75$; Appendix 1). Positive selection of *TRIM5 α* is especially strong in the hominid clade, with the

highest whole-gene dN/dS values of 5.7 and 3.0 found in the lineages leading to island siamang (gibbon) and human, respectively. We tested for the presence of positive selection in the gibbon and human branches by both comparative two-ratio likelihood tests by using PAML (15) and Monte-Carlo simulation using K-ESTIMATOR (16) (see *Materials and Methods* and Appendix 1). The average dS in *TRIM5 α* is not unusually low; it is 0.084 between hominids and Old World monkeys (OWMs) and 0.153 between hominids and New World monkeys (NWMs), compared with previous estimates of 0.08 and 0.15, respectively, for substitution rates in various intronic and noncoding regions of primate genomes (23). Thus, we can rule out the possibility that selection has led to deflated dS values in *TRIM5 α* , resulting in artificially high dN/dS ratios.

TRIM5 is a member of the large tripartite motif family in primate genomes, characterized by having RING finger, B-box, and coiled-coil domains (24). The α isoform of *TRIM5* has an additional SPRY protein domain (Fig. 1A), which is found in many proteins including those in the Ig superfamily (25). This SPRY domain has been shown previously to be essential for the restriction of HIV-1 (4). Although little is known about SPRY function, we found that this domain had undergone the most intense positive selection. This finding is evidenced by the high dN/dS values obtained in an analysis of this domain alone (in parenthesis on cladogram, Fig. 1B), including a striking dN/dS of 8.8 in the branch separating NWM from OWM and hominids (see also Fig. 4, which is published as supporting information on the PNAS web site). In addition to the strong positive selection, the SPRY domain has undergone an unusual number of insertions and deletions (Fig. 2A). A small deletion has occurred in the lineage leading to the NWMs, whereas there have been two distinct instances of internal duplications, one in African green monkeys and close relatives and a different triplication in the lineage leading to spider, woolly, and howler monkeys. The African green monkeys and tantalus duplication has been verified in *TRIM5 α* transcripts (GenBank accession nos. AY625003 and AY593973). We confirmed the woolly and spider monkey triplications by sequencing RT-PCR products (data not shown). When the duplicated and triplicated sequences within a single gene were aligned to each other, they showed an inflated number of nonsynonymous to synonymous changes (e.g., 25 replacement:2 synonymous changes for howler monkey; Fig. 2A). Thus, internal duplications in the SPRY domain followed by positive selection predict that these sequences are functionally important for the ability of *TRIM5 α* to restrict different viruses. Finally, in owl monkey, the entire SPRY domain of *TRIM5 α* has been replaced by a retrotransposed *CyclophilinA* gene, now referred to as *TRIM-Cyp* (26). *CyclophilinA* has been reported to directly interact with viral capsid (27), and, because the SPRY domain can be functionally replaced by *CyclophilinA*, this finding suggests that the SPRY also might be a capsid-interacting domain.

We used a maximum-likelihood approach, by using the PAML suite of programs (14), to determine whether specific codons of *TRIM5 α* have been repeatedly subjected to positive selection in primates. Statistical results support the presence of positive selection with extraordinary confidence, regardless of whether we include or exclude NWMs in our analysis (Table 1). Analysis of the full data set identifies $\approx 18\%$ of the codons as having evolved under positive selection with an average dN/dS ratio of >4 (Table 1, NSsites model 8 vs. 7). Of these, residues identified as being under positive selection with high confidence ($>95\%$ posterior probability) are illustrated on a schematic of the *TRIM5 α* protein (circles on stems, Fig. 2) and fall in either the coiled-coil or SPRY domains. Some TRIM proteins have been shown to homomultimerize through their coiled-coil domains (24), but our results also suggest that the coiled-coil domain of *TRIM5 α* may additionally participate in host defense. A secondary structure prediction of the SPRY domain suggests that

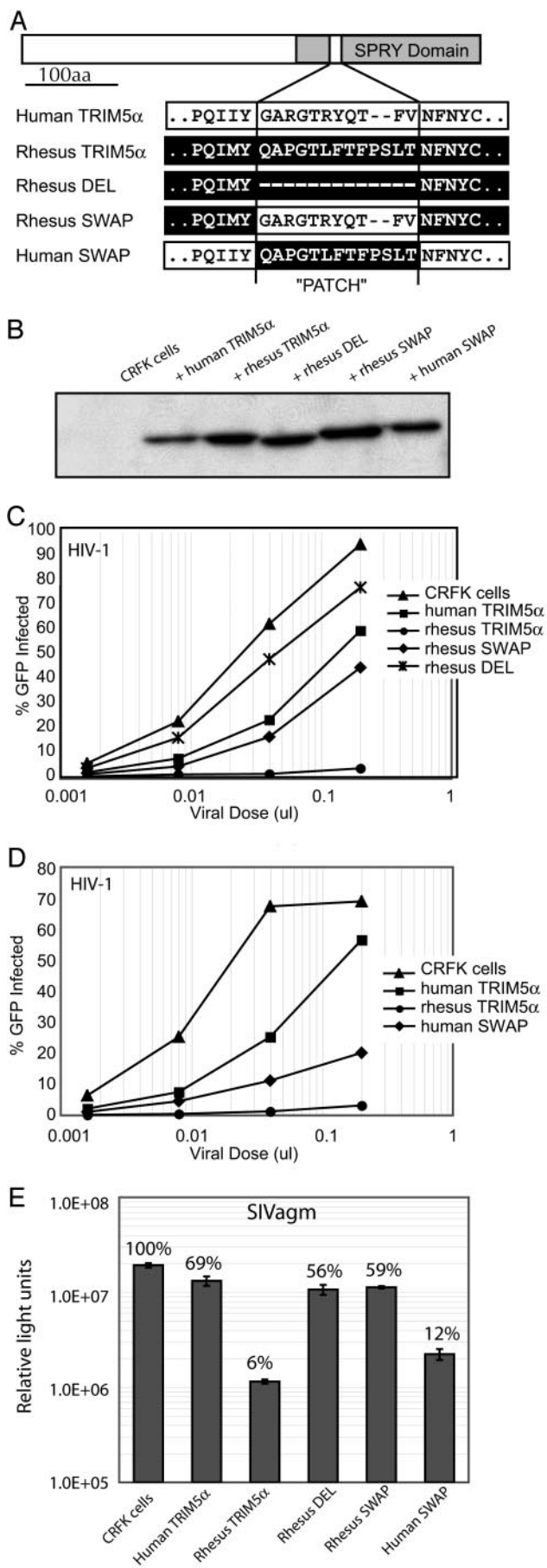
the positively selected residues and insertions/deletions fall exclusively in predicted coils (Fig. 2B), which could represent specific interaction surfaces.

Because there are NWM-specific deletions in *TRIM5 α* , excluding NWMs allows the opportunity to analyze all residues in hominids and OWMs, a more focused look at just the last 23 million years of primate phylogeny. Remarkably, the analysis of hominids and OWMs identified only five residues in the entire protein as being under positive selection with high confidence (triangles on stems, Fig. 2A). All five residues fall within an 11- to 13-aa segment of the SPRY domain (Fig. 2B), which we will refer to as the SPRY patch. Such tight clusters of positive selection are predicted to be points of physical contact between two proteins locked in genetic conflict. Similar PAML analyses have successfully highlighted the known binding surface between ZP3, an egg-receptor protein, and sperm (28). Additionally, the positive selection in the major histocompatibility complex (MHC) proteins is confined to small segments of the protein known to constitute the antigen-recognition site (29). In this case, we had no *a priori* knowledge of the *TRIM5 α* interaction interface crucial for viral restriction, so we tested our computational prediction that the SPRY patch identifies such a domain. Because *TRIM5 α* was originally identified because of the fact that it confers resistance to HIV-1 infection in rhesus, but not in human, cells (4) we used this species-specific example of *TRIM5 α* restriction to investigate the functional importance of the patch.

We constructed chimeric proteins between human and rhesus *TRIM5 α* that either differed only in the 11- to 13-aa patch region or had the patch region deleted altogether (Fig. 3A). CRFK cells (feline cells that have no known retroviral restriction) were stably transduced to express these proteins at approximately similar levels (Fig. 3B) and were tested for restriction of increasing titer of a recombinant HIV-1 virus that expresses GFP (20). As has been previously reported (4), human *TRIM5 α* showed a weak restriction of HIV-1, whereas rhesus *TRIM5 α* almost completely blocked infection (Fig. 3C). When the patch is deleted from rhesus *TRIM5 α* (rhesus DEL), it loses restriction not only against HIV-1 (Fig. 3C) but also against SIV from African green monkeys and N-tropic murine leukemia virus (Fig. 3E and data not shown). Thus, the patch we have identified based on positive selection may define a protein surface that is generally necessary for broad retroviral restriction.

We next asked whether the patch is responsible for the species specificity of *TRIM5 α* by substituting the human patch into rhesus *TRIM5 α* and vice versa. Remarkably, we found that substitution of the human patch into the rhesus *TRIM5 α* (rhesus SWAP) reduces its restrictive capability to a level that is close to that of human *TRIM5 α* (Fig. 3C). Thus, the patch we have defined by using positive selection identifies a region necessary for species-specific restriction to HIV-1. We then tested the reciprocal chimera, human *TRIM5 α* with the rhesus *TRIM5 α* patch (human SWAP, Fig. 3D). The Human SWAP was significantly more restrictive against HIV-1 than human *TRIM5 α* . Despite this gain-of-function effect, swapping the patch alone was not completely sufficient to recapitulate rhesus *TRIM5 α* -like restriction, suggesting that other minor determinants of restriction also may exist.

We wanted to know whether the patterns seen with these chimeric proteins are specific to HIV-1 or are generally responsible for defining broad species-specific ability to restrict lentiviruses. To address this question, we challenged the *TRIM5 α* -expressing CRFK cell lines with SIVagm and found that the reversal of species-specific restriction is the same as was observed for HIV-1 (Fig. 3E). Thus, human *TRIM5 α* does not usually restrict SIVagm, but can acquire the ability to restrict SIVagm when the 13-aa patch from rhesus *TRIM5 α* is added (human SWAP in Fig. 3E). Conversely, rhesus *TRIM5 α* loses the ability to restrict SIVagm when its patch is replaced by the 11-aa patch from human *TRIM5 α* (rhesus SWAP in Fig. 3E). Thus, the



patch we have identified through positive selection determines the species-specificity of TRIM5 α for restricting at least two highly divergent lentiviruses.

Discussion

Positive selection is a beacon for domains involved in genetic conflict. Based on analysis of positively selected codons, we were able to predict an important protein domain (potentially an interaction surface) that we then verified *in vivo*. In addition, we have illustrated strong positive selection of the TRIM5 α gene throughout primate evolution. The ancient and intense selective pressures that have shaped the evolution of APOBEC3G (11) and TRIM5 α indicate that the innate immune system is intricately evolved, just as the adaptive immune system is known to be (30). The tight clustering of positively selected residues in TRIM5 α is in stark contrast to our earlier findings in APOBEC3G, where the codons that were identified as having been repeatedly subject to positive selection were scattered throughout the length of the gene, with no significant clustering (11). There are two possible reasons for this difference. First, the TRIM5 α protein has a nonredundant domain structure, whereas APOBEC3G has resulted from two sequential duplications of a single cytidine deaminase domain to result in four potentially redundant structural copies of this domain (31, 32). Second, TRIM5 α and APOBEC3G play different roles in their respective conflicts. Although it is in the host genome's best interest that TRIM5 α improve binding to its target, it is in the host's interest that APOBEC3G avoid interactions with viral antagonists such as Vif in the case of HIV-1. Therefore, unlike TRIM5 α , APOBEC3G is under pressure to mutate whichever domain is the current viral target.

What drives the evolution of TRIM5 α ? The antiquity of the positive selection rules out primate lentiviruses like HIV-1 as being the sole, or even the major, cause because they are believed to be <1 million years old (33). In addition, TRIM5 α from human and OWM has been shown to be active against murine leukemia virus (6–9), a gammaretrovirus that is closely related to human endogenous retroviruses (34) that have episodically invaded primate genomes and continue to be active in the human genome (35–37). This finding suggests that TRIM5 α evolution may have been strongly influenced by distinct episodes of endogenous retrovirus infection and subsequent retrotransposition events (36). HIV and other primate lentiviruses are likely to be newcomers to this conflict, with the OWM TRIM5 α restriction against HIV-1 just an evolutionary coincidence.

Fig. 3. The positively selected patch is critical for the species-specific retroviral restriction by TRIM5 α . (A) Retroviral vectors (LPCX) containing hemagglutinin-tagged TRIM5 α alleles from human and rhesus genomes were modified to delete the rhesus SPRY patch (Rhesus DEL), swap the human patch into rhesus TRIM5 α (Rhesus SWAP), or swap the rhesus patch into the human TRIM5 α (Human SWAP). The sequences in the region of the swaps are shown. (B) A Western blot (using a hemagglutinin antibody) shows stable CRFK lines expressing each construct at roughly equivalent levels. (C) Variants of the rhesus TRIM5 α protein were assessed for their ability to restrict HIV-1. Single-cycle assays for HIV-1 infectivity were carried out in CRFK (feline renal fibroblasts) cells expressing exogenous TRIM5 α genes. Experiments were performed by using 5-fold dilutions of an HIV-1 provirus that was deleted for *env* and where GFP was inserted into the *nef* region (20). HIV-positive cells were detected by flow cytometry for GFP expression. (D) A variant of the human TRIM5 α (human SWAP) was tested in single-cycle assays for HIV-1 infectivity as in C. (E) CRFK cells that express the TRIM5 α alleles were challenged with SIVagm virus, which expresses *luciferase* in place of the *nef* gene. Preliminary experiments were performed to establish a dose of virus within the linear range of infection (data not shown). All infections were performed in triplicate, and luciferase activity was measured 2 days after infection. The relative light units are shown on the y-axis, and the percent infection (compared with the CRFK control) is shown above the bar. SD of the data also are shown.

Nevertheless, as we have shown, the positive selection has had a profound impact on the species-specific restriction of both HIV-1 and SIV_{agm}. Thus, the evolutionary histories of both *TRIM5* and *APOBEC3G* indicate ancient adaptation to endogenous retrovirus-like elements (11, 12), yet both restriction systems were discovered because of their incidental activity against HIV. These findings indicate that the cellular arsenal honed against endogenous retroviruses is large and mostly undiscovered and may strongly impact lentiviral restriction.

We thank Monica Rodriguez and the Fred Hutchinson Cancer Research Center flow cytometry facility for technical assistance. We thank the

National Institutes of Health AIDS Reagent Program for TRIM5 α alleles contributed by Joseph Sodroski; Ned Landau (Salk Institute) for SIV vectors; Jeremy Luban and David Sayah for communicating results and sequences before publication; and Coriell, FrozenZoo, and Trent Colbert for primate genomic DNA. We thank our colleagues Sue Biggins, Julie Kerns, Danielle Vermaak, and Janet Young for comments on the manuscript, and Jorja Henikoff, Trygve Bakken, and an anonymous reviewer for valuable suggestions and assistance with the PAML implementation. This work was supported by start-up funds from the Fred Hutchinson Cancer Research Center (to H.S.M.) and National Institutes of Health Grants R37 AI30927 (to M.E.) and T32 CA 09657 (to S.L.S.). H.S.M. is an Alfred P. Sloan Research Fellow in Computational and Evolutionary Molecular Biology.

1. Sheehy, A. M., Gaddis, N. C., Choi, J. D. & Malim, M. H. (2002) *Nature* **418**, 646–650.
2. Wiegand, H. L., Doehle, B. P., Bogerd, H. P. & Cullen, B. R. (2004) *EMBO J.* **23**, 2451–2458.
3. Zheng, Y.-H., Irwin, D., Kurosu, T., Tokunaga, K., Sata, T. & Peterlin, B. M. (2004) *J. Virol.* **78**, 6073–6076.
4. Stremlau, M., Owens, C. M., Perron, M. J., Kiessling, M., Autissier, P. & Sodroski, J. (2004) *Nature* **427**, 848–853.
5. Xu, L., Yang, L., Moitra, P. K., Hashimoto, K., Rallabhandi, P., Kaul, S., Meroni, G., Jensen, J. P., Weissman, A. M. & D'Arpa, P. (2003) *Exp. Cell Res.* **288**, 84–93.
6. Yap, M. W., Nisole, S., Lynch, C. & Stoye, J. P. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 10786–10791.
7. Hatzioannou, T., Perez-Caballero, D., Yang, A., Cowan, S. & Bieniasz, P. D. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 10774–10779.
8. Keckesova, Z., Ylinen, L. M. & Towers, G. J. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 10780–10775.
9. Perron, M. J., Stremlau, M., Song, B., Ulm, W., Mulligan, R. C. & Sodroski, J. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 11827–11832.
10. Van Valen, L. (1973) *Evol. Theory* **1**, 1–30.
11. Sawyer, S. L., Emerman, M. & Malik, H. S. (2004) *PLoS Biol.* **2**, 1278–1285.
12. Zhang, J. & Webb, D. M. (2004) *Hum. Mol. Genet.* **13**, 1785–1791.
13. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997) *Nucleic Acids Res.* **25**, 4876–4882.
14. Yang, Z. (1997) *Comput. Appl. Biosci.* **13**, 555–556.
15. Yang, Z. (1998) *Mol. Biol. Evol.* **15**, 568–573.
16. Comeran, J. M. (1999) *Bioinformatics* **15**, 763–764.
17. Nielsen, R. & Yang, Z. (1998) *Genetics* **148**, 929–936.
18. Yang, Z. & Bielawski, J. P. (2000) *Trends Ecol. Evol.* **15**, 496–503.
19. McGuffin, L. J., Bryson, K. & Jones, D. T. (2000) *Bioinformatics* **16**, 404–405.
20. Yamashita, M. & Emerman, M. (2004) *J. Virol.* **78**, 5670–5678.
21. Mariani, R., Chen, D., Schrofelbauer, B., Navarro, F., Konig, R., Bollman, B., Munk, C., Nymark-McMahon, H. & Landau, N. R. (2003) *Cell* **114**, 21–31.
22. Hurst, L. D. (2002) *Trends Genet.* **18**, 486–487.
23. Li, W.-H. (1997) *Molecular Evolution* (Sinauer, Sunderland, MA).
24. Reymond, A., Meroni, G., Fantozzi, A., Merla, G., Cairo, S., Luzi, L., Riganelli, D., Zanaria, E., Messali, S., Cainarca, S., *et al.* (2001) *EMBO J.* **20**, 2140–2151.
25. Henry, J., Mather, I. H., McDermott, M. F. & Pontarotti, P. (1998) *Mol. Biol. Evol.* **15**, 1696–1705.
26. Sayah, D. M., Sokolskaja, E., Berthoux, L. & Luban, J. (2004) *Nature* **430**, 569–573.
27. Luban, J., Bossolt, K. L., Franke, E. K., Kalpana, G. V. & Goff, S. P. (1993) *Cell* **73**, 1067–1078.
28. Swanson, W. J., Yang, Z., Wolfner, M. F. & Aquadro, C. F. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 2509–2514.
29. Hughes, A. L. & Nei, M. (1988) *Nature* **335**, 167–170.
30. Holmes, E. C. (2004) *PLoS Biol.* **2**, 1267–1269.
31. Jarmuz, A., Chester, A., Bayliss, J., Gisbourne, J., Dunham, I., Scott, J. & Navaratnam, N. (2002) *Genomics* **79**, 285–296.
32. Wedekind, J. E., Dancy, G. S. C., Sowden, M. P. & Smith, H. C. (2003) *Trends Genet.* **19**, 207–216.
33. Sharp, P. M., Bailes, E., Robertson, D. L., Gao, F. & Hahn, B. H. (1999) *Biol. Bull.* **196**, 338–342.
34. Andersson, M. L., Lindeskog, M., Medstrand, P., Westley, B., May, F. & Blomberg, J. (1999) *J. Gen. Virol.* **80**, 255–260.
35. Berkhout, B., Jebbink, M. & Zsiros, J. (1999) *J. Virol.* **73**, 2365–2375.
36. Smit, A. F. (1999) *Curr. Opin. Genet. Dev.* **9**, 657–663.
37. Belshaw, R., Pereira, V., Katzourakis, A., Talbot, G., Paces, J., Burt, A. & Tristem, M. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 4894–4899.
38. Purvis, A. (1995) *Philos. Trans. Biol. Sci.* **348**, 405–421.

Appendix 1. PAML summary results for TRIM5 α

A. PAML (1) Analysis of Entire Data Set (hominids+OWM+NWM)

1. Log Likelihood Scores and Parameter Estimates for Four Models of Variable ω 's Among Sites Assuming the F3x4 Model of Codon Frequencies.

Model	Parameter Estimates	Sites* with $\omega > 1$	ℓ
M0: one ratio	Average dN/dS for each branch = 1.160 S = 1.55260		-5772.38
Site models			
M1: neutral	$(\omega_0 = 0), f_0 = 0.354$ $(\omega_1 = 1) (f_1 = 0.646)$	Not allowed	-5,725.00
	Average dN/dS for each branch = 0.646 S = 1.56259		
M2: selection	$(\omega_0 = 0), f_0 = 0.281$ $(\omega_1 = 1), f_1 = 0.566$ $\omega_2 = 4.624 (f_2 = 0.153)$ Average dN/dS for each branch = 1.273 S = 1.62074	(47–51: gaps) 175 Q 0.9958** 213 S 0.9635* 215 T 0.9955** 228 L 0.9942** 310 C 0.9924** 311 A 0.9603* 317 K 0.9757* 324 K 0.9995** 325 P 0.9933** (326–338: gaps) 339 F 0.9994** 340 V 0.9999** (385–386: gaps) 389 K 0.9953** 407 G 0.9974** 418 F 0.9997** 421 P 0.9769* 423 V 0.9633* 471 Q 0.9535*	-5,667.01

Model	Parameter Estimates	Sites* with $\omega > 1$	ℓ
		483 G 0.9936**	
M3:discrete	$\omega_0 = 0.342$ $f_0 = 0.662$ $\omega_1 = 2.575$ ($f_1 = 0.305$) $\omega_2 = 7.922$ ($f_2 = 0.033$) Average dN/dS for each branch = 1.273 S = 1.63377	Too many to list. Note: M3 vs. M0 is a test of heterogeneity among sites and not an explicit test of positive selection.	-5,664.25
M7: β	$p = 0.02830$ $q = 0.01638$ Average dN/dS for each branch = 0.6160 S = 1.55777	Not allowed	-5,724.62
M8: β and ω	$p = 0.2299$ $q = 0.14801$ $f_0 = 0.829$ $\omega_1 = 4.308$ ($f_1 = 0.171$) Average dN/dS for each branch = 1.2426 S = 1.62038	(47-51: gaps) 139 Q 0.9534* 175 Q 0.9967** 213 S 0.9739* 215 T 0.9966** 228 L 0.9958** 257 T 0.9511* 310 C 0.9943** 311 A 0.9717* 317 K 0.9830* 324 K 0.9996** 325 P 0.9946** (326-338: gaps) 339 F 0.9995** 340 V 0.9999** 384 D 0.9539* (385-386: gaps) 389 K 0.9964** 390 N 0.9611* 407 G 0.9979** 418 F 0.9998** 421 P 0.9830* 423 V 0.9737*	-5,666.71

Model	Parameter Estimates	Sites* with $\omega > 1$	ℓ
		471 Q 0.9670* 483 G 0.9952**	

p and q are parameters of the β distribution. f is the proportion of sites assigned to an individual ω (= dN/dS) category or to a β distribution with shape parameters p and q . The proportion $f_{1,2}$ in parentheses is not a free parameter. Sites assigned to $\omega_{1,2}$ are those with posterior probabilities (P) > 0.95 , and those with $P > 0.99$ are in bold. Analyses were conducted using κ as a free parameter. S = tree length.

Sites assigned to $\omega > 1$ are those with posterior probabilities $P > 0.95$ () or $P > 0.99$ (**).

2. Log Likelihood Scores and Parameter Estimates for Four Models of Variable ω 's Among Sites Assuming the F61 Model of Codon Frequencies.

Model	Parameter estimates	Sites* with $\omega > 1$	ℓ
M0: one ratio	Average dN/dS for each branch = 1.144 S = 1.55005		-5,748.06
Site models			
M1: neutral	$(\omega_0 = 0) f_0 = 0.359$ $(\omega_1 = 1) (f_1 = 0.641)$	Not allowed	-5,696.78
	Average dN/dS for each branch = 0.641 S = 1.56922		
M2: selection	$(\omega_0 = 0) f_0 = 0.291$ $(\omega_1 = 1) f_1 = 0.551$ $\omega_2 = 4.451 (f_2 = 0.158)$ Average dN/dS for each branch = 1.2545 S = 1.62398	(47-51: gaps) 175 Q 0.9954** 213 S 0.9544* 215 T 0.9970** 228 L 0.9957** 310 C 0.9930** 311 A 0.9669* 317 K 0.9736* 324 K 0.9994** 325 P 0.9903**	-5,641.84

Model	Parameter estimates	Sites* with $\omega > 1$	ℓ
		(326-338: gaps) 339 F 0.9987** 340 V 0.9999** (385-386: gaps) 389 K 0.9935** 407 G 0.9958** 411 S 0.9536* 418 F 0.9994** 421 P 0.9784* 423 V 0.9680* 483 G 0.9909**	
M3:discrete	$\omega_0 = 0.301$ $f_0 = 0.631$ $\omega_1 = 2.379$ ($f_1 = 0.328$) $\omega_2 = 7.224$ ($f_2 = 0.041$) Average dN/dS for each branch = 1.2692 S = 1.63693	Too many to list. Note: M3 vs. M0 is a test of heterogeneity among sites and not an explicit test of positive selection.	-5,639.09
M7: β	$p = 0.029$ $q = 0.0167$ Average dN/dS for each branch = 0.6144 S = 1.56601	Not allowed	-5,696.39
M8: β and ω	$p = 0.16877$ $q = 0.11056$ $f_0 = 0.825$ $\omega_1 = 4.190$ ($f_1 = 0.175$) Average dN/dS for each branch = 1.2331 S = 1.62300	(47-51: gaps) 139 Q 0.9532* 175 Q 0.9963** 213 S 0.9659* 215 T 0.9975** 228 L 0.9967** 257 T 0.9607* 310 C 0.9945** 311 A 0.9749* 317 K 0.9805* 324 K 0.9995** 325 P 0.9920** (326-338: gaps) 339 F 0.9989** 340 V 0.9999** (385-386: gaps)	-5,641.64

Model	Parameter estimates	Sites* with $\omega > 1$	ℓ
		389 K 0.9949** 390 N 0.9587* 407 G 0.9966** 411 S 0.9653* 418 F 0.9995** 421 P 0.9832* 423 V 0.9756* 471 Q 0.9570* 483 G 0.9929**	

p and q are parameters of the β distribution. f is the proportion of sites assigned to an individual ω (= dN/dS) category or to a β distribution with shape parameters p and q . The proportion $f_{1,2}$ in parentheses is not a free parameter. Sites assigned to $\omega_{1,2}$ are those with posterior probabilities $P > 0.95$; those with $P > 0.99$ are in bold. Analyses were conducted by using κ as a free parameter. S = tree length.

Sites assigned to $\omega > 1$ are those with posterior probabilities $P > 0.95$ () or $P > 0.99$ (**).

3. Likelihood Ratio Test Statistics (2δ) for Models of Variable Selective Pressure Among Sites.

	2δ	df	P value
Model 0 (one dN/dS ratio for all branches)			
F3x4			
One ratio vs. M3 ($k = 3$)	216.26	4	$P < 0.0001$
M1 vs. M2	115.98	2	$P < 0.0001$
M7 vs. M8	115.82	2	$P < 0.0001$
F61			
One ratio vs. M3 ($k = 3$)	217.94	4	$P < 0.0001$
M1 vs. M2	109.88	2	$P < 0.0001$
M7 vs. M8	109.50	2	$P < 0.0001$

4. Likelihood Ratio Test Statistics (2δ) for Models of Variable Selective Pressure Along Branches.

	ℓ	2δ	df*	<i>P</i> value
TRIM5				
Model 0 (same dN/dS ratio for all branches)	-5,784.326			
Model 1 (different dN/dS ratio for each branch)	-5,755.534	57.584	37	$P < 0.02$
Apobec3G[†]				
Model 0 (same dN/dS ratio for all branches)	-4,209.048			
Model 1 (different dN/dS ratio for each branch)	-4,201.337	15.422	20	$P > 0.75$ (N.S.)

*df denotes degrees of freedom that is equal to one less than the total number of branches in the phylogeny. For a 20-taxa tree like in TRIM5, the total number of branches is 38, so df = 37. N.S., not significant.

[†]Sawyer, S. L., Emerman, M. & Malik, H. S. (2004) *PLoS Biol.* **2**, e275.

5. Likelihood Ratio Test Statistics for Models of Selection Along the Human and Gibbon Lineages (OWM + hominids only).

	ω (dN/dS) (human or gibbon branch)	ω (dN/dS) (remainder of tree)	ℓ	2δ	df*	<i>P</i> value
Human						
Model 0 (two-ratio, human $\omega = 1$)	1.00	1.132	-5,748.233			
Model 1 [two-ratio, human $\omega > 1$ (estimated by PAML; ref. 1)]	2.985	1.133	-5,747.543	1.38	1	0.24 (N.S)
Gibbon						
Model 0 (two-ratio, gibbon $\omega = 1$)	1.00	1.087	-5,748.718			
Model 1 [two-ratio, gibbon $\omega > 1$ (estimated by PAML; ref. 1)]	5.594	1.091	-5,744.345	8.746	1	0.0031

df denotes degrees of freedom; here, $df = 1$.

6. P Value Statistics for Positive Selection Along the Human and Gibbon Lineages, Based on Monte Carlo Simulation.

	Confidence ancestral sequence (0-1) per site	ω (dN/dS)	P value
Human vs. Ancestral Node	0.99971	3.6	$P < 10^{-4}$
Gibbon vs. Ancestral Node	0.99633	6.5	$P < 10^{-4}$

B. PAML (1) Analysis of Partial Data Set (hominids+OWM)

1. Log Likelihood Scores and Parameter Estimates for Four Models of Variable ω 's Among Sites Assuming the F3x4 Model of Codon Frequencies.

Model	Parameter estimates	Sites* with $\omega > 1$	ℓ
M0: one ratio	Average dN/dS for each branch = 1.1923		-3,519.244
Site models			
M1: neutral	$(\omega_0 = 0) f_0 = 0.307$ $(\omega_1 = 1) (f_1 = 0.693)$ Average dN/dS for each branch = 0.693	Not allowed	-3,513.813
M2: selection	$(\omega_0 = 0) f_0 = 0.185$ $(\omega_1 = 1) f_1 = 0.783$ $\omega_2 = 11.283 (f_2 = 0.032)$ Average dN/dS for each branch = 1.1429	330 G 0.9960** 332 R 0.9929** 335 R 0.9890* 337 Q 0.9981** 340 V 0.9993**	-3,492.320
M3:discrete	$\omega_0 0.00001 f_0 = 0.380$ $\omega_1 = 1.667 (f_1 = 0.601)$ $\omega_2 = 16.866 (f_2 = 0.019)$ average dN/dS for each branch = 1.3219	Too many to list. Note: M3 vs. M0 is a test of heterogeneity among sites and not an explicit test of positive selection.	-3,490.526
M7: β	$p = 0.00592 q = 0.00273$ Average dN/dS for each branch = 0.7	Not allowed	-3,513.818

Model	Parameter estimates	Sites* with $\omega > 1$	ℓ
M8: β and ω	$p = 0.01037$ $q = 0.00248$ $f_0 = 0.968$ $\omega_1 = 11.191$ ($f_1 = 0.032$) Average dN/dS for each branch = 1.1308	330 G 0.9959** 332 R 0.9927** 335 R 0.9888* 337 Q 0.9980** 340 V 0.9992**	-3,492.324

p and q are parameters of the β distribution. f is the proportion of sites assigned to an individual ω (= dN/dS) category or to a β distribution with shape parameters p and q . The proportion $f_{1,2}$ in parentheses is not a free parameter. Sites assigned to $\omega_{1,2}$ are those with posterior probabilities $P > 0.95$; those with $P > 0.99$ are in bold. Analyses were conducted using κ as a free parameter. S = tree length.

Sites assigned to $\omega > 1$ are those with posterior probabilities $P > 0.95$ () or $P > 0.99$ (**).

2. Log Likelihood Scores and Parameter Estimates for Four Models of Variable ω 's Among Sites Assuming the F61 Model of Codon Frequencies

Model	Parameter estimates	Sites* with $\omega > 1$	ℓ
M0: one ratio	Average dN/dS for each branch = 1.1375		-3,469.858
Site models			
M1: neutral	$(\omega_0 = 0) f_0 = 0.331$ $(\omega_1 = 1) (f_1 = 0.669)$ Average dN/dS for each branch = 0.669	Not allowed	-3,462.818
M2: selection	$(\omega_0 = 0) f_0 = 0.215$ $(\omega_1 = 1) f_1 = 0.749$ $\omega_2 = 10.019$ ($f_2 = 0.036$) Average dN/dS for each branch = 1.1057	330 G 0.9914** 332 R 0.9934** 335 R 0.9840* 337 Q 0.9984** 338 T 0.6807 340 V 0.9993**	-3,443.237
M3: discrete	$\omega_0 = 0.00001$ $f_0 = 0.415$ $\omega_1 = 1.761$ ($f_1 = 0.568$)	Too many to list. Note: M3 vs. M0 is a test	-3,441.147

Model	Parameter estimates	Sites* with $\omega > 1$	ℓ
	$\omega_2 = 17.069$ ($f_2 = 0.017$) Average dN/dS for each branch = 1.2835	of heterogeneity among sites and not an explicit test of positive selection.	
M7: β	$p = 0.00484$ $q = 0.00310$	Not allowed	-3,463.372
	Average dN/dS for each branch = 0.6		
M8: β and ω	$p = 0.00553$ $q = 0.00144$ $f_0 = 0.964$ $\omega_1 = 10.263$ ($f_1 = 0.036$) Average dN/dS for each branch = 1.2368	330 G 0.9919** 332 R 0.9938** 335 R 0.9849* 337 Q 0.9985** 340 V 0.9994**	-3,443.266

p and q are parameters of the β distribution. f is the proportion of sites assigned to an individual ω (= dN/dS) category or to a β distribution with shape parameters p and q . The proportion $f_{1,2}$ in parentheses is not a free parameter. Sites assigned to $\omega_{1,2}$ are those with posterior probabilities $P > 0.95$; those with $P > 0.99$ are in bold. Analyses were conducted using κ as a free parameter. S = tree length.

* Sites assigned to $\omega > 1$ are those with posterior probabilities $P > 0.95$ (*) or $P > 0.99$ (**).

3. Likelihood Ratio Test Statistics (2 δ) for Models of Variable Selective Pressure Among Sites.

	2 δ	df	P value
Model 0 (one ω for all branches)			
F3x4			
One ratio vs. M3 ($k = 3$)	57.436	4	$P < 0.0001$
M1 vs. M2	42.986	2	$P < 0.0001$
M7 vs. M8	42.988	2	$P < 0.0001$
F61			
One ratio vs. M3 ($k = 3$)	57.422	4	$P < 0.0001$
M1 vs. M2	39.162	2	$P < 0.0001$
M7 vs. M8	40.212	2	$P < 0.0001$

1. Yang, Z. (1997) *Comput. Appl. Biosci.* **13**, 555–556.

RING→

Human ATGGCTTCTGGAATCCTGGTTAATGTAAAGGAGGAGGTGACCTGCCCCATCTGCCTGGAA (20)
Chimp ATGGCTTCTGGAATCCTGGTTAATGTAAAGGAGGAGGTGACCTGCCCCATCTGCCTGGAA
Gorilla ATGGCTTCTGGAATCCTGGTTAATGTAAAGGAGGAGGTGACCTGCCCCATCTGCCTGGAA
Orangutan ATGGCTTCTGGAATCCTGGTTAATGTAAAGGAGGAGGTGACCTGCCCCATCTGCCTGGAA
Gibbon ATGGCTTCTGGAATCCTGGTTAATGTAAAGGAGAAGGTGACCTGCCCCATCTGCCTGGAA
Rhes_cDNA ATGGCTTCTGGAATCCTGCTTAATGTAAAGGAGGAGGTGACCTGTCCCATCTGCCTGGAA
Baboon ATGGCTTCTGGAATCCTGCTTAATGTAAAGGAGGAGGTGACCTGTCCCATCTGCCTGGAA
AGM ATGGCTTCTGGAATCCTGCTTAATGTAAAGGAGGAGGTGACCTGTCCCATCTGCCTGGAA
AGM_cDNA ATGGCTTCTGGAATCCTGGTTAATGTAAAGGAGGAGGTGACCTGTCCCATCTGCCTGGAA
Tant_cDNA ATGGCTTCTGGAATCCTGCTTAATGTAAAGGAGGAGGTGACCTGTCCCATCTGCCTGGAA
Patas ATGGCTTCTGGAATCCTGCTTAATGTAAAGGAGGAGGTGACCTGTCCATCTGCCTGGAA
Colobus ATGGCTTCTGGAATCCTGGTTAATATAAAGGAGGAGGTGACCTGCCCCATCTGCCTGGAA
DLangur ATGGCTTCTGGAATCCTGGTTAATATAAAGGAGGAGGTGACCTGCCCCATCTGCCTGGAA
PMarmoset ATGGCTTCCAGAATCCTGGTGAATATAAAGGAGGAGGTAACCTGCCCCATCTGCCTGGAA
Tamarin ATGGCTTCCAGAATCCTGGTGAATATAAAGGAGGAGGTGACCTGCCCCATCTGCCTGGAA
Squirrel ATGGCTTCCAGAATCCTGGTGAATATAAAGGAGGAGGTGACCTGCCCCATCTGCCTGGAA
Owl ATGGCTTCCAGAATCCTGGTCAATATAAAGGAGGAGGTGACCTGCCCCATCTGCCTGGAA
Titi ATGGCTTCCAGAATCCTGGTGAATATAAAGGAGGAGGTGACCTGCCCCATCTGCCTGGAA
Saki ATGGCTTCCAGAATCCTGATGAACATAAAGGAGGAGGTGACCTGCCCCATCTGCCTGGAA
Howler ATGGCTTCCAAAATCCTGGTGAATATAAAGGAGGAGGTGACCTGCCCCATCTGCCTGGAA
Spider ATGGCTTCCGAAATCCTGTTGAATATAAAGGAGGAGGTGACCTGCCCCATCTGCCTGGAA
Woolly ATGGCTTCCGAAATCCTGGTGAATATAAAGGAGGAGGTGACCTGCCCCATCTGCCTGGAC

Human CTCCTGACACAACCCCTGAGCCTGGACTGCGGCCACAGCTTCTGCCAAGCATGCCTCACT (40)
Chimp CTCCTGACACAACCCCTGAGCCTGGACTGCGGCCACAGCTTCTGCCAAGCATGCCTCACT
Gorilla CTCCTGACACAACCCCTGAGCCTGGACTGCGGCCACAGCTTCTGCCAAGCATGCCTCACT
Orangutan CTCCTGACACAACCCCTGAGTCTGGACTGCGGCCACAGCTTCTGCCAAGCATGCCTCACT
Gibbon CTCCTGACACAACCCCTGAGTCTGGACTGCGGCCACAGCTTCTGCCAAGCATGCCTCACT
Rhes_cDNA CTCCTGACAGAACCCCTGAGTCTGCACTGCGGCCACAGCTTCTGCCAAGCGTGCATCACT
Baboon CTCCTGACAGAACCCCTGAGTCTGCCCTGTGGCCACAGCTTCTGCCAAGCGTGCATCACT
AGM CTCCTGACAGAACCCCTGAGTCTGCCCTGCGGCCACAGCTTCTGCCAAGCGTGCATCACT
AGM_cDNA CTCCTGACAGAACCCCTGAGTCTGCCCTGCGGCCACAGCTTCTGCCAAGCGTGCATCACT
Tant_cDNA CTCCTGACAGAACCCCTGAGTCTGCCCTGCGGCCACAGCTTCTGCCAAGCGTGCATCACT
Patas CTCCTGACAGAACCCCTGAGTCTGCCCTGCGGCCACAGCTTCTGCCAAGCGTGCATCACT
Colobus CTCCTGACAGAACCCCTGAGTCTGCACTGCGGCCACAGCTTCTGCCAAGCGTGCATCACT
DLangur CTCCTGACAGAACCCCTGAGTCTGCACTGCGGCCACAGCTTCTGCCAAGCGTGCATCACT
PMarmoset CTCCTGACAGAACCTCTGAGCCTAGACTGTGGCCACAGCTTCTGCCAAGCCTGCATCACT
Tamarin CTCCTGACAGAACCTCTGAGCCTAGACTGTGGCCACAGCTTCTGCCAAGCATGCATCACT
Squirrel CTCCTGACAGAACCTCTGAGCCTAGACTGTGGCCACAGCTTCTGCCAAGCATGCATCACT
Owl CTCCTGACAGAACCCCTGAGCCTGGACTGTGGCCATAGCTTCTGCCAAGCATGCATCACT
Titi CTCCTGACAGAACCCCTGAGCCTAGACTGTGGCCACAGCTTCTGCCAAGCATGCATCACC
Saki CTCCTGACAGAACCCCTGAGCCTAGACTGTGGCCACAGCTTCTGCCAAGCATGCATCACT
Howler CTA CTGACAGAACCTCTGAGCCTAGACTGTGGCCACAGCTTCTGCCAAGCATGCATCACT
Spider CTA CTGACAGAACCTCTGAGCCTAGACTGTGGCCACAGCTTCTGCCAAGCATGCATCACT
Woolly CTA CTGACAGAACCTCTGAGCCTAGACTGTGGCCACAGCTTCTGCCAAGCATGCATCACT

Human GCAAACCACAAGAAGTCCATGCTAGACAAAGGA---GAGAGTAGCTGCCCTGTGTGCCGG (59)
Chimp GCAAACCACAAGAAGTCCATGCTAGACAAAGGA---GAGAGTAGCTGCCCTGTGTGCCGG
Gorilla GCAAACCACAAGAAGTCCATGCTAGACAAAGGA---GAGAGTAGCTGCCCTGTGTGCCGG
Orangutan GCAAACCACAAGAAGTCCACGCTAGACAAAGGA---GAGAGAAGCTGCCCTGTGTGCCGG
Gibbon GCAAACCACAAAACGTCCATGCCAGACGAAGGA---GAGAGAAGCTGCCCTGTGTGCCGG
Rhes_cDNA GCGAACCACAAGAAGTCCATGCTATACAAAGAAGGAGAGAGAAGCTGCCCTGTGTGCCGG
Baboon GCAAACCACAGGAAGTCCATGCTATACAAAGAAGGAGAGAGAAGCTGCCCTGTGTGCCGG
AGM GCAAACCACAAGGAGTCCATGCTATACAAAGAAGAAGAGAGAAGCTGCCCTGTGTGCCGG
AGM_cDNA GCAAACCACAAGGAGTCCATGCTATACAAAGAAGAAGAGAGAAGCTGCCCTGTGTGCCGG
Tant_cDNA GCAAACCACAAGGAGTCCATGCTATACAAAGAAGAAAGAGAGAAGCTGCCCTGTGTGCCGG
Patas GCAAACCACAAGAAGTCCATGCTATACAAAGAAGAAAGAGAGAAGCTGCCCTGTGTGCCGG
Colobus GCAAACCACAAGAAGTCCATGCTATACAAAGAAGGAGAGAGAAGCTGCCCTGTGTGCCGG
DLangur GCAAACCACAAGAAGTCCATGCTATACAAAGAAGGAGAGAGAAGCTGCCCTGTGTGCCGG
PMarmoset GCAAACCACAAAAGAGTCTACGCCACACCAAGGA---GAGAGAAGCTGCCCTTTGTGCCGG
Tamarin GCAAACCACAAAAGAGTCTACGCCACACCAAGGA---GAGAGAAGCTGCCCTTTGTGCCGG
Squirrel GCAAATCACAAAAGAGTCTATGCTACACCAAGGA---GAGAGAAGCTGCCCTTTGTGCCGG
Owl GCAAATCACAAAAAGTCTATGCCACACCAAGGA---GAGAGAAGCTGCCCTTTGTGCCGG
Titi GCAAACCACAAAAGAGTCTACGCTACACCAAGGA---GAGAGAAGCTGCCCTTTGTGCCGG
Saki GCAAACCACAAAAGAGTCTATGCTACACCAAGGA---GAGAGAAGCTGCCCTTTGTGCCGG
Howler GCAAACCACAAAAGAGTCC-----AGA---GAGAGAAGCTGCCCTTTGTGCCGG
Spider GCAAACCACAAAAGAGTCTACGCTACACCAAGGA---GAGAGAAGCTGCCCTTTGTGCCGG
Woolly GCAGACCACAAAAGAGTCTACGCTACACCAAGGA---GAGAGAAGCTGCCCTTTGTGCCGG

Human ATCAGTTACCAGCCTGAGAACATACGGCCTAATCGGCATGTAGCCAACATAGTGGAGAAG (79)
Chimp ATCAGTTACCAGCCTGAGAACATACGGCCTAATCGGCATGTAGCCAACATAGTGGAGAAG
Gorilla ATCAGTTACCAGCCTGAGAACATACGGCCTAATCGGCATGTAGCCAACATAGTGGAGAAG
Orangutan GTCAGTTACCAGCCTAAGAACATACGGCCTAATCGGCATGTAGCCAACATAGTGGAGAAG
Gibbon ATCAGTTACCAGCATAAGAACATACGGCCTAATCGGCATGTAGCCAACATAGTGGAGAAG
Rhes_cDNA ATCAGTTACCAGCCTGAGAACATACAGCCTAATCGGCATGTAGCCAACATAGTGGAGAAG
Baboon ATCAGTTACCAGCCTGAGAACATACAGCCTAATCGGCATGTAGCCAACATAGTGGAGAAG
AGM ATCAGTTACCAGCCTGAGAATATACAGCCTAATCGGCATGTAGCCAACATAGTGGAGAAG
AGM_cDNA ATCAGTTACCAGCCTGAGAATATACAGCCTAATCGGCATGTAGCCAACATAGTGGAGAAG
Tant_cDNA ATCAGTTACCAGCCTGAGAATATACAGCCTAATCGGCATGTAGCCAACATAGTGGAGAAG
Patas ATCAGTTACCAGCCTGAGAACATACAGCCTAATCGGCATGTAGCCAACATAGTGGAGAAG
Colobus ATCAGTTACCAGCCTGAGAACATACGGCCTAATCGGCATGTGGCCAACATAGTGGAGAAG
DLangur ATCAGTTACCAGCCTGAGAACATACGGCCTAATCGGCATGTGGCCAACATAGTGGAGAAG
PMarmoset ATGAGTTACCCGTCTGAGAACTTGCGGCCTAATCGGCATTTGGCCAATATAGTGGAGAGG
Tamarin ATGAGTTACCCGTCTGAGAACTTGCGGCCTAATCGGCATTTGGCCAACATAGTGGAGAGG
Squirrel CTCCCTTACCAGTCTGAGAACCTGCGGCCTAATCGGCATTTGGCCAGCATCGTGGAGAGG
Owl ATCAGTTACTCGTCTGAGAACCTGCGGCCTAATCGGCATTTGGTCAACATAGTGGAGAGG
Titi ATCAGTTACCCGTCTGAGAACCTGCGGCCTAATCGGCATTTGGCCAACATAGTGGAGAGG
Saki ATCAGTTACCCATCTGAGAACCTGCGGCCTAATCGGCATTTGGCCAACATAGTGGAGAGG
Howler GTCAGTTACCACTCTGAGAACCTGCGGCCTAATCGGCATTTGGCCAACATAGCGGAGAGG
Spider GTCAGTTACCAGTCTGAGAACCTGCGGCCTAATCGGCATTTGGCAAACATAGCGGAGAGG
Woolly GTCGGTTACCAGTCTGAGAACCTGCGGCCTAATCGGCATTTGGCAAACATAGCCGAGAGG

B-box2→

Human CTCAGGGAGGTCAAGTTGAGCCCA---GAGGGGCAGAAAGTTGATCATTGTGCACGCCAT (98)
Chimp CTCAGGGAGGTCAAGTTGAGCCCA---GAGGGGCAGAAAGTTGATCATTGTGCACACCCAT
Gorilla CTTAGGGAGGTCAAGTTGAGCCCA---GAGGGGCAGAAAGTTGATCATTGTGCACGCCAT
Orangutan CTCAGGGAGGTCAAATTTGAGCCCA---GAGGGGCAGAAAGTTGATCACTGTGCACGCCAT
Gibbon CTCAGGGAGGTCAAGTTGAGCCCA**GAG**GAGGGGCAGAAAGTTGATCACTGTGCACGCCAC
Rhes_cDNA CTCAGGGAGGTCAAGTTGAGCCCA**GAA**GAGGGACAGAAAGTTGATCACTGTGCACGCCAT
Baboon CTCAGGGAGGTCAAGTTGAGCCCA**GAA**GAGGGGCTGAAGTTGATCACTGTGCACGCCAT
AGM CTCAGAGAGGTCAAGTTGAGCCCA**GAA**GAGGGGCAGAAAGTTGATCACTGTGCACGCCAT
AGM_cDNA CTCAGAGAGGTCAAGTTGAGCCCA**GAA**GAGGGGCAGAAAGTTGATCACTGTGCACGCCAT
Tant_cDNA CTCAGAGAGGTCAAGTTGAGCCCA**GAA**GAGGGGCAGAAAGTTGATCACTGTGCACGCCAT
Patas CTCAGAGAGGTCAAGTTGAGCCCA**GAA**GAGGGGCAGAAAGTTGATCACTGTGCACGCCAT
Colobus CTCAGGGAGGTCAAGTTGAGCCCA**GAA**GAGGGGCAGAAAGTTGATCACTGTGCACGCCAT
DLangur CTCAGGGAGGTCAAGTTGAGCCCA**GAA**GAGGGGCAGAAAGTTGATCACTGTGCACGCCAT
PMarmoset CTCAAAGAGGTCATGCTGAGCCCA**GAG**GAGGGGCAGAAAGTTGATCACTGTGCACGCCAT
Tamarin CTCAAAGAGGTCATGCTGAGCCCA**GAG**GAGGGGCAGAAAGTTGGTCACTGTGCACGCCAT
Squirrel CTCAGGGAGGTCATGCTGAGACCA**GAG**GAAAGGCAGAAACGTTGATCACTGTGCACGCCAT
Owl CTCAGGGAGGTCATGCTGAGCCCA**GAG**GAGGGGCAGAAAGTTGATCACTGTGCACACCCAT
Titi CTCAGGGAGGTCGCTGCTGAGCCCA**GAG**GAGGGGCAGAAAGTTGATCTCTGTGCACGCCAT
Saki CTCAGGGAGGTCATGCTGAGCCCA**GAG**GAGGGGCAGAAAGTTGATCACTGTGCACGCCAT
Howler CTCAGGGAGGTCATGTTGAGCCCA**GAG**GAGGGGCAGAAAGTTGATCGCTGTGCACGCCAT
Spider CTCAGGGAGGTCATGTTGAGCCCA**GAG**GAAAGGCAGAAAGTTGATCGCTGTGCACGCCAT
Woolly CTCAGGGAGGTCATGTTGAGCCCA**GAG**GAAAGGCAGAAAGTTGATCGCTGTGCACGCCAT

Human GGAGAGAAACTTCTACTCTTCTGTCAGGAGGACGGGAAGGTCATTTGCTGGCTTTGTGAG (118)
Chimp GGAGAGAAACTTCTACTCTTCTGTCAGGAGGACGGGAAGGTCATTTGCTGGCTTTGTGAG
Gorilla GGAGAGAAACTTCTACTCTTCTGTCAGGAGGACGGGAAGGTCATTTGCTGGCTTTGCGAG
Orangutan GGAGAGAAACTTCTACTCTTCTGTAAGGAGGACGGGAAGGTCATTTGCTGGCTTTGTGAG
Gibbon GGAAAGAAACTTCTACTCTTCTGTCAGGAGGACAGGAAGGTCATTTGCTGGCTTTGTGAG
Rhes_cDNA GGAGAGAAACTCCTACTCTTCTGTCAGGAGGACAGCAAGGTCATTTGCTGGCTTTGTGAG
Baboon GGAGAGAAACTCCTACTCTTCTGTCAGGAGGACAGCAAGGTCATTTGCTGGCTTTGTGAG
AGM GGAGAGAAACTCCTACTCTTCTGTCAGGAGGACAGCAAGGTCATTTGCTGGCTTTGTGAG
AGM_cDNA GGAGAGAAACTCCTACTCTTCTGTCAGGAGGACAGCAAGGTCATTTGCTGGCTTTGTGAG
Tant_cDNA GGAGAGAAACTCCTACTCTTCTGTCAGGAGGACAGCAAGGTCATTTGCTGGCTTTGTGAG
Patas GGAGAGAAACTCCTACTCTTCTGTCAGGAGGACAGGAAGGTCATTTGCTGGCTTTGTGAG
Colobus GGAGAGAAACTCCTACTCTTCTGTCAGGAGGACAGGAAGGTCATTTGCTGGCTTTGTGAG
DLangur GGAGAGAAACTCCTACTCTTCTGTCAGGAGGACAGGAAGGTCATTTGCTGGCTTTGTGAG
PMarmoset GGAGAGAAACTTCTACTCTTCTGTCAGCAGGATGGAAATGTCATTTGCTGGCTTTGTGAG
Tamarin GGAGAGAAACTTCTACTCTTCTGTGAGCAGGATGGAAATGTCATTTGCTGGCTTTGTGAG
Squirrel GGAGAGAAACTTCTACTCTTCTGTGAGCAGGATGGAAATATCATTGCTGGCTTTGTGAG
Owl GGAGAGAAACTTGTACTCTTCTGTCAGCAGGATGGAAATGTCATTTGCTGGCTTTGTGAG
Titi GGAGAGAAACTTCTACTCTTCTGTCAGCAGGATGGAAATGTCATTTGCTGGCTTTGTGAG
Saki GGAGAGAAACTTCTACTCTTCTGTCAGCAGGATGGAAATGTCATTTGCTGGCTTTGTGAG
Howler GGAGAGAAACTTCTACTCTTCTGTCAGCAGCATGGAAATGTCATTTGCTGGCTTTGTGAG
Spider GGAGAGAAACTTCTACTCTTCTGTCAGCAGCATGGAAATGTCATTTGCTGGCTTTGTGAG
Woolly GGAGAGAAACTTCTACTCTTCTGTCAGCAGCATGGAAATGTCATTTGCTGGCTTTGTGAG

Coiled-Coil→

Human	<u>CGGTCTCAGGAGCACCGTGGTCACCAC</u> ACGTTCTCACAGAGGAGGTTGCCCGGGAGTAC (138)
Chimp	CGGTCTCAGGAGCACCGTGGTCACCACACGTTCCCTCACAGAGGAGGTTGCCCGGGAGTAC
Gorilla	CGGTCTCAGGAGCACCGTGGTCACCACACGTTCCCTCACAGAGGAGGTTGCCCGGGAGTAC
Orangutan	CGGTCTCAGGAGCACCGTGGTCACCACACATTCCCTCACGGAGGAGGTTGCCCGAGAGTAC
Gibbon	CGGTCTCAGGAGCACCGTGGTCACCACACATTCCCTCACGGAGGAGGTTGCCCGAGGAGTAC
Rhes_cDNA	CGGTCTCAGGAGCACCGTGGTCACCACACTTTCCCTCATGGAGGAGGTTGCCCGAGGAGTAC
Baboon	CGGTCTCAGGAGCACCGTGGTCACCACACTTTCCCTCATGGAGGAGGTTGCCCGAGGAGTAC
AGM	CGGTCTCAGGAGCACCGTGGTCACCACACTTTCCCTCATGGAGGAGGTTGCCCGAGGAGTAC
AGM_cDNA	CGGTCTCAGGAGCACCGTGGTCACCACACTTTCCCTCATGGAGGAGGTTGCCCGAGGAGTAC
Tant_cDNA	CGGTCTCAGGAGCACCGTGGTCACCACACTTTCCCTCATGGAGGAGGTTGCCCGAGGAGTAC
Patas	CGGTCTCAGGAGCACCGTGGTCACCACACTTTCCCTCATGGAGGAGGTTGCCCGAGGAGTAC
Colobus	CGGTCTCAGGAGCACCGTGGTCACCACACGTTCCCTCATGGAGGAGGTTGCCCGAGGAGTAC
DLangur	CGGTCTCAGGAGCACCGTGGTCACCACACGTTCCCTCATGGAGGAAGTTGCCCGAGGAGTAC
PMarmoset	CGGTCTCAAGAACACCGTGGTCACCACACATTCCCTCGTGGAGGAGGTTGCAGAGAAATAC
Tamarin	CGGTCTCAAGAACATCGTGGTCACCACACATTACTCGTGGAGGAGGTTGCAGAGAAATAC
Squirrel	CGGTCTCAAGAACACCGTGGTCACAACACATTCCCTCGTGGAGGAGGTTGCACAGAAATAC
Owl	CGGTCTCAAGAACACCGTGGGCACCAGACATTCCCTTGTGGAGGAGGTTGCACAGAAATAC
Titi	CGGTCTCAAGAACACCGTGGTCACCACACATTCCCTCGTGGAGGAGGTTGCACAGACATAC
Saki	CGGTCTCAAGAACACCGTGGTCACCACACATTACTCGTGGAGGAGGTTGCACAGACATAC
Howler	CGGTCTGAAGAACACCGTGGTCACCGCACATCCCTCGTGGAGGAGGTTGCACAGAAATAC
Spider	CGGTCTCAAGAACACCGTGGTCACAGCACATTCCCTCGTGGAGGAGGTTGCACAGAAATAC
Woolly	CGGTCTCAAGAACACCGTGGTCACAGCACATTCCCTCGTGGAGGAGGTTGCACAGAAATAC

Exon3→

Human	<u>CAA</u> GTGAAGCTCCAGGCAGCTCTGGAGATGCTGAGGCAGAAGCAGCAGGAAGCTGAAGAG (158)
Chimp	CAA GTGAAGCTCCAGGCAGCTCTGGAGATGCTGAGGCAGAAGCAGCAGGAAGCTGAAGAG
Gorilla	CAA GTGAAGCTCCAGGCAGCTCTGGAGATGCTGAGGCAGAAGCAGCAGGAAGCTGAAGAG
Orangutan	CAA GTGAAGCTCCAGGCAGCTCTGGAGATGCTGAGGCAGAAGCAGCAGGAAGCTGAAGAG
Gibbon	CAA ATGAAGCTCCAGGCAGCTCTGCAGATGCTGAGGCAGAAGCAGCAGGAAGCTGAAGAG
Rhes_cDNA	CAT GTGAAGCTCCAGACAGCTCTGGAGATGCTGAGGCAGAAGCAGCAGGAAGCTGAAAAG
Baboon	CAT GTGAAGCTCCAGACAGCTCTGGAGATGCTGAGGCAGAAGCAGCAGGAAGCTGAAAAG
AGM	CAT GTGAAGCTCCAGACAGCTCTGGAGATGCTGAGGCAGAAGCAGCAGGAAGCTGAAAAG
AGM_cDNA	CAT GTGAAGCTCCAGACAGCTCTGGAGATGCTGAGGCAGAAGCAGCAGGAAGCTGAAAAG
Tant_cDNA	CAT GTGAAGCTCCAGACAGCTCTGGAGATGCTGAGGCAGAAGCAGCAGGAAGCTGAAAAG
Patas	CAT GTGAAGCTCCAGACAGCTCTGGAGATGCTGAGGCAGAAGCAGCAGGAAGCTGAAAAG
Colobus	CAC GTGAAGCTCCAGACAGCTCTGGAGATGCTGAGGCAGAAGCAGCAGGAAGCTGAAAAG
DLangur	CAC GTGAAGCTCCAGACAGCTCTGGAGATGCTGAGGCAGAAGCAGCAGGAAGCTGAAAAG
PMarmoset	CAA GGAAAGCTCCAGGTAGCTCTGGAGATGATGAGGCAGAAGCAGCAGGATGCTGAAAAG
Tamarin	CAA GAAAAGCTCCAGGTAGCTCTGGAGATGATGAGGCAGAAGCAGCAGGATGCTGAAAAG
Squirrel	CGA GAAAAGCTCCAGGTAGCTCTGGAGACAATGAGGCAGAAGCAGCAGGATGCTGAAAAG
Owl	CGA GAAAAGCTCCAGGTAGCTCTGGAGATGATGAGGCAGAAGCAGAAGGATGCTGAAAAG
Titi	CGA GAAAATCTCCAGGTAGTTCTGGAGATGATGAGGCAGAAGCATCAGGATGCTGAAAAG
Saki	CGA GAAAATCTCCAGGTAGCTCTGGAGACGATGAGGCAGAAGCAGCAGGATGCTGAAAAG
Howler	CGA GAAAAGCTCCAGGCAGCTCTGGAGATGATGAGGCAGAAGGAGCAGGATGCTGAAAATG
Spider	CAA GAAAAGCTCCAGGTAGCTCTGGAGATGATGAGGCAGAAGCAGCAGGATGCTGAAAAG
Woolly	CGA GAAAAGCTCCAGGTAGCTCTGGAAATGATGAGGGAGAAGCAGCAGGATGCTGAAAAG

Exon4→

Human	TTGGAAGCTGACATCAGAGAAGAGAAAGCTTCCTGGAAG	ACTCAAATACAGTATGACAAA	(178)
Chimp	TTGGAAGCTGACATCAGAGAAGAGAAAGCTTCCTGGAAG	ACTCAAATACAGTATGACAAA	
Gorilla	TTGGAAGCTGACATCAGAGAAGAGAAAGCTTCCTGGAAG	ACTCAAATACAGTATGACAAA	
Orangutan	TTGGAAGCTGACATCAGAGAAGAGAAAGCTTCCTGGAAG	ACTCAAATACAGTATGACAAA	
Gibbon	TTGGAAGCTGACATCAGAGAAGAGAAAGCTTCCTGGAAG	ACTCAAATACAGTATGACAAA	
Rhes_cDNA	TTGGAAGCTGACATCAGAGAAGAGAAAGCTTCCTGGAAG	ATTCAAATACAGTACGACAAA	
Baboon	TTGGAAGCTGACATCAGAGAAGAGAAAGCTTCCTGGAAG	ATTCAAATACAGTACGACAAA	
AGM	TTGGAAGCTGACATCAGAGAAGAGAAAGCTTCCTGGAAG	ATTCAAATACAGTACGACAAA	
AGM_cDNA	TTGGAAGCTGACATCAGAGAAGAGAAAGCTTCCTGGAAG	ATTCAAATACAGTACGACAAA	
Tant_cDNA	TTGGAAGCTGACATCAGAGAAGAGAAAGCTTCCTGGAAG	ATTCAAATACAGTACGACAAA	
Patas	TTGGAAGCTGACATCAGAGAAGAGAAAGCTTCCTGGAAG	ATTCAAATACAGTACGACAAA	
Colobus	TTGGAAGCTGACATCAGAGAAGAGAAAGCTTCCTGGAAG	ATTCAAATACAGTATGACAAA	
DLangur	TTGGAAGCTGACATCAGAGAAGAGAAAGCTTCCTGGAAG	ATTCAAATACAGTGCACAAA	
PMarmoset	TTAGAAGCTGATGTCAGAGAAGAGCAAGCTTCCTGGAAG	ATTCAAATACAAAATGACAAA	
Tamarin	TTGGAAGCTGACGTCAGAGAAGAGCAAGCTTCTTGAAG	ATTCAAATACGAAATGACAAA	
Squirrel	TTGGAAGCTGACGTCAGAGAAGAGCAAGCTTCCTGGAAG	ATTCAAATACAAAATGACAAA	
Owl	TTGGAAGCTGACGTCAGAGAAGAGCAAGCTTCCTGGAAG	ATTCAAATACAAAATGACAAA	
Titi	TTGGAAGCTGACGTCAGAGAAGAGCAAGCTTCCTGGAAG	ATTCAAATACAAAATGACAAA	
Saki	TTAGAAGCTGACGTCAGAGAAGAGCAAGCTTCCTGGAAG	ATTCAAATACGAGATGACAAA	
Howler	TTGGAAGCTGACGTCAGAGAAGAGCAAGCTTCCTGGAAG	ATTCAAATAGAAAATGACAAA	
Spider	TTGGAAGCTGATGTCAGAGAAGAGCAAGCTTCCTGGAAG	ATTCAAATAGAAAATGACAAA	
Woolly	TTGGAAGCTGATGTCAGAGAAGAGCAAGCTTCCTGGAAG	ATTCAAATAAAAACGACAAA	

Human	ACCAACGTCTTGGCAGATTTTGAGCAACTGAGAGACATCCTGGACTGGGAGGAGAGCAAT	(198)
Chimp	ACCAACGTCTTGGCAGATTTTGAGCAACTGAGAGACATCCTGGACTGGGAGGAGAGCAAT	
Gorilla	ACCAACGTCTTGGCAGATTTTGAGCAACTGAGAGACATCCTGGACTGGGAGGAGAGCAAT	
Orangutan	ACCAGCGTCTTGGCAGATTTTGAGCAACTGAGAGACATCCTGGACTGGGAGGAGAGCAAT	
Gibbon	ACCAACATCTTGGCAGATTTTGAGCAACTGAGACACATCCTGGACTGGGTGGAGAGCAAT	
Rhes_cDNA	ACCAACGTCTCGGCAGATTTTGAGCAACTGAGAGAGATCCTGGACTGGGAGGAGAGCAAT	
Baboon	ACCAACGTCTCGGCAGATTTTGAGCAACTGAGAGAGATCCTGGACTGGGAGGAGAGCAAT	
AGM	ACCAACGTCTCGGCAGATTTTGAGCAACTGAGAGAGATCCTGGACTGGGAGGAGAGCAAT	
AGM_cDNA	ACCAACGTCTCGGCAGATTTTGAGCAACTGAGAGAGATCCTGGACTGGGAGGAGAGCAAT	
Tant_cDNA	ACCAACGTCTCGGCAGATTTTGAGCAACTGAGAGAGATCCTGGACTGGGAGGAGAGCAAT	
Patas	ACCAACGTCTTGGCAGATTTTGAGCAACTGAGAGAGATCCTGGACTGGGAGGAGAGCAAT	
Colobus	ACCAACGTCTTGGCAGATTTTGAGCAACTGAGAGAGATCCTGGACTGGGAGGAGAGCAAT	
DLangur	ACCAATGTCTTGGCAGATTTTGAGCAACTGAGAGAGATCCTGGACTGGGAGGAGAGCAAT	
PMarmoset	ACCAACATCATGGCAGAGTTTAAGCAACTGAGAGACATCCTGGACTGTGAGGAGAGCAAA	
Tamarin	ACCAACATCATGGCAGAGTTTAAGCAACTGAGAGACATCCTGGACTGTGAGGAGAGCAAA	
Squirrel	ACCAACATCATGGCAGAGTTTAAGCAACTGAGAGACATCCTGGACTGTGAGGAGAGCAAT	
Owl	ACCAACATCATGGCAGAGTTTAAAAACGGAGAGACATCCTGGACTGTGAGGAGAGCAAA	
Titi	ACCAACATCATGGCAGAGTTTAAGCAACTGAGAGACATCCTGGACTGTGAGGAGAGCAAT	
Saki	ACCAACATTATGGCAGAGTTTAAGCAACTGAGAGACATCCTGGACTGTGAGGAGAGCAAT	
Howler	ACCAGCACCTGGCAGAGTTTAAGCAACTGAGAGACATCCTGGACTGTGAGGAGAGCAAC	
Spider	ACCAACATCCTGGCAGAGTTTAAGCAACTGAGAGACATCCTGGACTGTGAGGAGAGCAAT	
Woolly	ACCAACATCCTGGCAGAGTTTAAGCAACTGAGAGACATCCTGGACTGTGAGGAGAGCAAT	

Human GAGCTGCAAAACCTGGAGAAGGAGGAGGAAGACATTCTGAAAAGCCTTACGAACTCTGAA (218)
Chimp GAGCTGCAAAACCTGGAGAAGGAGGAGGAAGACATTCTGAAAAGCCTTACGAAGTCTGAA
Gorilla GAGCTGCAAAACCTGGAGAAGGAGGAGGAAGACATTCTGAAAAGCCTTACGAAGTCTGAA
Orangutan GAGCTGCAAAACCTGGAGAAGGAGGAGGAAGACATTCTAAAAAGCCTTACGAAGTCTGAA
Gibbon GAGCTGCAAAACCTGGAGAAGGAGGAGAAAAGACGTTCTGAAAAGGCTTATGAGGTCTGAA
Rhes_cDNA GAGCTGCAGAACCTGGAGAAGGAGGAAGAAGACATTCTGAAAAGCCTTACGAAGTCTGAA
Baboon GAGCTGCAGAACCTGGAGAAGGAGGAAGAAGACATTCTGAAAAGCCTTACGAAGTCTGAA
AGM GAGCTGCAGAACCTGGAGAAGGAGGAAGAAGACATTCTGAAAAGCCTTACGAAGTCTGAA
AGM_cDNA GAGCTGCAGAACCTGGAGAAGGAGGAAGAAGACATTCTGAAAAGCCTTACGAAGTCTGAA
Tant_cDNA GAGCTGCAGAACCTGGAGAAGGAGGAAGAAGACATTCTGAAAAGCCTTACGAAGTCTGAA
Patas GAGCTGCAGTACCTGGAGAAGGAGGAAGAAGACATTCTGAAAAGCCTTACGAAGTCTGAA
Colobus GAGCTGCAGAACCTGGAGAAGGAGGAGGAAGACATTCTGAAAAGCCTTACGAAGTCTGAA
DLangur GAGCTGCAGAACCTGGAGAAGGAGGAGGAAGACATTCTGAAAAGCCTTACGAAGTCTGAA
PMarmoset GAGCTGCAAAACCTGGAGAAGGAGGAGAAAACATTCTGAAAAGACTTGTACAGTCTGAA
Tamarin GAGCTGCAAAACCTGGAGAAGGAGGAGAAAACATTCTGAAAAGACTTGTACAGTCTGAA
Squirrel GAGCTGCAAAACCTGGAGAAGGAGGAGAAAACATTCTGAAAAGACTTGTACAGTCTGAA
Owl GAGTTGCAAAACCTGGAGAAGGAGGAGAAAACATTCTGAAAAGACTTGTACAGTCTGAA
Titi GAGCTGCAAAACCTAGAGAAGGAGGAGAAAACATTCTGAAAAGACTTGTACAGTCTGAG
Saki GAGCTGCAAAATCCTAGAGAAGGAGGAGAAAACATTCTGAAAAGACTTACAAGTCTGAA
Howler GAGCTGCAAAAACCTGGAGAAGGAGGAGAAAACCTTCTGAAAAGACTTGTACAGTCTGAA
Spider GAGCTACAAAACCTGGAGAAGGAGGAGAAAACCTTCTGAAAACACTTGCACAGTCTGAA
Woolly GAGCTGCAAAACCTGGAGAAGGAGGAGAAAACCTTCTGAAAATACTTGCACAGTCTGAA

Human ACTGAGATGGTGCAGCAGACCCAGTCCCTGAGAGAGCTCATCTCAGATCTGGAGCATCGG (238)
Chimp ACTGAGATGGTGCAGCAGACCCAGTCCCTGAGAGAGCTCATCTCAGATCTGGAGCGTCGG
Gorilla ACTGAGATGGTGCAGCAGACCCAGTCCCTGAGAGAGCTCATCTCAGATCTGGAGCATCGG
Orangutan ACTGAGATGGTGCAGCAGACCCAGTCCCTGAGAGAGCTCATCTCAGATGTGGAGCATCGG
Gibbon ATTGAGATGGTGCAGCAGACCCAGTCCCTGAGAGAGCTCATCTCAGATCTGGAGCATCGG
Rhes_cDNA ACGGAGATGGTGCAGCAGACCCAGTACATGAGAGAGCTCATCTCAGAAGTGGAGCATCGG
Baboon ACGGAGATGGTGCAGCAGACCCAGTACATGAGAGAGCTCATCTCAGATCTGGAGCATCGG
AGM ACGGAGATGGTGCAGCAGACCCAGTACATGAGAGAGCTCATCTCAGATCTGGAGCATCGG
AGM_cDNA ACGGAGATGGTGCAGCAGACCCAGTACATGAGAGAGCTCATCTCAGATCTGGAGCATCGG
Tant_cDNA ACGGAGATGGTGCAGCAGACCCAGTACATGAGAGAGCTCATCTCAGATCTGGAGCATCGG
Patas ACGAAGATGGTGCAGCAGACCCAGTACCTGAGAGAGCTCATCTCAGATCTGGAGCATCGG
Colobus ACTGAGATGGTGCAGCAGACCCAGTACATGAGAGAGCTCGTCTCAGATCTGGAGCATCGG
DLangur ACTGAGATGGTGCAGCAGACCCAGTACATGAGAGAGCTCATCTCAGATCTGGAGCATCGG
PMarmoset AGTGACATGGTGTGCTGCAGACCCAGTCCATTAGAGTGCTCATCTCAGATCTGGAGCGTCGC
Tamarin AGTGACATGGTGTGCTGCAGACCCAGTCCATGAGAGTGCTCATCTCAGATCTGGAGCGTCGC
Squirrel AATGACATGGTGTGCTGCAGACCCAGTCCCTGAGAGTGCTCATCTCAGATCTGGAGCGTCGC
Owl AATGACATGGTGTGCTGCAGACCCAGTCCCTGAGAGTGCTCATCTCAGATCTGGAGCATCGC
Titi AATGACATGGTGTGCTGCAGACCCAGTCCATAAGCGTGCTCATCTCGGATCTGGAGCATCGC
Saki AATGACATGGTGTGCTGCAGACCCAGTCCATGGAGTGCTCATCTCAGATCTGGAGCATCGC
Howler AATGACATGGTGTGCTGCAGACCCAGTCCATAAGAGTGCTCATGTCAGACCTGGAGCGTCGC
Spider AATGACATGGTGTGCTGCAGACCCAGTCCATGAGAGTGCTCATGTCAGATCTGGAGCACCGC
Woolly AATGACATGGTGTGCTGCAGACCCAGTCCATGAGAGTGCTCATGTCAGATCTGGAGCATCGC

		Exon5→	Exon6→
Human	CTGCAGGGGTCAGTGATGGAGCTGCTTCAG	GGTGTGGATGGCGTCATAAAAAAG	GACGGAG (258)
Chimp	CTGCAGGGGTCAGTGATGGAGCTGCTTCAG	GGTGTGGATGGCGTCATAAAAAAG	GATGGAG
Gorilla	CTGCAGGGGTCAGTGATGGAGCTGCTTCAG	GGTGTGGATGGCGTCATAAAAAAG	GATGGAG
Orangutan	CTGCAGGGGTCAGTGATGGAGCTGCTTCAG	GGTGTGGATGGCATCATAAAAAAG	GATGCAG
Gibbon	CTGCAGGGGTCAGTGATGGAGCTGCTTCAG	GGTGTGGATGGCGTCATAAAAAAG	GATGAAG
Rhes_cDNA	TTGCAGGGGTCAATGATGGATCTACTGCAG	GGTGTGGATGGCATCATTAAAAAG	GATTGAG
Baboon	TTGCAGGGGTCAATGATGGAGCTACTGCAG	GGTGTGGATGGCATCATTAAAAAG	GATTGAG
AGM	TTGCAGGGGTCAATGATGGAGCTGCTGCAG	GGTGTGGATGGCATCATTAAAAAG	GATTGAG
AGM_cDNA	TTGCAGGGGTCAATGATGGAGCTGCTGCAG	GGTGTGGATGGCATCATTAAAAAG	GGTTGAG
Tant_cDNA	TTGCAGGGGTCAATGATGGAGCTGCTGCAG	GGTGTGGATGGCATCATTAAAAAG	GATTGAG
Patas	TTGCAGGGGTCAATGATGGAGCTGCTGCAG	GGTGTGGATGGCATCATTAAAAAG	GATTGAG
Colobus	TTGCAGGGGTCAATGATGGAGCTGCTGCAG	GGTGTGGATGGCATCATAAAAAAG	GATTGAG
DLangur	TTGCAGGGGTCAATGATGGAGCTGCTGCAG	GGTGTGGATGGCATCATAAAAAAG	GATTGAG
PMarmoset	CTGCAGGGGTCAATGATGGAGCTGCTGCAG	GGTGTGGATGGCATCATAAAAAAG	GATTGAG
Tamarin	CTGCAGGGGTCAATGATGGAGCTGCTGCAG	GGTGTGGATGGCATCATAAAAAAG	GATTGAG
Squirrel	CTGCAGGGGTCAATGATGGAGCTGCTGCAG	GGTGTGGATGGCATCATAAAAAAG	GATTGAG
Owl	CTGCAGGGGTCAATGATGGAGCTGCTGCAG	GGTGTGGATGGCATCATAAAAAAG	GATTGAG
Titi	CTGCAGGGGTCAATGATGGAGCTGCTGCAG	GGTGTGGATGGCATCATAAAAAAG	GGTTAAG
Saki	CTGCAGGGGTCAATGATGGAGCTGCTGCAG	GGTGTGGATGGCATCATAAAAAAG	GGTTAAG
Howler	CTGCAGGGGTCAATGATGGAGCTGCTGCAG	GGTGTGGATGGCATCATAAAAAAG	GATTAAG
Spider	CTGCAGGGGTCAATGATGGAGCTGCTGCAG	GGTGTGGATGGCATCATAAAAAAG	GATTAAG
Woolly	CTGCAGGGGTCAATGATGGAGCTGCTGCAG	GGTGTGGATGGCATCATAAAAAAG	GACTACG

Human	AACGTGACCTTGAAGAAGCCAGAAAACCTTTTCCAAAAAATCAAAGGAGAGTGTTTCGAGCT (278)
Chimp	AACGTGACCTTGAAGAAGCCAGAAAACCTTTTCCAAAAAATCAAAGGAGAGTGTTTCGAGCT
Gorilla	AACGTGACCTTGAAGAAGCCAGAAAACCTTTTCCAAAAAATCAAAGGAGAGTGTTTCGAGCT
Orangutan	AACGTGACCTTGAAGAAGCCAGAAAACCTTTTCCAAAAAATCAAAGGAGAGTGTTTCGAGCT
Gibbon	AACGTGACCTTGAAGAAGCCAGAAAACCTTTTCCAAAAAATCAAAGGAGAGTGTTTCGAGCT
Rhes_cDNA	AACATGACCTTGAAGAAGCCAAAAACCTTTTCCAAAAAATCAAAGGAGAGTGTTTCGAGCT
Baboon	AACATGACCTTGAAGAAGCCAAAAACCTTTTCCAAAAAATCAAAGGAGAGTGTTTCGAGCT
AGM	AACATGACCTTGAAGAAGCCAAAAACCTTTTCCAAAAAATCAAAGGAGAGTGTTTCGAGCT
AGM_cDNA	AACATGACCTTGAAGAAGCCAAAAACCTTTTCCAAAAAATCAAAGGAGAGTGTTTCGAGCT
Tant_cDNA	AACATGACCTTGAAGAAGCCAAAAACCTTTTCCAAAAAATCAAAGGAGAGTGTTTCGAGCT
Patas	AACATGACCTTGAAGAAGCCAGAAAACCTTTTCCAAAAAATCAAAGGAGAGTGTTTCGAGCT
Colobus	GACATGACCTTGAAGAAGCCAAAAACCTTTTCCAAAAAATCAAAGGAGAGTGTTTCGAGCT
DLangur	AACATGACCTTGAAGAAGCCAAAAACCTTTTCCAAAAAATCAAAGGAGAGTGTTTCGAGCT
PMarmoset	AAAGTTACTTTGCAGAAGCCAAAAACCTTTTCTTAATGAAAAAAGGAGAGTATTTTCGAGCT
Tamarin	ACAGTGACTTTGCAGAAGCCAAAAACCTTTTCTTAATGAAAAAAGGAGAGTATTTTCGAGCT
Squirrel	AAAGTGACTTTGCAGAAGCCAAAAACCTTTTCTTAATGAAAAAAGGAGAGTATTTTCGAGCT
Owl	AAAGTGACTTTGCAGAATCCAAAAACCTTTTCTTAATGAAAAAAGGAGAGTATTTTCAAACCT
Titi	AATGTGACTTTGCAGAAGCCAAAAACCTTTTCTTAATGAAAAAAGGAGAGTATTTTCGAGTT
Saki	AACGTGACTTTGCAGAAGCCGAAAACCTTTTCTTAATGAAAAAAGGAGAGTATTTTCGAGCT
Howler	AACGTGACTTTGCAGAAGCCAGAAAACCTTTTCTTAATGAAAAAAGGAGAGTATTTTCAAGCT
Spider	AATGTGACTTTGCAGAAGCCAAAAACCTTTTCTTAATGAAAAAAGGAGAGTATTTTCGAGCT
Woolly	AATGTGACTTTGCAGAAGCCAAAAACCTTTTCTTAATGAAAAAAGGAGAGTATTTTCGAGCT

		Exon7→	SPRY→	
Human	CCTGATCTGAAAGGAATGCTAGAAGTGTTTAGAG	AGCTGACAGATGTC	CGACGCTACTGG	(298)
Chimp	CCTGATCTGAAAGGAATGCTAGAAGTGTTTAGAG	AGCTGACAGATGTCCGACGCTACTGG		
Gorilla	CCTGATCTGAAAGGAATGCTAGAAGTGTTTAGAG	AGCTGACAGATGTCCGACGCTACTGG		
Orangutan	CCTAATCTGAAAGGAATGCTAGAAGTGTTTAGAG	AGCTGACAGATGTCCGACGCTACTGG		
Gibbon	GCTGATCTGAAAGTAATGCTAGAAGTGTTGAGAG	AGCTGAGAGATGTCCGACGCTACTGG		
Rhes_cDNA	CCTGATCTGAAAGGAATGCTAGACATGTTTAGAG	AGCTAACAGATGCCCCGACGCTACTGG		
Baboon	CCTGATCTGAAAGGAATGCTAGACATGTTTAGAG	AGCTAACAGATGTCCGACGCTACTGG		
AGM	CCTGATCTGAAAGGAATGCTAGACATGTTTAGAG	AGCTAACAGATGTCCGACGCTACTGG		
AGM_cDNA	CCTGATCTGAAAGGAATGCTAGACATGTTTAGAG	AGCTAACAGATGTCCGACGCTACTGG		
Tant_cDNA	CCTGATCTGAAAGGAATGCTAGACATGTTTAGAG	AGCTAACAGATGTCCGACGCTACTGG		
Patas	CCTGCTCTGAAAGGAATGCTAGACATGTTTAGAG	AGCTAACAGATGTCCGGCGCTACTGG		
Colobus	CCTGATCTGAAAGGAATGCTAGACATGTTTAGAG	AGCTAACAGATGTCCGACGCTACTGG		
DLangur	CCTGATCTGAAAGGAATCCTAGACATGTTTAGAG	AACTAACAGATGTCCGACGCTACTGG		
PMarmoset	CCTGATCTGAAAGGAATGCTACAAGCATTAAAG	AGCTGACAGAAGTCCAACGCTACTGG		
Tamarin	CCTGATCTGAAAGCAATGCTACAAGCATTAAAG	AGCTGACAGAAGTCCAACGCTACTGG		
Squirrel	CCTGATCTGAAAGGAATGCTCCAAGTGTAAAG	AACTGACAGAAGTCCAACGCTACTGG		
Owl	CCTGATCTGAAAGGAACACTACAAGTGTAAAG	AGCCGACAGAAGTCCAACGCTACTGG		
Titi	CCTGATCTGAAAGGAATGCTACAAGTGTCTAAAG	AGTTGACAGAAGTCCAACGCTACTGG		
Saki	CCTGATCTGAAAGGAATGCTACAAGTGTCTAAAG	AGCTGACAGAAGTCCAACGCTACTGG		
Howler	CCTGATCTGAAAGGAATGCTACAAGTGTCTAAAG	AGCTGAAAGAAGTCCAGTGCTACTGG		
Spider	CCTGATCTGAAAGGAATGCTACAAGTGTCTAAAG	AGCTGAAAGAAGTCCAGTGCTACTGG		
Woolly	CCTAATCTGAAAGGAATGCTACAAGTGTCTAAAG	AGCTGAAAGAAGTCCAATGCTACTGG		

	Exon8→	C	A	V	I	S	E	D	K	R	
Human	G TTGATGTGACAGTGGCTCCAAACAACATTTCA	TGTGCT	GTCATT	TCTGAA	GAT	AAG	AGA				(318)
Chimp	G TTGATGTGACAGTGGCTCCAAACAACATTTCA	TGTGCT	GTCATTT	TCTGAAG	ATG	GAGA					
Gorilla	G TTGATGTGACAGTGGCTCCAAACAACATTTCA	TGTGCT	GTCATTT	TCTGAAG	ATG	GAGA					
Orangutan	G TTGATGTGACAGTGGCTCCAAACGACATTTCA	TATGCT	GTCATTT	TCTGAAG	ATG	GAGA					
Gibbon	G TTGATGTGACAGTGGCTCCAAACAACATTTCA	TATGCT	GTCATTT	TCTGAAG	ATG	GAGA					
Rhes_cDNA	G TTGATGTGACACTGGCTACAAACAACATTTCA	CATGCT	GTCATTG	TCTGAAG	AAG	GAGA					
Baboon	G TTGATGTGACACTGGCTCCAAACAACATTTCA	CATGCT	GTCATTG	TCTGAAG	AAG	GAGA					
AGM	G TTGATGTGACACTGGCTCCAAACAACATTTCA	CATGCT	GTCATTG	TCTGAAG	AAG	GAGA					
AGM_cDNA	G TTGATGTGACACTGGCTCCAAACAACATTTCA	CATGCT	GTCATTG	TCTGAAG	AAG	GAGA					
Tant_cDNA	G TTGATGTGACACTGGCTCCAAACAACATTTCA	CATGCT	GTCATTG	TCTGAAG	AAG	GAGA					
Patas	G TTGATGTGACACTGGCTCCAAACAACATTTCA	CATGCT	GTCATTG	TCTGAAG	AAG	GAGA					
Colobus	G TTGATGTGACACTGGCTCCAAACAACATTTCA	CATGCT	GTCATTG	TCTGAAG	AAG	GAGA					
DLangur	G TTGATGTGACACTGGCTCCAAACAACATTTCA	CATGCT	GTCATTG	TCTGAAG	AAG	GAGA					
PMarmoset	G CTCATGTGACACTGGTTCCAAGTCACCCTTCA	TGTACT	GTCATTT	TCTGAAG	GAG	GAGA					
Tamarin	G CTCATGTGACACTGGTTCCAAGTCACCCTTCA	TATGCT	GTTATTT	TCTGAAG	GAG	GAGA					
Squirrel	G CTCATGTGACACTGGTTCCAAGTCACCCTTCA	TATACT	ATCATTT	TCTGAAG	GGG	GAGA					
Owl	G CTCATGTGACACTGGTTCCAAGTCACCCTTCA	TGTACT	GTCATTT	TCTGAAG	GAG	GAGA					
Titi	G CTCATGTGACACTGGTTGCAAGTCACCCTTCA	CGTGCT	GTCATTT	TCTGAAG	GAA	GAGA					
Saki	G TTCATGTGACACTGGTTCCAAGTCACCCTTCA	TGTGCT	GTCATTT	TCTGAAG	GAG	GAGA					
Howler	G CTCATGTGACACTGATTCCGAATCACCCTTCA	TGTACT	GTCATTT	TCTGAAG	AAG	GAGA					
Spider	G CTCATGTGACACTGGTTCCAAGTCACCCTTCA	TGTACT	GTCATTT	TCTGAAG	GAG	GAGA					
Woolly	G CTCATGTGACACTGGTTCCAAGTCACCCTTCA	TGTGCT	GTCATTT	TCTGAAG	CAG	GAGA					

(Patch) →

	Q	V	S	S	P	K	P	Q	I	I	Y	G	A	R	G	T	R	Y	Q	T
Human	CAAGT	GAGCT	TCTCCG	AAACCA	CAGATA	AATATAT	GGGGC	ACGA	GGGACA	AAGATAC	CAG	---	ACA	(338)						
Chimp	CAAGT	GAGCT	TCTCCG	AAACCA	CAGATA	AATATAT	GGGGC	ACGA	GGGACA	AAGATAT	CAG	---	ACA							
Gorilla	CAAGT	GAGCT	TCTCCG	AAACCA	CAGATA	AATATAT	GGGGC	ACAA	GGGACA	AAGATAT	CAG	---	ACA							
Orangutan	CAAGT	GAGCT	GTCCG	GAACCA	CAGATA	AATATAT	GGGGC	ACAA	GGGACA	AACATAT	CAG	---	ACA							
Gibbon	CAAGT	GAGCT	TCTCCG	GAACCA	CAGATA	AATATTT	GAGGC	ACAA	GGGACA	AATATCT	CAG	---	ACA							
Rhes_cDNA	CAAGT	GAGCT	TCTCGG	AAACCA	CAGATA	AATGTAT	CAGGC	ACCA	GGGAC	ATTTACG	---	TTT								
Baboon	CAAGT	GAGCT	TCTCGG	AAACCA	CAGATA	AACGTAT	CAGGC	ACCA	GGGAC	ATTTTCG	---	TTT								
AGM	CAAGT	GAGCT	ATCAGA	AAACCA	CAGATA	AATGTAT	CAGGC	ACCA	GGGTC	ATTTGGG	---	TCA								
AGM_cDNA	CAAGT	GAGCT	ATCGGA	AAACCA	CAGATA	AATGTAT	CAGTC	ACCA	GGGTC	ATTTGGG	---	TCA								
Tant_cDNA	CAAGT	GAGCT	ATCAGA	AAACCA	CAGATA	AATGTAT	CAGGC	ACCA	GGGTC	ATTTGGG	---	TCA								
Patas	CAAGT	GAGCT	TCTCGG	AAACCA	CAGATA	AATGTAT	TGGGC	ACAA	GGGAA	ATTTTCAG	---	TCA								
Colobus	CGAGT	GAGCT	TCTCCG	AAACCA	CAGATA	AATGTAT	CGGGC	ACAA	GGGAC	ATTTTCAG	---	TCA								
DLangur	CAAGT	GAGCT	TCTCCG	AAACCA	CAGATA	AATGTGT	CGGGC	ACGA	GGGAC	ATTTTCAG	---	TCA								
PMarmoset	CAAGT	GAGATAT	CAGGTT	CCG	-----	-----	-----	-----	-----	ATACAT	CAA	---	CCA							
Tamarin	CAAGT	GAGATAT	CAGTTT	CAG	-----	-----	-----	-----	-----	ATACAT	CAA	---	CCA							
Squirrel	CAAGT	GAGATAT	CAGAA	ACCT	-----	-----	-----	-----	-----	ATAC	GCAC	---	CTA							
Owl	CAAGT	GAGATAT	CAGAA	ACGG	-----	-----	-----	-----	-----	ATATAT	CAA	---	CCA							
Titi	CAAGT	GAGATAT	CAGGA	ATGG	-----	-----	-----	-----	-----	ATACAT	CAA	---	TCA							
Saki	CAAGT	GAGATAT	CAGGA	ACGG	-----	-----	-----	-----	-----	ATACAT	CAA	---	TCA							
Howler	GAAGT	GAGATAT	CAGGA	ACAG	-----	-----	-----	-----	-----	ATACAT	CATC	ACCCG								
Spider	CAAGT	GAGATAT	CAGGA	ACAG	-----	-----	-----	-----	-----	ATACAT	CAA	---	CCA							
Woolly	CAAGT	GAGATAT	CAGAA	ACAG	-----	-----	-----	-----	-----	AGACAT	CGA	---	CCA							

Human	-----
Chimp	-----
Gorilla	-----
Orangutan	-----
Gibbon	-----
Rhes_cDNA	CCGTCA-----
Baboon	CCGTCA-----
AGM	CTCACGAATTTCAATTATTGTACTGGCGTCCTGGGCTCCCAAAGTATCACATCAAGGAAA
AGM_cDNA	CTCACGAATTTCAATTATTGTACTGGCGTCCTGGGCTCCCAAAGTATCACATCAGGGAAA
Tant_cDNA	CTCACGAATTTCAATTATTGTACTGGCGTCCTGGGCTCCCAAAGTATCACATCAAGGAAA
Patas	-----
Colobus	-----
DLangur	-----
PMarmoset	-----
Tamarin	-----
Squirrel	-----
Owl	-----
Titi	-----
Saki	-----
Howler	-----
Spider	-----
Woolly	-----

F V N F N Y C T G I

Human TTTGTG AATTTCAATTATTGTACTGGCATCCTGGGCTCTCAAAGTATCACATCAGGGAAA (358)
Chimp TTTATG AATTTCAATTATTGTACTGGCATCCTGGGCTCTCAAAGTATCACATCAGGGAAA
Gorilla TTTATG AATTTCAATTATTGTACTGGCATCCTGGGCTCTCAAAGTATCACATCAGGGAAA
Orangutan TATGTG AATTTCAATTATTGTACTGGCATCCTGGGCTCTCAAAGTATCACATCAGGGAAA
Gibbon TTTGTG AATTTCAATTATTGTACTGGCATCCTGGGCTCTCAAAGTATCACATCAGGGAAA
Rhes_cDNA CTCACG AATTTCAATTATTGTACTGGCGTCCTGGGCTCCCAAAGTATCACATCAGGGGAA
Baboon CTCACG AATTTCAATTATTGTACTGGCGTCCTGGGCTCCCAAAGTATCACATCAGGGGAA
AGM CTCACG AATTTCAATTATTGTACTGGCGTCCTGGGCTCCCAAAGTATCACATCAGGGGAA
AGM_cDNA CTCACG AATTTCAATTATTGTACTGGCGTCCTGGGCTCCCAAAGTATCACATCAGGGGAA
Tant_cDNA CTCACG AATTTCAATTATTGTACTGGCGTCCTGGGCTCCCAAAGTATCACATCAGGGGAA
Patas CTCAAG AATTTCAATTATTGTACTGGCATCCTGGGCTCCCAAAGTATCACATCAGGGGAA
Colobus CTCAAG AATTTCAATTATTGTACTGGCATCCTGGGCTCCCAAAGTATCACATCAGGGGAA
DLangur CTCAAG AATTTCAATTATTGTACTGGCGTCCTGGGCTCCCAAAGTATCACATCAGGGGAA
PMarmoset CTTGTG AAAGTCAAGTATTTTTATGGCGTCCTGGGCTCCCAAAGTATCACATCAGGGGAA
Tamarin TCTGTG AAAGTCAACTATTTTTATGGCGTCCTGGGCTCCCAAAGTATCACATCAGGGGAA
Squirrel CTTGTG AAAGTCCAGTATTTTTATGGCGTCCTGGGCTCCCAAAGTATCACATCAGGGGAA
Owl TTTCTG AAAGTCAAGTATTTTTATGGCGTCCTGGGCTCCCAAAGTATCACATCAGGGGAA
Titi TCTGGG AGAGTCAAGTATTTTTATGGCGTCCTGGGCTCCCAAAGTATCACATCAGGGGAA
Saki TTTGGG AAAGTCAAGTATTTTTATGGCGTCCTGGGCTCCCAAAGTATCAGATCAGGGGAA
Howler TCTATG GAAGTCAAGTATTTTTATGGCATCCTGGGCTCCCAAAGTATCACATCAGGGGAA
Spider TCTGTG AAAGTCAAGTATTTTTATGGCGTCCTGGGCTCCCAAAGTATCACATCAGGGGAA
Woolly TCTGTG AAAGCCAAATATTTTTATGGCGTCCTGGGCTCCCAAAGTATCACATCAGGGGAA

Human CATTACTGGGAGGTAGACGTGTCCAAGAAAAGTGCTTGGATCCTGGGGGTATGTGCTGGC (378)
Chimp CATTACTGGGAGGTAGACGTGTCCAAGAAAAGTGCTTGGATCCTGGGGGTATGTGCTGGC
Gorilla CATTACTGGGAGGTAGACGTGTCCAAGAAAAGTGCTTGGATCCTGGGGGTATGTGCTGGC
Orangutan CATTACTGGGAGGTAGACGTGTCCAAGAAAAGTGCTTGGATCCTGGGGGTATGTGCTGGC
Gibbon CATTACTGGGAGGTAGACGTGTCCAAGAAAAGTGCTTGGATCCTGGGGGTATGTGCTGGC
Rhes_cDNA CATTACTGGGAGGTAGATGTGTCCAAGAAAAGTGCTTGGATCCTGGGGGTATGTGCTGGC
Baboon CATTACTGGGAGGTAGATGTGTCCAAGAAAAGTGCTTGGATCCTGGGGGTATGTGCTGGC
AGM CATTACTGGGAGGTAGATGTGTCCAAGAAAAGTGCTTGGATCCTGGGGGTATGTGCTGGC
AGM_cDNA CATTACTGGGAGGTAGATGTGTCCAAGAAAAGTGCTTGGATCCTGGGGGTATGTGCTGGC
Tant_cDNA CATTACTGGGAGGTAGATGTGTCCAAGAAAAGTGCTTGGATCCTGGGGGTATGTGCTGGC
Patas CATTACTGGGAGGTAGATGTGTCCAAGAAAAGTGCTTGGATCCTGGGGGTATGTGCTGGC
Colobus CATTACTGGGAGGTAGATGTGTCCAAGAAAAGTGCTTGGATCCTGGGGGTATGTGCTGGC
DLangur CATTACTGGGAGGTAGATGTGTCCAAGAAAAGTGCTTGGATCCTGGGGGTATGTGCTGGC
PMarmoset CATTACTGGGAAGTAGACGTGTCCAATAAAAAGGGATTGGATCCTGGGGGTATGTGTTAGC
Tamarin CATTACTGGGAGGTAGACGTGTCCAATAAAAAGGGATTGGATCCTGGGGGTATGTGTTAGC
Squirrel CATTACTGGGAGGTAGACGTGTCCAATAAAAAGGGATTGGATCCTGGGGGTATGTGTTAGC
Owl CATTACTGGGAGGTAGACGTGTCCAATAAAAAGTGAGTGGATCCTGGGGGTATGTGTTAGC
Titi CATTACTGGGAGGTAGACGTGTCCAATAAAAAGTGCTTGGATCCTGGGGGTATGTGTTAGC
Saki CATTACTGGGAGGTAGACGTGTCCAATAAAAAGTGCTTGGATCCTGGGAGTATGTGTTAGC
Howler CATTACTGGGAGGTAGACGTGTCCAATAAAAAGTGCTTGGATCCTGGGGGTATGTGTCAGC
Spider CATTACTGGGAGGTAGACGTGTCCGATAAAAAGTGCTTGGATCCTGGGGGTATGTGTTAGC
Woolly CATTACTGGGAGGTAGACGTGTCCAATAAAAAGTGCTTGGATCCTGGGGGTATGTGTTAGC

Human	TTCCAACCTGATGCAATG	(384)
Chimp	TTCCAACCTGATGCAATG	
Gorilla	TTCCAACCTGATGCAACG	
Orangutan	TTCCAACCTGATGCAATG	
Gibbon	TTGCAACCTGATGCAATG	
Rhes_cDNA	TTCCAATCCGATGCAATG	
Baboon	TTCCAACCTGATGCAATG	
AGM	TTCCAACCCGATGCAACG	
AGM_cDNA	TTCCAACCCGATGCAACG	
Tant_cDNA	TTCCAACCCGATGCAACG	
Patas	TTCCAACCCGATGCAATG	
Colobus	TTCCAACCCGATGCAATG	
DLangur	TTCCAACCCGATGCAATG	
PMarmoset	TGGAATGCAATGCAAAA	
Tamarin	TTTAAATGCAATGCAAAA	
Squirrel	TTGAAATGTACTGCAAAAT	
Owl	TTGAAGCGCACTGCAAGT	
Titi	TTGAAATGCGCTGCAAAAT	
Saki	TTGAAATGCACTGCAAAAT	
Howler	TTGAAATGCATTGGAAATTTTCCAGGAATTGAAAATTATCAACCTCAAAAATGGCTACTGG	
Spider	TTGAAATGCACTGCAAAATGTTCCAGGAATTGAAAATTATCAACCTAAAAATGGCTACTGG	
Woolly	TTGAAATGCACTGCAAAATGTTCCAGGAATTGAAAATTATCAACCTAAAAATGGCTACTGG	

Human	-----TGTAAT-----	ATTGAAAAAAAT	(390)
Chimp	-----TGTAAT-----	ATTGAAAAAAAT	
Gorilla	-----TGTAAT-----	ATTGAAAAAAAT	
Orangutan	-----TATAAT-----	ATTGAACAAAAT	
Gibbon	-----TATAAT-----	ATTGAACAAAAT	
Rhes_cDNA	-----TATAAT-----	ATTGAACAAAAT	
Baboon	-----TATAAT-----	ATTGAACAAAAT	
AGM	-----TATAAT-----	ATTGAACAAAAT	
AGM_cDNA	-----TATAAT-----	ATTGAACAAAAT	
Tant_cDNA	-----TATAAT-----	ATTGAACAAAAT	
Patas	-----TATGAT-----	GTTGAACAAAAT	
Colobus	-----TATAAT-----	ATTGAACAAAAT	
DLangur	-----TATAAT-----	ATTGAACAAAAT	
PMarmoset	-----TGGAAT-----	GTTCTAAGACCT	
Tamarin	-----TGGAAT-----	GTTCTAAGACCT	
Squirrel	-----CAGAGT-----	GTTTCAGGAACT	
Owl	-----TGTAGT-----	GTTCCAAGAATT	
Titi	-----CGGAAT-----	GGTCCAGGAGTT	
Saki	-----CGGAAT-----	GGTCCAAGAATT	
Howler	GTTATAGGGTTACGGAATGCAGATAACTATAGTGCTTTCCAAGATGCA	GTTCCAGAACT	
Spider	GTTATAGGGTTACAGAATGCAATAACTATAGTGCTTTCCAGGATGCA	GTTCCAGGAACT	
Woolly	GTTATAGGGTTACAGAATGCAGATAACTATAGTGCTTTCCAGGATGCA	GTTCCAGGAACT	

Human	<u>GAAAATTATCAACCTAAATACGGC</u> -----	(398)
Chimp	GAAAATTATCAACCTAAATATGGC-----	
Gorilla	GAAAATTATCAACCTAAATATGGC-----	
Orangutan	GAAAATTATCAACCTCAATATGGC-----	
Gibbon	GAAAATTATCAACCTAAATATGGC-----	
Rhes_cDNA	GAAAATTATCAACCTAAATATGGC-----	
Baboon	GAAAATTATCAACCTAAATATGGC-----	
AGM	GAAAATTATCAACCTAAATATGGA-----	
AGM_cDNA	GAAAATTATCAACCTAAATATGGC-----	
Tant_cDNA	GAAAATTATCAACCTAAATATGGC-----	
Patas	GAAAATTATCAACCTAAATATGGC-----	
Colobus	GAAAATTATCAACCTAAATATGGC-----	
DLangur	GAAAATTATCAACCTAAATATGGC-----	
PMarmoset	GAAAATTATCAACCTAAAAATGGC-----	
Tamarin	GAAAATTATCAACCTAAAAATGGC-----	
Squirrel	GAAAATTATCAACCTAAAAATGGC-----	
Owl	GAAAATGATCAACCTAAAAATGGC-----	
Titi	GAAAATATCAACCTAAAAATGGC-----	
Saki	GAAAATTATCAACCTAAAAATGGC-----	
Howler	GAAAATTATCAACCTAAAAATCGCAACCGG---TTTACAGGGTTACAGAATGCAGATAATTGT	
Spider	GAAAATTATCAACCTAAAAATGGCAACCGGAGGAATAAAGGGTTACGGAATGCAGATAACTAT	
Woolly	GAAGATTATCAACCTAAAAATGGCTGCTGGAGGAATACAGGGTTACGGAATGCAGATAACTAT	

Human	<u>-----TACTGG</u>	(400)
Chimp	-----TACTGG	
Gorilla	-----TACTGG	
Orangutan	-----TACTGG	
Gibbon	-----TACTGG	
Rhes_cDNA	-----TACTGG	
Baboon	-----TACTGG	
AGM	-----TACTGG	
AGM_cDNA	-----TACTGG	
Tant_cDNA	-----TACTGG	
Patas	-----TACTGG	
Colobus	-----TACTGG	
DLangur	-----TACTGG	
PMarmoset	-----TACTGG	
Tamarin	-----TACTGG	
Squirrel	-----TACTGG	
Owl	-----TACTGG	
Titi	-----TACTGG	
Saki	-----TACTGG	
Howler	AGTGCTTTCCAGAATGCATTTCCAGGAATTCAAAGTTATCAACCTAAAAAGAGCCACTTG	
Spider	AGTGCTTTCCGGGATACATTT-----CAACCTATAAATGACTCCTGG	
Woolly	AGTGCTTTCCAGGATGTATTT-----CAACCTAAAAATGACTACTGG	

Human	GTTATAGGGTTA-----GAGGAA GGA GTTAAATGTAGT	(411)
Chimp	GTTATAGGGTTA-----GAGGAA GGA GTTAAATGTAGT	
Gorilla	GTTATAGGGTTA-----GAGGAA GGA GTTAAATGCAGT	
Orangutan	GTTATAGGGTTA-----GAGGAA GGA GTTAAATGTAGT	
Gibbon	GTTATAGGGTTA-----GAGGAA GGA GTTAAATGTAAT	
Rhes_cDNA	GTTATAGGGTTA-----CAGGAA GGA GTTAAATATAGT	
Baboon	GTTATAGGGTTA-----CAGGAA GGA GTTAAATATAGT	
AGM	GTTATAGGGTTA-----CAGGAA GGA GATAAATATAGT	
AGM_cDNA	GTTATAGGGTTA-----CAGGAA GGA GATAAATATAGT	
Tant_cDNA	GTTATAGGGTTA-----CAGGAA GGA GATAAATATAGT	
Patas	GTTATAGGGTTA-----CAGGAA GGA GTAATAATATAGT	
Colobus	GTTATAGGGTTA-----CAGGAA GGA GTTAAATATAGT	
DLangur	GTTATAGGGTTA-----CAGGAA GGA GTTAAATATAAT	
PMarmoset	GTTATAGGGTTA CGGAATACAGATAACTATAGTGCTTTC CAGGAT GCA GTTAAATATAGT	
Tamarin	GTTATAGGGTTA CAGAATACAAATAACTATAGTGCTTTC CAGGAT GC AGTTAAATATAGT	
Squirrel	GTTATAGGGTTA CGGAATGCAGGTA ACTATAGGGCTTTCAGAGT TC ATTTGAATTTTCGT	
Owl	GTTATAGGGTTA CGGAATGCAGATA ACTATAGTGCTTTCAGGAT GC AGTTGAATATAGT	
Titi	GTGATAGGGTTA CGGAATGCAGATA ACTATAGTGCTTTCAGGAT TC AGTTAAATATAAT	
Saki	GTTATAGGGTTA TGGAATGCAGGTA ACTATAGTGCTTTCAGGAT TC AGTTAAATATAGT	
Howler	TTTACAGGGTTA CAGAATCTAAGTA ACTATAATGCTTTCAGAA AAA GTTCAATATAAC	
Spider	GTTACAGGGTTA CGGAATGTAGATA ACTATAATGCTTTCAGGAT GC AGTTAAATATAGT	
Woolly	GTTACAGGGTTA TGGAACGCAGATA ACTATAATGCTTTCAGGAT GC AGTTAAATATAGT	

Human	-----GCTTTCCAGGATAGTTCC TTC CATACT CCT TCT GTT CCTTTTCATTGTGCCCCCTC	(429)
Chimp	-----GCTTTCCAGGATGGTTCC TTC CATACT CCT TCT GCT CCTTTTCATTGTGCCCCCTC	
Orangutan	-----GCTTTCCAGGATGGTTCC TTC CATACT CCT TCT GCT CCTTTTCATTGTGCCCCCTC	
Gibbon	-----GCTTTCCAGGATGGTTCC ATC CATACT CCT TCT GCT CCTTTTCATTGTGCCCCCTC	
Rhes_cDNA	-----GTTTTCCAGGATGGTTCC TCA CATACT CCT TTT GCT CCTTTTCATTGTGCCCCCTC	
Baboon	-----GTTTTCCAGGATGGTTCC TCA CATACT CCT TTT GCT CCTTTTCATTGTGCCCCCTC	
AGM	-----GTTTTCCAGGATAGTTCC TCA CATACT CCT TTT GCT CCTTTTCATTGTGCCCCCTC	
AGM_cDNA	-----GTTTTCCAGGATGGTTCC TCA CATACT CCT TTT GCT CCTTTTCATTGTGCCCCCTC	
Tant_cDNA	-----GTTTTCCAGGATGGTTCC TCA CATACT CCT TTT GCT CCTTTTCATTGTGCCCCCTC	
Patas	-----GTTTTCCAGGATGGTTCC TCA CATACT CCT TTT GCT CCTTTTCATTGTGCCCCCTC	
Colobus	-----GTTTTCCAGGATGGTTCC TCA CATACT CCT TTT GCT CCTTTTCATTGTGCCCCCTC	
DLangur	-----GTTTTCCAGGATGGTTCC TCA CATACT CCT TTT GCT CCTTTTCATTGTGCCCCCTC	
PMarmoset	-----GATGTCCAGGATGGTTCT CGC TCTGTT TCT TCT GGT CCTTTTGATCGTGCCCCCTC	
Tamarin	-----GATTTCCAGATTGGTTCC CGC TCTACT GCT TCT GTT CCTTTTGATCGTGCCCCCTC	
Squirrel	-----GATTTCCCTGGCTGGTTCC CGC CTTACT CTT TCT CCT CCTTTTGATCGTGCCCCCTC	
Owl	-----GATTTCCAGGATGGTTCC CGC TCTACT CCT TCT GCT CCTTTTGATCGTGCCCCCTC	
Titi	-----GATTTCCAGGATGGTTCC CGC TCTACT ACT TAT GCT CCTTTTGATCGTGCCCCCTC	
Saki	-----GATTTCCAGGATGGTTCC CAC TCTGCT ACT TAT GGT CCTTTTGATCGTGCCCCCTC	
Howler	TATATT GATTTTCAGGATGATTCC CTC TCTACT CCT TCT GCT CCTTTTGATCGTACCCCTC	
Spider	-----GATTTCCAGGATGGTTCC TGC TCTACT CCT TCT GCT CCTTTTGATGGTGCCCCCTC	
Woolly	-----GATTTCCAGGATGGTTCC TGC TCTACT CCT TTT GCT CCTTTTGATTGTGCCCCCTC	

Human TCTGTGATTATTTGTCCTGATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC (449)
Chimp TCTGTGATTATTTGTCCTGATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
Gorilla TCTGTGATTATTTGTCCTGATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
Orangutan TCTGTGATTATTTGTCCTGATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
Gibbon TCTGTGAATATTTGTCCTGATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
Rhes_cDNA TCTGTGATTATTTGTCCTGATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
Baboon TCTGTGATTATTTGTCCTGATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
AGM TCTGTGATTATTTGTCCTGATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
AGM_cDNA TCTGTGATTATTTGTCCTGATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
Tant_cDNA TCTGTGATTATTTGTCCTGATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
Patas TCTGTGATTTTTGTCCTGATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
Colobus TCTGTGATCATTGTCCTGATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
DLangur TCTGTGATTATTTGTCCTGATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
PMarmoset TTTATGACTATTTGTCCTAATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTATC
Tamarin TTTATGACTATTTATCCTAATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
Squirrel TTTATGACTATTTGTCCTAATCGGGTCGGAGTTTTCTAGACTATGAGGCTCGCACTATC
Owl TTTATGACTATTTGTCCTAATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
Titi TTTATGACTATTTGTCCTAATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
Saki TTTATGACTATTTGTCCTAATCGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
Howler TTTATGACTATTTGTCCTAAACGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC
Spider TTTATGACTATTTGTCCTAAACGTGTGGAGTTTTCTAGACTGTAAGGCTTGCACTGTC
Woolly TTTATGACTATTCGTCTAAACGTGTGGAGTTTTCTAGACTATGAGGCTTGCACTGTC

Human TCATTCTTCAATATCACAAACCATGGATTTCTCATCTATAAGTTTTCTCACTGTTCTTTT (469)
Chimp TCATTCTTCAATATCACAAACCATGGATTTCTCATCTATAAGTTTTCTCACTGTTCTTTT
Gorilla TCATTCTTCAATATCACAAACCATGGATTTCTCATCTATAAGTTTTCTCACTGTTCTTTT
Orangutan TCATTCTTCAATATCACAAACCATGGATTTCTCATCTATAAGTTTTCTCACTGTTCTTTT
Gibbon TCATTCTTCAATATCACAGACCATGGATTTCTCATCTATAAGTTTTCTCACTGTTCTTTT
Rhes_cDNA TCATTCTTCAATATCACAAACCATGGATTTCTCATCTATAAGTTTTCTCAGTGTCTTTT
Baboon TCATTCTTCAATATCACAAACCATGGATTTCTCATCTATAAGTTTTCTCAGTGTCTTTT
AGM TCATTCTTCAATATCACAAACCATGGATTTCTCATCTATAAGTTTTCTCAGTGTCTTTT
AGM_cDNA TCATTCTTCAATATCACAAACCATGGATTTCTCATCTATAAGTTTTCTCAGTGTCTTTT
Tant_cDNA TCATTCTTCAATATCACAAACCATGGATTTCTCATCTATAAGTTTTCTCAGTGTCTTTT
Patas TCATTCTTCAATATCACAAACCATGGATTTCTCATCTATAAGTTTTCTCAGTGTCTTTT
Colobus TCATTCTTCAATATCACAAACCATGGATTTCTCATCTATAAGTTTTCTCAGTGTCTTTT
DLangur TCATTCTTCAATATCACAAACCATGGATTTCTCATCTATAAGTTTTCTCAGTGTCTTTT
PMarmoset TCATTCTTCAATGTCACAAGCAATGGATTTCTCATCTATAAGTTTTCTAACTGTCATTTT
Tamarin TCATTCTTCAATGTCACAAGCAATGGATTTCTCATCTATAAGTTTTCTAACTGTCATTTT
Squirrel TCATTCTTCAATGTCACAAGCAATGGATTTCTCATCTATAAGTTTTCTGACTGTCATTTT
Owl TCATTCTTCAATGTCACAAGCAATGGATTTCTCATCTATAAGTTTTCTAACTGTCATTTT
Titi TCATTCTTCAATGTCACAAGCAATGGATTTCTCATCTATAAGTTTTCTAACTGTCATTTT
Saki TCATTCTTCAATGTCACAAGCAATGGATTTCTCATCTATAAGTTTTCTAACTGTCGTTTTT
Howler TCATTCTTCAATGTCACAAGCAATGGATATCTCATCTATAAGTTTTCTAACTGTCAGTTT
Spider TCATTCTTCAATGTCACAAGCAATGGATGTCCTCATCTATAAGTTTTCTAAGTGTCATTTT
Woolly TCATTCTTCAATGTCACAAGCAATGGATGTCCTCATCTATAAGTTTTCTAACTGTCATTTT

Human	<u>TCTCAGCCTGTATTTCCATATTTAAATCCTAGAAAATGTGGAGTCCCCATGACTCTGTGC</u> (489)
Chimp	TCTCAGCCTGTATTTCCATATTTAAATCCTAGAAAATGTGGAGTCCCCATGACTCTGTGC
Gorilla	TCTCAGCCTGTATTTCCATATTTAAATCCTAGAAAATGTAGAGTCCCCATGACTCTGTGC
Orangutan	TCTCAGCCTGTATTTCCATATTTAAATCCTAGAAAATGTAGAGTCCCCATGACTCTGTGC
Gibbon	TCTCAGCCTGTATTTCCATATTTAAATCCTAGAAAATGTACAGTCCCCATGACTCTGTGC
Rhes_cDNA	TCTAAGCCTGTATTTCCATATTTAAATCCAGAAAATGTACAGTCCCCATGACTCTGTGC
Baboon	TCTAAGCCTGTATTTCCATATTTAAATCCAGAAAATGTACAGTCCCCATGACTCTGTGC
AGM	TCTAAGCCTGTATTTCCATATTTAAATCCAGAAAATGTACAGTCCCCATGACTCTGTGC
AGM_cDNA	TCTAAGCCTGTATTTCCATATTTAAATCCAGAAAATGTACAGTCCCCATGACTCTGTGC
Tant_cDNA	TCTAAGCCTGTATTTCCATATTTAAATCCAGAAAATGTACAGTCCCCATGACTCTGTGC
Patas	TCTAAGCCTGTATTTCCATATTTAAATCCAGAAAATGTACAGTCCCCATGACTCTGTGC
Colobus	TCTAAGCCTGTATTTCCATATTTAAATCCTAGAAAATGTACAGTCCCCATGACTCTGTGC
DLangur	TCTAAGCCTGTATTTCCATATTTAAATCCTAGAAAATGTACAGTCCCCATGACTCTGTGC
PMarmoset	TCTTATCCTGTATTTCCATATTTAAGTCCAGCAGCAGTGAATTACCCATGACTCTGTGC
Tamarin	TCTTATCCTGTATTTCCATATTTAAGTCCAGCAGCAGTGAATTACCCATGACTCTGTGC
Squirrel	TCTTATCCTGTATTTCCATATTTCAATCCTATGACGTTGAATTACCCATGACTCTGTGC
Owl	TGTTATCCTGTATTTCCATATTTAAGTCCAGCAGCAGTGAATTACCCATGACTCTGTGC
Titi	TCTTATCCTGTATTTCCATATTTAAGTCCAGCAGCAGTGAATTACCCATGACTCTGTGC
Saki	TCTGATCCTGTATTTCCATATTTAAGTCCAGCAGCAGTGAATTACCCATGACTCTGTGC
Howler	TCTTATCCTGTATTTCCATATTTAAGTCCAGCAGCAGTGAATTACCCATGACTCTGTGC
Spider	TCTTATCCTGTATTTCCATATTTAAGTCCAGCAGCAGTGAATTACCCATGACTCTGTGC
Woolly	TCTTGTCCGTATTTCCATATTTAAGTCCAGCAGCAGTGAATTACCCATGACTCTGTGC

Human	<u>TCACCAAGCTCTTGA</u> (493)
Chimp	TCACCAAGCTCTTGA
Gorilla	TCGCAAGCTCTTGA
Orangutan	TCACCAAGCTCTTGA
Gibbon	TCACCAAGCTCTTGA
Rhes_cDNA	TCACCAAGCTCTTGA
Baboon	TCACCAAGCTCTTGA
AGM	TCACCAAGCTCTTGA
AGM_cDNA	TCACCAAGCTCTTGA
Tant_cDNA	TCACCAAGCTCTTGA
Patas	TCACCAAGCTCTTGA
Colobus	TCACCAAGCTCTTGA
DLangur	TCACCAAGCTCTTGA
PMarmoset	TCACCAAGCTCTTGA
Tamarin	TCACCAAGCTCTTGA
Squirrel	TCACCAAGCTCTTGA
Owl	TCACCAAGCTCTTGA
Titi	TCACCAAGCTCTTGA
Saki	TCACCAAGCTCTTGA
Howler	TCACCAAGCTCTTGA
Spider	TCACCAAGCTCTTGA
Woolly	TCACCAAGCTCTTGA

Institution: **INIST CNRS** | [Sign In via User Name/Password](#)

 Sawyer *et al.* 10.1073/pnas.0409853102.

Supporting Information

Files in this Data Supplement:

[Supporting Table 2](#)
[Supporting Appendix 1](#)
[Supporting Data Set 1](#)
[Supporting Figure 4](#)

[Supporting Data Set 1](#)

Data Set 1. Alignment of TRIM5 α sequences from primates analyzed in this study. Protein domains are indicated by boxes on the human sequence. Numbers at the end of each line (in parentheses) indicate the human codon number. Vertical lines denote exon breakpoints. PAML-highlighted codons (codons evolving under positive selection) are denoted with blue (highlighted in analysis of full data set) or yellow (highlighted in hominid+OWM data set) highlighting. Purple boxed human sequence and purple translated residues above the alignment indicate the SPRY "patch." Blocks of sequence in orange type contain gaps and were omitted from the PAML analysis of the full data set. The small section blocked with orange lines (and orange type) cannot be unambiguously aligned to the rest of the alignment. Two AGM sequences are present in the alignment. "AGM-cDNA" is cDNA read no. AY625003 from GenBank. This sequence was not used in the final analysis, and instead the read titled "AGM" (our sequence from Cos-7 DNA) was used. The owl monkey sequence in this alignment (gift from J. Luban, Columbia University, New York) also was not used in the analysis, because it has not been shown that TRIM5 α is transcribed as well as TRIM-Cyp.

[Supporting Figure 4](#)

Fig. 4. Positive selection of the SPRY domain. dN is plotted vs. dS for all pairwise combinations of primate sequences. These values are calculated by using the Nei–Gojobori method by PAML software [Yang, Z. (1997) *Comput. Appl. Biosci.* **13**, 555–556]. Pairwise combinations of hominids (hom),

This Article
<ul style="list-style-type: none"> ▶ Abstract ▶ Full Text
Services
<ul style="list-style-type: none"> ▶ Email this article to a colleague ▶ Alert me to new issues of the journal ▶ Request Copyright Permission
Citing Articles
<ul style="list-style-type: none"> ▶ Citing Articles via CrossRef

Old World monkeys (owm), and New World monkeys (nwm) are plotted with different symbols as shown in the legend. Points that are making comparisons between human and another primate are in a black box. In the first graph, the entire gene length is analyzed, and in the second and third graphs, the gene has been divided into two at the beginning of the SPRY (as defined in the alignment shown). The SPRY domain is predominantly evolving under positive selection, whereas the remainder of the gene shows a net signature of purifying selection.

<i>This Article</i>
▶ Abstract
▶ Full Text
<i>Services</i>
▶ Email this article to a colleague
▶ Alert me to new issues of the journal
▶ Request Copyright Permission
<i>Citing Articles</i>
▶ Citing Articles via CrossRef

[Current Issue](#) | [Archives](#) | [Online Submission](#) | [Info for Authors](#) | [Editorial Board](#) | [About](#)
[Subscribe](#) | [Advertise](#) | [Contact](#) | [Site Map](#)

[Copyright © 2008 by the National Academy of Sciences](#)

