Table S1. Primate cell lines utilized and cell culture methods

Common Name		Source		Cell type
Gorilla	*Gorilla gorilla*	Coriell	PR00280	Fibroblasts
Borneo Orangutan	*Pongo pygmaeus*	Coriell	PR00650	B-Lymphocyte
Bonobo	*Pan paniscus*	Coriell	PR00748	B-Lymphocyte
White-handed Gibbon	*Hylobates lar*	Coriell	PR01131	Fibroblasts
Pileated Gibbon	*Hylobates pileatus*	Coriell	PR00243	B-Lymphocyte
Yellow-Cheeked Gibbon	*Hylobates gabriellae*	Coriell	PR00381	B-Lymphocyte
Siamang	*Hylobates syndactylus*	Coriell	PR00722	Fibroblasts
White-Cheeked Gibbon	*Hylobates leucogeny*	Coriell	PR01037	Fibroblasts
Agile Gibbon	*Hylobates agilis*	Coriell	PR00773	Fibroblasts
Talapoin	*Miopithecus talapoin*	Coriell	PR00716	Fibroblasts
Colobus	*Colobus guereza*	Coriell	PR00980	Fibroblasts
Leaf Monkey	*Trachypithecus francoisi*	Coriell	PR01099	Fibroblasts
Olive Baboon	*Papio anubis*	Coriell	PR00978	Fibroblasts
Black Mangabey	*Lophocebus albigena*	Coriell	PR01215	Fibroblasts
Wolf's Guenon	*Cercopithecus wolfi*	Coriell	PR01241	Fibroblasts
Squirrel Monkey	*Saimiri sciureus*	Coriell	PR00603	Fibroblasts
Titi Monkey	*Callicebus cupreus*	Coriell	PR00793	Fibroblasts

METHODS: *DARC* was sequenced from these 17 primate species. Primary and immortalized primate cell lines (sources and individual primate identifiers are listed above) were grown in standard media supplemented with 15% fetal bovine serum at 37^{o}C and in 5% CO2. Total DNA was harvested from cell lines using the AllPrep DNA/RNA kit (Qiagen). PCR was performed from total genomic DNA with PCR SuperMix High Fidelity (Invitrogen) using the primers shown in Table S2.

Table S2. *DARC* PCR and sequencing primers

This table lists PCR and sequencing primers used to amplify and sequence the *DARC* gene from genomic DNA
(PCR primers), *sequencing primer
All samples were sequenced directly from pooled PCR products

Agile Gibbon	(AD469/AD470) AD473*, AD474* (KAT031/AD502) KAT035R*
Black Mangabey	(AD469/AD470) AD473*, AD474* (KAT031*/AD502) KAT035R*
Bonobo	(AD469/AD470) AD473*, AD474* (AD499/KAT038R*)
Borneo Orangutan	(AD469/AD470) AD473*, AD474* (KAT031*/ KAT038R*)
Colobus	(AD469/AD470) AD473*, AD474* (KAT031*/KAT035R*)
Gorilla	(AD469/AD470) AD473*, AD474* (KAT031*/KAT038R*), KAT032*, KAT038R*
Leaf Monkey	(AD469/AD470) AD473*, AD474* (KAT031*/KAT035R)
Olive Baboon	(AD469/AD470) AD473*, AD474* (KAT031/KAT035R*) KAT032*, KAT038R*
Pileated Gibbon	(AD469/AD470) AD473*, AD474* (KAT031*/KAT038R*)
Siamang	(AD469/AD470) AD473*, AD474* (KAT031*/KAT035R*) KAT038R*
Squirrel Monkey	(AD469/AD470*) AD473*, AD474* (KAT031*/ AD502*) KAT035*
Talapoin	(AD469/AD470) AD473*, AD474* (KAT031*/KAT035R*) AD500*
Titi Monkey	(AD469/AD470) AD473*, AD474* (KAT031*/AD501*) AD500*, KAT035R*
White-cheeked Gibbon	(AD469/AD470) AD473*, AD474* (KAT031*/KAT038R*)
White-handed Gibbon	(AD469/AD470) AD473*, AD474* (KAT031*/KAT038R*)
Wolfs Guenon	(AD469/AD470) AD473*, AD474* (KAT031*/KAT035R*) KAT038R*
Yellow-cheeked Gibbon	(AD469/AD470) AD473*, AD474* (AD499*/KAT038R*)

Primers used for amplification and sequencing of *DARC*

Primer Name	Sequence
AD469	CAGACAAAATAAGAAACCACCCGC
AD470	CTGTCCTCCCCTCCCACC
AD473	GGGCTACTGTGTCTGGTATGG
AD474	CCTTGGCTCCAAACAAACCC
AD499	CGGTAAAATGCCCACTTTCTGGTCCC
AD500	CAAATCCAACCTCAAAACAGG
AD501	CCATACCAGACACAGTAGCCC
AD502	GAGGCATGGCACCCTAGCAGC
KAT031	CTTCGGTAAAATGCCCACTTTCTGG
KAT032	GGTAAAATGCCCACTTTCTGGTCC
KAT035 R	GGAACTGAGTCAAAGGCAAAGG
KAT038 R	GGGAGAAAAGATGAGGAGGAAAAGG

Table S3. PAML methods and results, NSsites models, 35 species.

DARC[a] dataset	ω_0[c]	codon freq.[d]	*M1a-M2a* [b]		*M7-M8* [b]		*M8a-M8* [b]		tree length[f]	dN/dS (%)[g]	AA Positions of dN/dS>1[h] * p>0.95 ** p>0.99	
			$2\Delta\ell$ [e]	p-value	$2\Delta\ell$ [e]	p-value	$2\Delta\ell$ [e]	p-value			NEB	BEB
35 primate dataset	0.4	f61	8.9	p=0.012	9.2	p=0.010	8.6	p=0.003	1.18	3.65 (4.2%)	31G*	7R, 31G*, 68T*
	0.4	f3x4	7.0	p=0.030	9.4	p=0.009	7.0	p=0.008	1.16	3.70 (2.5%)	31G*	7R, 31G*, 68T
	1.6	f61	8.9	p=0.012	9.2	p=0.010	8.6	p=0.003	1.18	3.65 (4.2%)	31G*	7R, 31G*, 68T*
	1.6	f3x4	7.0	p=0.030	9.4	p=0.009	7.0	p=0.008	1.16	3.70 (2.5%)	31G*	7R, 31G*, 68T

[a] Dataset consisted of the aligned *DARC* sequences from *Homo sapiens, Pan troglodytes, Pan paniscus, Gorilla gorilla, Pongo pygmaeus* (Sumatran Orangutan), *Pongo pygmaeus* (Borneo Orangutan), *Hylobates lar, Hylobates pileatus, Hylobates syndactylus, Hylobates leucogenys, Hylobates agilis, Hylobates gabriellae, Macaca mulatta, Lophocebus albigena, Papio anubis, Miopithecus talapoin, Cercopithecus wolfi, Colobus guereza, Trachypithecus francoisi, Theropithecus gelada, Mandrillus sphinx, Cercocebus agilis, Allenopithecus nigroviridis, Cercopithecus mona, Saimiri sciureus, Callithrix jacchus, Saguinus midas, Pithecia pithecia, Chiropotes satanas, Ateles geoffroyi, Saguinus imperator, Aotus trivirgatus, Saimiri boliviensis, Cebus apella* and *Callicebus cupreus*.

[b] Maximum likelihood analysis was performed with codeml in the PAML 4.1 software package (Yang Z. 1997. Comput. Appl. Biosci. 13: 555–556). To detect selection, multiple alignments were fitted to the NSsites models M1a (null model, codon values of dN/dS are fit into two site classes, one with value between 0 and 1, and one fixed at dN/dS=1), M2a (positive selection model, similar to M1a but with an extra class of dN/dS > 1 allowed), M7 (null model, codon values of dN/dS fit to a beta distribution, dN/dS > 1 disallowed), M8a (null model, similar to M7 except with a fixed codon class of at dN/dS = 1) and M8 (positive selection model, similar to M7 but with an extra class of dN/dS >1 allowed). All three comparisons are between a null model and a model of positive selection. The p-value reflects the confidence with which the null model can be rejected.

[c] Initial seed value for ω (dN/dS) used in the maximum likelihood simulation

[d] Model of codon frequency

[e] Twice the difference in the natural logs of the likelihoods ($\Delta\ell$ x 2) of the two models being compared. This value is used in a likelihood ratio test along with the degrees of freedom. In all cases (M1a-M2a), (M7-M8), (M8a-M8), a model that allows positive selection is compared to a null model. The p-value indicates the confidence with which the null model can be rejected.

[f] The tree length is the number of substitutions per site along all branches in the phylogeny. It is calculated as the sum of the branch lengths, and is a representation of total diversity in the dataset

[g] dN/dS value of the class of codons evolving under positive selection in M8, and the percent of codons falling in that class.

[h] Amino acid positions identified in the class of codons evolving under positive selection in M8 with a posterior probability >0.90. Coordinates correspond to the human protein.

Table S4. Datamonkey methods and results, 35 species.

DARC dataset [a]	SLAC [b]	FEL [b]	REL [b]
35 primate dataset	68T	25V*, 68T*	31G, 68T*

[a] Dataset consisted of the aligned *DARC* sequences from *Homo sapiens, Pan troglodytes, Pan paniscus, Gorilla gorilla, Pongo pygmaeus* (Sumatran Orangutan), *Pongo pygmaeus* (Borneo Orangutan), *Hylobates lar, Hylobates pileatus, Hylobates syndactylus, Hylobates leucogenys, Hylobates agilis, Hylobates gabriellae, Macaca mulatta, Lophocebus albigena, Papio anubis, Miopithecus talapoin, Cercopithecus wolfi, Colobus guereza, Trachypithecus francoisi, Theropithecus gelada, Mandrillus sphinx, Cercocebus agilis, Allenopithecus nigroviridis, Cercopithecus mona, Saimiri sciureus, Callithrix jacchus, Saguinus midas, Pithecia pithecia, Chiropotes satanas, Ateles geoffroyi, Saguinus imperator, Aotus trivirgatus, Saimiri boliviensis, Cebus apella* and *Callicebus cupreus.*

[b] Positive selection analysis was performed using Datamonkey (Delport W et.al. 2010. Bioinformatics 26: 2455–2457), a web-based implementation of the HyPhy software package, using the single likelihood ancestor counting (SLAC), fixed effects likelihood (FEL), and random effects likelihood (REL) tests. SLAC tests for positive selection using a modified version of the Suzuki–Gojobori counting approach. The FEL test uses a likelihood-based approach, in which dS and dN are evaluated at each site based on a codon-substitution model. The REL test for positive selection is a variant of the likelihood methods used in PAML (Yang Z. 1997. Comput. Appl. Biosci. 13: 555–556). Codon positions identified as evolving under positive selection with a p-value <0.10 (*<0.05) are listed for the SLAC and FEL analyses, and sites with a posterior probability >0.90 (* >0.95) are listed for the REL analysis. Coordinates correspond to the human protein.

Table S5. PAML methods and results, NSsites models, hominoid species only.

DARC[a] dataset	ω_0[c]	codon freq.[d]	*M1a-M2a* [b]		*M7-M8* [b]		*M8a-M8* [b]		tree length[f]	dN/dS (%)[g]	AA Positions of dN/dS>1[h] * p>0.95 ** p>0.99	
			$2\Delta\ell$[e]	p-value	$2\Delta\ell$[e]	p-value	$2\Delta\ell$[e]	p-value			NEB	BEB
13 primate dataset	0.4	f61	11.9	p=0.003	13.1	p=0.001	11.8	p<0.001	0.36	10.0 (2.7%)	R7**, 25V, 197L	R7**, 25V*, 197L*
	0.4	f3x4	18.9	p<0.001	19.0	p<0.001	18.9	p<0.001	0.40	28.0 (0.9%)	R7**, 25V, 197L*	R7**, 25V*, 197L**
	1.6	f61	11.9	p=0.003	11.9	p=0.003	11.8	p<0.001	0.36	10.0 (2.7%)	R7**, 25V, 197L	R7**, 25V*, 197L*
	1.6	f3x4	18.9	p<0.001	19.0	p<0.001	18.9	p<0.001	0.40	28.0 (0.9%)	R7**, 25V, 197L*	R7**, 25V*, 197L**

[a] Dataset consisted of the aligned *DARC* sequences from *Homo sapiens, Pan troglodytes, Pan paniscus, Gorilla gorilla, Pongo pygmaeus* (Sumatran Orangutan)*, Pongo pygmaeus* (Borneo Orangutan)*, Hylobates lar, Hylobates pileatus, Hylobates syndactylus, Hylobates leucogenys, Hylobates agilis, Hylobates gabriellae,* and *Macaca mulatta* (as an outgroup).

[b] Maximum likelihood analysis was performed with codeml in the PAML 4.1 software package (Yang Z. 1997. Comput. Appl. Biosci. 13: 555–556). To detect selection, multiple alignments were fitted to the NSsites models M1a (null model, codon values of dN/dS are fit into two site classes, one with value between 0 and 1, and one fixed at dN/dS=1), M2a (positive selection model, similar to M1a but with an extra class of dN/dS > 1 allowed), M7 (null model, codon values of dN/dS fit to a beta distribution, dN/dS > 1 disallowed), M8a (null model, similar to M7 except with a fixed codon class of at dN/dS = 1) and M8 (positive selection model, similar to M7 but with an extra class of dN/dS >1 allowed). All three comparisons are between a null model and a model of positive selection. The p-value reflects the confidence with which the null model can be rejected.

[c] Initial seed value for ω (dN/dS) used in the maximum likelihood simulation

[d] Model of codon frequency

[e] Twice the difference in the natural logs of the likelihoods ($\Delta\ell$ x 2) of the two models being compared. This value is used in a likelihood ratio test along with the degrees of freedom. In all cases (M1a-M2a), (M7-M8), (M8a-M8), a model that allows positive selection is compared to a null model. The p-value indicates the confidence with which the null model can be rejected.

[f] The tree length is the number of substitutions per site along all branches in the phylogeny. It is calculated as the sum of the branch lengths, and is a representation of total diversity in the dataset

[g] dN/dS value of the class of codons evolving under positive selection in M8, and the percent of codons falling in that class.

[h] Amino acid positions identified in the class of codons evolving under positive selection in M8 with a posterior probability >0.90. Coordinates correspond to the human protein.

Table S6. Branch-site test for positive selection in the hominoid clade

branch-site model[a]	estimate of parameters [b]		*Test 2* $2\Delta\ell$ [c]	p-value	AA Positions of dN/dS>1[d]
Model A with ω_2 fixed at 1	ℓ = -3625.75	p_0 = 0.321 p_1=0.298 p_{2a}+p_{2b}=0.38 ω_0 = 0.017 ω_1 = 1.000 ω_2 = 1.000	16.2	p<0.001	7R**, 25V**, 197L**
Model A	ℓ = -3617.65	p_0 = 0.520 p_1=0.437 p_{2a}+p_{2b}=0.04 ω_0 = 0.082 ω_1 = 1.000 ω_2=11.3			

[a] Dataset consisted of the aligned *DARC* sequences from *Homo sapiens, Pan troglodytes, Pan paniscus, Gorilla gorilla, Pongo pygmaeus* (Sumatran Orangutan), *Pongo pygmaeus* (Borneo Orangutan), *Hylobates lar, Hylobates pileatus, Hylobates syndactylus, Hylobates leucogenys, Hylobates agilis, Hylobates gabriellae, Macaca mulatta, Lophocebus albigena, Papio anubis, Miopithecus talapoin, Cercopithecus wolfi, Colobus guereza, Trachypithecus francoisi, Theropithecus gelada, Mandrillus sphinx, Cercocebus agilis, Allenopithecus nigroviridis, Cercopithecus mona, Saimiri sciureus, Callithrix jacchus, Saguinus midas, Pithecia pithecia, Chiropotes satanas, Ateles geoffroyi, Saguinus imperator, Aotus trivirgatus, Saimiri boliviensis, Cebus apella* and *Callicebus cupreus*. The hominoid clade (the first 12 primates listed) were defined as the foreground clade in the models.

[b] To implement the branch-sites test (Zhang J, Nielsen R, and Yang Z. 2005. Mol Biol Evol 22: 2472–2479), multiple alignments were fitted to the branch-sites Model A (positive selection model, codon values of dN/dS along background branches are fit into two site classes, one (ω0) between 0 and 1 and one (ω1) equal to 1, on the foreground branches a third site class is allowed (ω2) with dN/dS > 1), and Model A with fixed ω2 = 1 (null model, similar to Model A except the foreground ω2 value is fixed at 1). The branch-sites model for positive selection (Model A) allows certain codons to evolve with dN/dS > 1 exclusively along the lineages of the foreground clade. p2a + p2b is the sum of the proportion of sites that fall into the categories of background-branch < 1/foreground branch > 1 (p2a) and background branch =1/foreground branch >1 (p2b). Models were run using the f61 codon model and an initial seed value of ω=1.5.

[c] A likelihood ratio test (branch-site test 2) was performed to assess whether permitting codons to evolve under positive selection in the hominoid clade (Model A) gives a significantly better fit to the data than disallowing this (null model; Model A with ω_2 fixed at 1). Twice the difference in the natural logs of the likelihoods ($\Delta\ell$ x 2) of the two models being compared is shown. This value is used in a likelihood ratio test along with the degrees of freedom (1 in this case). The p-value indicates the confidence with which the null model can be rejected.

[d] Residue positions identified in the class of codons evolving under positive selection with a posterior probability > 0.90 (* p>0.95 and ** p>0.99). Coordinates correspond to the human protein.