**Supplemental Materials**

Relaxed selection on ultraconservation following widespread ultraconservation divergence in primates

To measure the relative level of purifying selection acting on UCEs versus the rest of the conserved, non-repetitive human DNA, [which have been also previously shown to evolve under purifying selection (Waterston et al. 2002; Drake et al. 2006)]. I compared corresponding DAF distributions (Figure S1). It was observed that the purifying pressure acting on UCEs/sUCEs is much stronger than the purifying pressure acting on other conserved non-repetitive sequences in the human genome. Similarly, DAF distributions of UCEs and sUCEs were significantly different from the corresponding DAF distributions of the remaining fraction of noncoding conserved human DNA (P<<0.01) with the exception of ASN SNPs in UCEs, for which statistical significance was not reached. Thus, despite observed rapid eradication of ultraconservation in primates, UCEs/sUCEs are under highly elevated purifying selection and their divergence is the result of active predominant mutational processes.

Figure S1. Derived allele frequency (DAF) distributions comparing allelic adaptation in UCEs (light gray) and sUCEs (white) to non-repetitive human genome DNA (black) and non-repetitive evolutionary conserved DNA (light blue; defined as 100bps/70% identity between human and mouse; generated by ECR Browser(Ovcharenko et al. 2004)). YRI population (A), CEU population (B), ASN populations (C). Standard deviation (σ) in UCE/sUCE DAF binning (B-D) was estimated using the Binomial distribution as $\sigma^2 = np(1-p)$, where p is the fraction of SNPs in a particular bin, and n is the total number of SNPs in that distribution.
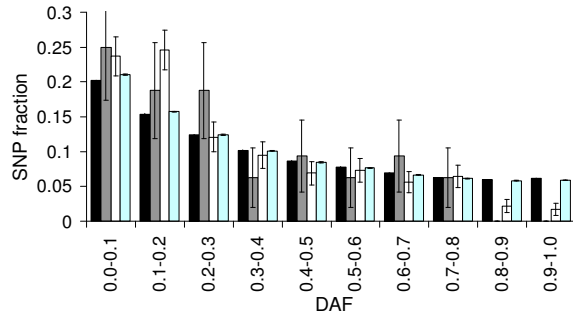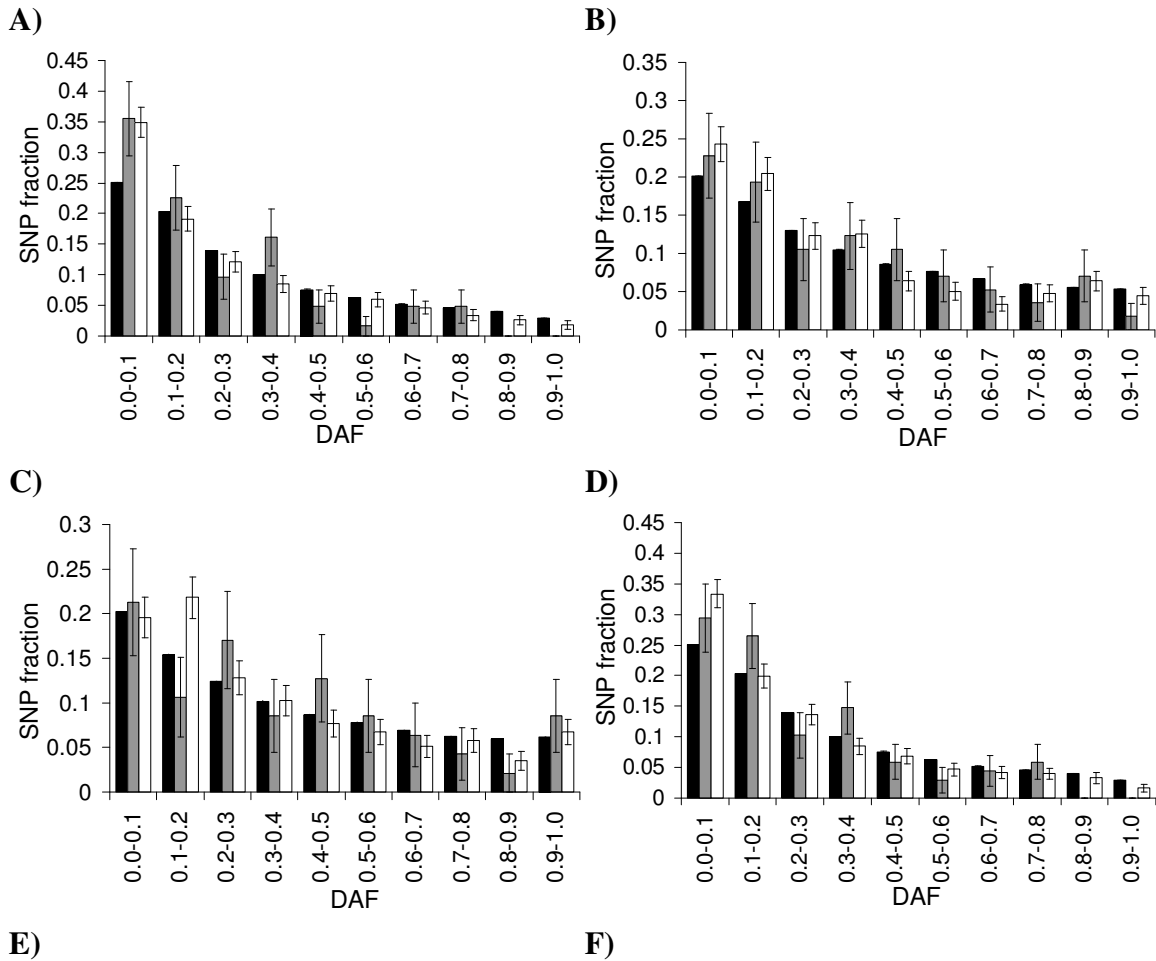
A)

B)

C)
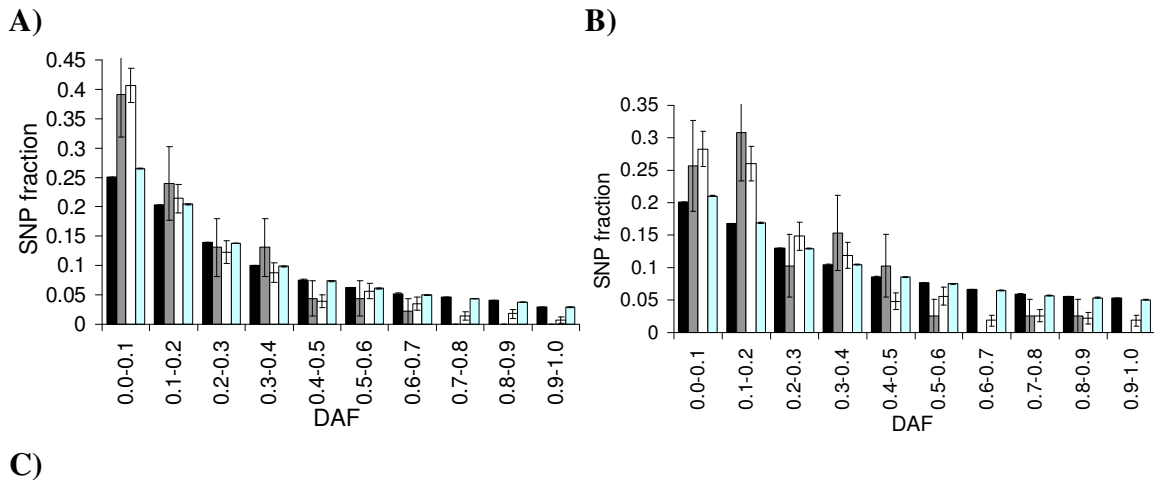
Figure S2. Derived allele frequency (DAF) distributions comparing allelic adaptation in human homologs of C/M UCEs (light gray) and sUCEs (white) to non-repetitive human genome DNA (black); YRI population (A), CEU population (B), ASN populations (C). Analogous DAF distributions for R/M data – YRI population (D), CEU population (E), ASN populations (F).

A)

B)

C)

D)

E)

F)
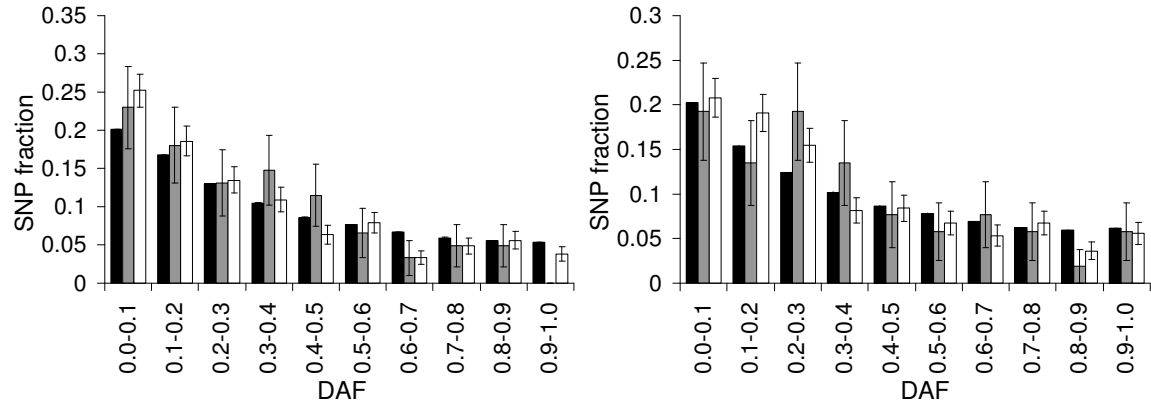
Figure S3. HapMap sUCE SNPs binned into different genic categories (full set of SNPs in gray). The 0.2 DAF threshold was used to separate ancient (black) from derived (white) alleles. DAF indicators were averaged across 3 populations for individual SNPs.
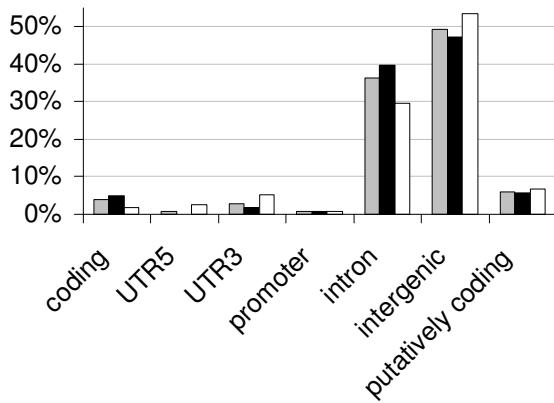

Figure S4. Binning of H/M UCE and sUCE by genic categories. Comparing elements with SNPs to the original distribution. Statistically significant difference is marked by asterisk.
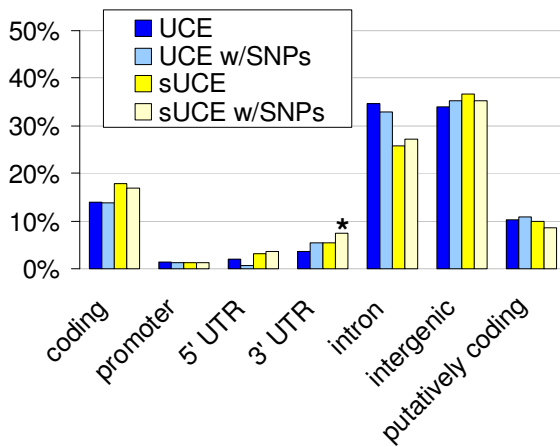
Table S1. Binning of UCEs and sUCEs into coding, promoter, UTR, intronic, interegenic, and putatively coding categories.

	UCE			sUCE		
	Chimp	**Rhesus**	**Human**	**Chimp**	**Rhesus**	**Human**
3-species match	656	696	635	6,701	6,885	6,867
Coding	12.8%	13.2%	14.0%	16.9%	17.0%	17.9%
Promoter	1.1%	1.3%	1.4%	1.4%	1.2%	1.3%
5' UTR	2.3%	2.6%	2.0%	2.8%	2.8%	3.1%
3' UTR	3.5%	3.4%	3.6%	5.1%	5.2%	5.4%
Intron	35.1%	32.6%	34.6%	26.8%	26.6%	25.8%
Intergenic	35.2%	35.8%	34.0%	37.3%	37.6%	36.6%
Putatively coding	10.1%	11.1%	10.2%	9.8%	9.6%	9.9%

Table S2. Putatively coding UCEs and sUCEs in introns and intergenic intervals.

	UCE			sUCE		
	Chimp	**Rhesus**	**Human**	**Chimp**	**Rhesus**	**Human**
Putatively coding	66	77	65	658	664	677
Putatively coding in introns	68.2%	68.8%	67.7%	64.6%	63.7%	64.0%
Putatively coding in intergenic intervals	31.8%	31.2%	32.3%	35.4%	36.3%	36.0%

Table S3. Nonsynonymous mutations in H/M sUCEs. Affected genes include HOXA7, CNOT4, CALU, HECTD1, EFNA3, SCN2A2, and SNX10.

Population	sUCE location	DAF	SNP	Gene name	Substitution
ASN	chr7:128181801-128181911	0.039	A/G	**CALU**	A:V:non
YRI	chr7:128181801-128181911	0.092	A/G	**CALU**	A:V:non
CEU	chr7:128181801-128181911	0.3	A/G	**CALU**	A:V:non
YRI	chr7:134699202-134699351	0.225	C/T	**CNOT4**	V:I:non
CEU	chr7:134699202-134699351	0.233	C/T	**CNOT4**	V:I:non
ASN	chr7:134699202-134699351	0.466	C/T	**CNOT4**	V:I:non
CEU	chr1:153325191-153325301	0.158	A/G	**EFNA3**	V:M:non
CEU	chr14:30667978-30668128	0.119	A/G	**HECTD1**	F:L:non
ASN	chr14:30667978-30668128	0.191	A/G	**HECTD1**	F:L:non
YRI	chr7:27162631-27162745	0.067	C/T	**HOXA7**	A:T:non
JTP_CHB	chr7:27162631-27162745	0.843	C/T	**HOXA7**	A:T:non
CEU	chr7:27162631-27162745	0.85	C/T	**HOXA7**	A:T:non
YRI	chr2:165860589-165860707	0.042	A/G	**SCN2A2**	R:K:non
CEU	chr2:165860589-165860707	0.085	A/G	**SCN2A2**	R:K:non
ASN	chr2:165860589-165860707	0.089	A/G	**SCN2A2**	R:K:non
ASN	chr7:26378557-26378677	0.011	G/T	**SNX10**	S:I:non

Table S4. Primate UCEs and short UCEs (sUCEs). SNP annotation in H/M elements and in HHs of C/M and R/M elements was done using the human dbSNP data (Sherry, Ward, and Sirotkin 1999).

	UCEs			sUCEs		
	Human	**Chimp**	**Rhesus**	**Human**	**Chimp**	**Rhesus**

Total	635	653	695	6867	6692	6875
Diverged with SNPs	n/a	12.1%	14.8%	n/a	5.2%	7.8%
Diverged without SNPs	n/a	14.1%	23.5%	n/a	11.1%	21.9%
Identical with SNPs	26.3%	16.8%	15.1%	17.2%	14.0%	11.5%
Identical without SNPs	73.7%	57.0%	46.6%	82.8%	69.6%	58.7%
Total diverged and identical with SNPs	**26.3%**	**43.0%**	**53.4%**	**17.2%**	**30.4%**	**41.3%**

Table S5. HapMap SNP density per kilobase of genome sequence averaged across YRI, CEU and ASN populations. Genome average corresponds to non-repetitive DNA sequence.

UCEs			sUCE			Genome average
H/M	**C/M**	**R/M**	**H/M**	**C/M**	**R/M**	
0.24	0.32	0.34	0.27	0.37	0.41	1.20

Table S6. Fraction of SNPs with the DAF < 0.2. Regular conserved subset corresponds to evolutionary conserved regions in the human genome obtained using standard 70% identity, 100 bps length thresholds.

	YRI	CEU	ASN
H/M UCEs	0.63	0.56	0.43
HHs of C/M UCEs	0.58	0.42	0.32
HHs of R/M UCEs	0.56	0.41	0.33
H/M sUCEs	0.62	0.54	0.48
HHs of C/M sUCEs	0.54	0.45	0.41
HHs of R/M sUCEs	0.53	0.44	0.40
regular conserved	0.47	0.38	0.37
non-repetitive genome	0.45	0.37	0.36

REFERENCES:

Drake, J. A., C. Bird, J. Nemesh, D. J. Thomas, C. Newton-Cheh, A. Reymond, L. Excoffier, H. Attar, S. E. Antonarakis, E. T. Dermitzakis, and J. N. Hirschhorn. 2006. Conserved noncoding sequences are selectively constrained and not mutation cold spots. Nat Genet **38**:223-227.

Ovcharenko, I., M. A. Nobrega, G. G. Loots, and L. Stubbs. 2004. ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. Nucleic Acids Res **32**:W280-286.

Sherry, S. T., M. Ward, and K. Sirotkin. 1999. dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. Genome Res **9**:677-679.

Waterston, R. H.K. Lindblad-TohE. BirneyJ. RogersJ. F. AbrilP. AgarwalR. AgarwalaR. AinscoughM. AlexanderssonP. AnS. E. AntonarakisJ. AttwoodR. BaertschJ. BaileyK. BarlowS. BeckE. BerryB. BirrenT. BloomP. BorkM. BotcherbyN. BrayM. R. BrentD. G. BrownS. D. BrownC. BultJ. BurtonJ. ButlerR. D. CampbellP. CarninciS. CawleyF. ChiaromonteA. T. ChinwallaD. M. ChurchM. ClampC. CleeF. S. CollinsL. L. CookR. R. CopleyA. CoulsonO. CouronneJ. CuffV. CurwenT. CuttsM. DalyR. DavidJ. DaviesK. D. DelehauntyJ. DeriE. T. DermitzakisC. DeweyN. J. DickensM. DiekhansS. DodgeI. DubchakD. M. DunnS. R. EddyL. ElnitskiR. D. EmesP. EswaraE. EyrasA. FelsenfeldG. A. FewellP. FlicekK. FoleyW. N. FrankelL. A. FultonR. S. FultonT. S. FureyD. GageR. A. GibbsG. GlusmanS. GnerreN. GoldmanL. GoodstadtD. GrafhamT. A. GravesE. D. GreenS. GregoryR. GuigoM. GuyerR. C. HardisonD. HausslerY. HayashizakiL. W. HillierA. HinrichsW. HlavinaT. HolzerF. HsuA. HuaT. HubbardA. HuntI. JacksonD. B. JaffeL. S. JohnsonM. JonesT. A. JonesA. JoyM. KamalE. K. KarlssonD. KarolchikA. KasprzykJ. KawaiE. KeiblerC. KellsW. J. KentA. KirbyD. L. KolbeI. KorfR. S. KucherlapatiE. J. KulbokasD. KulpT. LandersJ. P. LegerS. LeonardI. LetunicR. LevineJ. LiM. LiC. LloydS. LucasB. MaD. R. MaglottE. R. MardisL. MatthewsE. MauceliJ. H. MayerM. McCarthyW. R. McCombieS. McLarenK. McLayJ. D. McPhersonJ. MeldrimB. MeredithJ. P. MesirovW. MillerT. L. MinerE. MonginK. T. MontgomeryM. MorganR. MottJ. C. MullikinD. M. MuznyW. E. NashJ. O. NelsonM. N. NhanR. NicolZ. NingC. NusbaumM. J. O'ConnorY. OkazakiK. OliverE. Overton-LartyL. PachterG. ParraK. H. PepinJ. PetersonP. PevznerR. PlumbC. S. PohlA. PoliakovT. C. PonceC. P. PontingS. PotterM. QuailA. ReymondB. A. RoeK. M. RoskinE. M. RubinA. G. RustR. SantosV. SapojnikovB. SchultzJ. SchultzM. S. SchwartzS. SchwartzC. ScottS. SeamanS. SearleT. SharpeA. SheridanR. ShownkeenS. SimsJ. B. SingerG. SlaterA. SmitD. R. SmithB. SpencerA. StabenauN. Stange-ThomannC. SugnetM. SuyamaG. TeslerJ. ThompsonD. TorrentsE. TrevaskisJ. TrompC. UclaA. Ureta-VidalJ. P. VinsonA. C. Von NiederhausernC. M. WadeM. WallR. J. WeberR. B. WeissM. C. WendlA. P. WestK. WetterstrandR. WheelerS. WhelanJ. WierzbowskiD. WilleyS. WilliamsR. K. WilsonE. WinterK. C. WorleyD. WymanS. YangS. P. YangE. M. ZdobnovM. C. Zody, and E. S. Lander. 2002. Initial sequencing and comparative analysis of the mouse genome. Nature **420**:520-562.