Supplemental Materials

THE RISE AND FALL OF THE CHEMOATTRACTANT RECEPTOR GPR33

Holger Römpler[1], Angela Schulz[1], Christian Pitra[2], Graham Coop[3], Molly Przeworski[4], Svante Pääbo[5], and Torsten Schöneberg[1]

From the [1]Institute of Biochemistry, Molecular Biochemistry, Medical Faculty, University of Leipzig, Leipzig, Germany; [2]Institute for Zoo and Wildlife Research, Department of Evolutionary Genetics, Berlin, Germany; [3]Department of Statistics, University of Oxford, Oxford, United Kingdom; [4]Department of Ecology and Evolutionary Biology, Brown University, Providence, USA; [5]Max-Planck-Institute of Evolutionary Anthropology, Leipzig, Germany

Running title: GPR33 evolution

Address correspondence to: Torsten Schöneberg, Institute of Biochemistry, Molecular Biochemistry (Max-Planck-Institute Interim), Medical Faculty, University of Leipzig, Deutscher Platz 6, 04103 Leipzig, Germany, Tel.: +49-341-3550-850, Fax: +49-341-3550-855, E-mail: schoberg@medizin.uni-leipzig.de

RESULTS

Nonsense mutation of GPR33 causes truncation of the receptor protein

Because all GPR33 pseudogenes found in humans and apes were due to stop codons we addressed the question whether the stop codons indeed terminates the translation of the receptor proteins. An increasing number of studies showed that not all stop codons are equal and some allow a leaky translation mainly in plants and yeast (44-46). To test whether the stop codon within the human GPR33 pseudogene leads to a truncated protein the receptor was N- and C-terminally epitope tagged with an HA-epitope and a FLAG-epitope, respectively. Immunofluorescence studies revealed the presence of the N-terminal epitope, however, the C terminus could not be detected (Suppl. Fig. S1). Similar results were obtained with the pseudogenes of chimpanzee and orangutan (data not shown). Interestingly, aminoglycoside antibiotics are able to suppress premature stop codons, thereby permitting protein translation to continue to the normal end of the gene. This phenomenon is most likely due to the interaction of the aminoglycosides with ribosomes, reducing the usual stringency of codon-anticodon pairing (47). Recent *in vivo* studies have shown, for example, that aminoglycoside antibiotics can suppress premature stop codons in the cystic fibrosis transmembrane regulator, dystrophin and the V2 vasopressin receptor genes (48-50). Similarly, COS-7 cells transiently transfected with the human GPR33 regained the ability to produce a full length receptor following incubation with the aminoglycoside geneticin (Suppl. Fig. S1).

Human and rat GPR33 pseudogenes are expressed

To gain more information beyond structural relations, we set out to clarify the expression pattern of GPR33 in mouse tissue by RT-PCR. The mouse GPR33 is a single copy gene at chromosome 12 (12B3) and the structure of the mRNA was unknown. To avoid false positive results in the RT-PCR analysis with primers derived from the coding region because of genomic contamination of the cDNA we first analyzed the structure of the mouse GPR33 transcript. Thus, 5' rapid amplification of cDNA ends (RACE) PCR was performed with a cDNA library from mouse spleen (Clontech, Palo Alto, CA). PCR fragments were cloned and sequenced. In several clones the coding region of the mouse GPR33 was 5' joined with noncoding DNA sequence which matched with the genomic mouse database sequence (Acc.

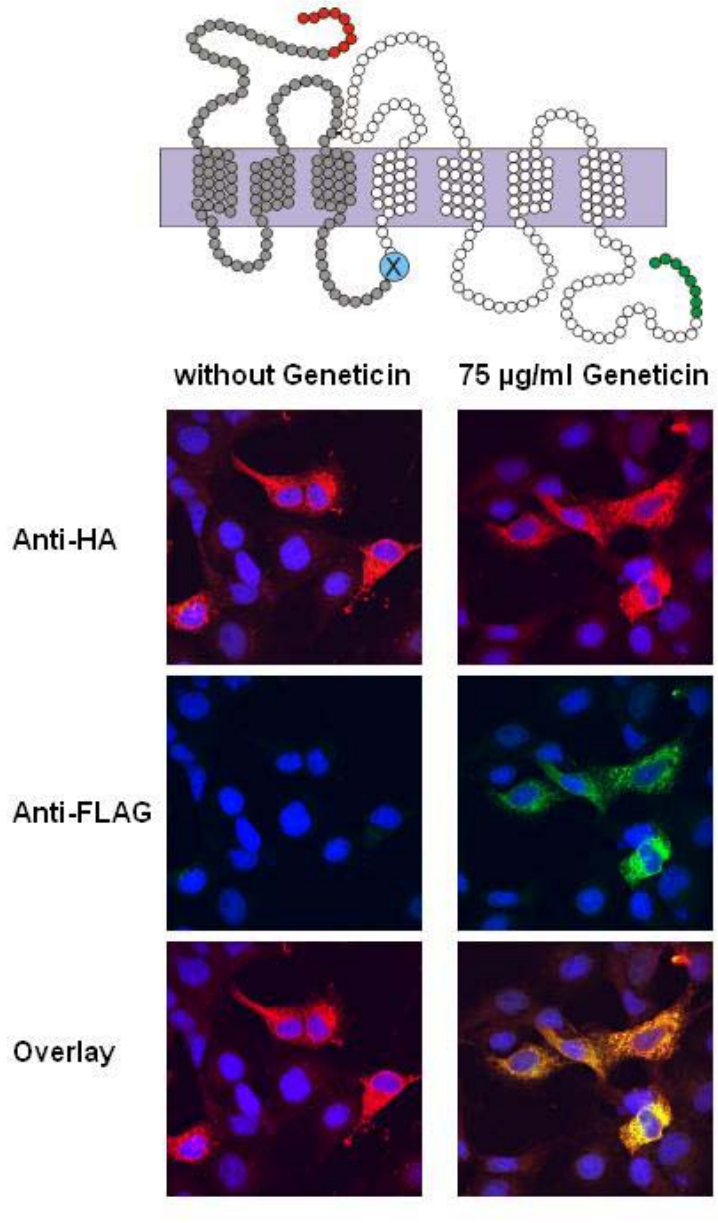
number NT_039551) about 3.6 kb upstream the start ATG of GPR33. The intron of the 5'-UTR is flanked by a classical consensus sequence. The sequence of the GPR33 5'-UTR was verified by direct genomic sequencing. For RT-PCR a primer pair was chosen which flanks the intron and a 413-bp PCR product is obtained only when the exons are properly spliced. As shown in Fig. 4A, GPR33 is mainly expressed in mouse spleen, lung and testis. Then, several murine cell lines were analyzed for GPR33 expression by RT-PCR. In consensus with a predominant expression in immunologically relevant tissues GPR33 transcripts were identified in RAW 264.7 cells, a murine macrophage cell line.

Next, we asked whether GPR33 pseudogenes are also transcribed. Comparison of the rat genomic sequence with the mouse GPR33 locus revealed high sequence similarities in the coding region and the 5'-UTR. Similarly, primers were designed to amplify a 284-bp fragment from spleen cDNA of *R. norvegicus*. As found for the mouse GPR33 a specific product was amplified and its identity was verified by sequencing.

Except of the coding sequence and the splice acceptor site (5' of the translation start ATG), no similarities are obvious when the mouse and the human GPR33 locus were compared in the 5' region. Therefore, primers derived from the coding region had to be used in the human RT-PCR. As shown in Fig. 4B, the human GPR33 transcript was highly expressed in spleen and lung but also in heart, liver, kidney, pancreas, thymus, gonads, and leucocytes. The tissue panel was controlled for genomic contamination by using the identical antisense primer together with a sense primer directly upstream of the putative splice acceptor site. Although a specific PCR product was amplified with human genomic DNA as control, all samples of the cDNA panel revealed no such product indicating that there is no contamination. Sequence analysis of rat and human pseudogene cDNA revealed no evidence for exon skipping or intergenic splicing which may produce fusion proteins with neighbored genes as found e.g. for $P2Y_{11}$ and SSF1 (51).

REFERENCES (Suppl.)

44. Skuzeski, J. M., Nichols, L. M., Gesteland, R. F. and Atkins, J. F. (1991) *J. Mol. Biol.* **218**, 365-373
45. Namy, O., Hatin, I. and Rousset, J. P. (2001) Impact of the six nucleotides downstream of the stop codon on translation termination. *EMBO Rep.* **2**, 787-793
46. Keeling, K. M., Lanier, J., Du, M., Salas-Marco, J., Gao, L., Kaenjak-Angeletti, A. and Bedwell, D. M. (2004) *RNA* **10**, 691-703
47. Mankin, A. S. and Liebman, S. W. (1999) *Nat. Genet*. **23**, 8-10
48. Barton-Davis, E. R., Cordier, L., Shoturma, D. I., Leland, S. E. and Sweeney, H. L. (1999) *J. Clin. Invest*. **104**, 375–381
49. Wilschanski, M., Yahav, Y., Yaacov, Y., Blau, H., Bentur, L., Rivlin, J., Aviram, M., Bdolah-Abram, T., Bebok, Z., Shushi, L., Kerem, B. and Kerem, E. (2003) *N. Engl. J. Med*. **349**, 1433-1441
50. Sangkuhl, K., Schulz, A., Römpler, H., Yun, J., Wess, J. and Schöneberg, T. (2004) *Hum. Mol. Genet.* **13**, 893-903
51. Communi, D., Suarez-Huerta, N., Dussossoy, D., Savi, P., Boeynaems, J. M. (2001) *J. Biol. Chem.* **276**, 16561-16566



Suppl. Figure S1. ***Reconstitution of the polypeptide chain of the human GPR33 pseudogene.*** Aminoglycosides are able to suppress premature stop codons, thereby permitting protein translation to continue to the normal end of the gene. To examine the ability of geneticin to suppress the stop codon 140, COS-7 cells were transfected with expression plasmids coding for the human GPR33 Stop140. Transfected COS-7 cells were incubated with or without 75 µg geneticin/ml one day prior to immunostaining. Permeabilized cells were incubated with a monoclonal anti-HA-antibody (red) and a polyclonal antibody against the FLAG epitope (green). The primary antibodies were detected using species specific TRITC- and FITC-labeled secondary antibodies. Nuclei were stained with DAPI (blue). Confocal immunofluorescence pictures are representative of three independent experiments.

Suppl. Table S1
Accession numbers, description and sources of genomic DNA samples

Abbreviations: NT, N terminus; CT, C terminus; TMD, transmembrane domain, IL, intracellular loop

accession number	common name	species	cloned sequence	Source
Primates				
AY490588 AY490589 AY493998 AY493999 AY494000	human	*Homo sapiens*	NT-CT	Dr M. Kayser, MPI, Leipzig, Germany Dr M. Stoneking, MPI, Leipzig, Germany D. Sere, MPI, Leipzig, Germany own source
AY490569 AY490570	gorilla	*Gorilla gorilla*	NT-CT	Dr J. Bullerdiek, University of Bremen Dr T. Haaf, MPI Molecular Genetics Berlin, Germany Dr J. Rhyne, University of Maryland, Baltimore, USA Dr M. Rocchi, University of Bari, Italy Dr L. Vigilant, MPI, Leipzig, Germany
AY490571 AY490572	siamang	*Hylobates syndactylus*	NT-CT	Dipl. Biol. C. Roos, Primate Center Goettingen, Germany Dr T. Haaf, MPI Molecular Genetics Berlin, Germany
AY490573	common gibbon	*Hylobates lar*	NT-CT	Dipl. Biol. C. Roos, Primate Center Goettingen, Germany Dr W. Enard, MPI, Leipzig, Germany Dr J. Bullerdiek, University of Bremen Dr M. Rocchi, University of Bari, Italy
AY490574	red-cheeked gibbon	*Hylobates gabriellae*	NT-CT	Dipl. Biol. C. Roos, Primate Center Goettingen, Germany
AY490575	white-cheeked gibbon	*Hylobates leucogenys*	NT-CT	Dipl. Biol. C. Roos, Primate Center Goettingen, Germany
AY490576 AY490577	white-tufted-ear marmoset	*Callithrix jacchus*	NT-CT	Dr T. Haaf, MPI Molecular Genetics Berlin, Germany Dr W. Enard, MPI, Leipzig, Germany
AY490578	Geoffroy's marmoset	*Callithrix geoffroyi*	NT-CT	Dr J. Rhyne, University of Maryland, Baltimore, USA
AY490579	silvered langur	*Presbytis cristata*	NT-CT	Dr T. Haaf, MPI Molecular Genetics Berlin, Germany

AY490580 AY490581 AY490582 AY490583 AY490584 AY490585 AY490586 AY493989 AY493990 AY493991 AY493992 AY493993 AY493994 AY493995	chimpanzee	*Pan troglodytes*	NT-CT	Dr T. Haaf, MPI Molecular Genetics Berlin, Germany Dr M. Rocchi, University of Bari, Italy Dr W. Enard, MPI, Leipzig, Germany
AY490587 AY493996 AY493997	pygmy chimpanzee, bonobo	*Pan paniscus*	NT-CT	Dr W. Enard, MPI, Leipzig, Germany Dr L. Vigilant, MPI, Leipzig, Germany
AY490590	hamadryas baboon	*Papio hamadryas*	NT-CT	Dr M. Rocchi, University of Bari, Italy Dr T. Haaf, MPI Molecular Genetics Berlin, Germany
AY490591	rhesus monkey	*Macaca mulatta*	NT-CT	Dr W. Enard, MPI, Leipzig, Germany Dr M. Rocchi, University of Bari, Italy Dr J. Rhyne, University of Maryland, Baltimore, USA
AY490592 AY490593 AY490594	orangutan	*Pongo pygmaeus*	NT-CT	Dr J. Rhyne, University of Maryland, Baltimore, USA Dr M. Rocchi, University of Bari, Italy Dr W. Enard, MPI, Leipzig, Germany
AY490595	potto	*Perodicticus potto*	NT-IL3	Dr T. Haaf, MPI Molecular Genetics Berlin, Germany
AY490596 AY490597	emperor tamarin	*Saguinus imperator*	NT-CT	Dr J. Rhyne, University of Maryland, Baltimore, USA
AY490598	ruffed lemur	*Varecia variegata*	NT-CT	Dr W. Enard, MPI, Leipzig, Germany
AY490599	ring-tailed lemur	*Lemur catta*	NT-CT	Dr W. Enard, MPI, Leipzig, Germany Dr J. Rhyne, University of Maryland, Baltimore, USA
AY490600	mandrill	*Mandrillus sphinx*	NT-CT	Dr W. Enard, MPI, Leipzig, Germany
AY490601 AY490602	common squirrel monkey	*Saimiri sciureus*	NT-CT	Dr W. Enard, MPI, Leipzig, Germany
AY490603	red guenon	*Erythrocebus patas*	NT-CT	Dr J. Rhyne, University of Maryland, Baltimore, USA
AY490604 AY490605	Azara's night monkey	*Aotus azarai*	NT-CT	Dr W. Enard, MPI, Leipzig, Germany
AY490606	brown capuchin	*Cebus apella*	NT-CT	Dr W. Enard, MPI, Leipzig, Germany

AY490607	guereza	*Colobus guereza*	NT-CT	Dr W. Enard, MPI, Leipzig, Germany
AY490608	slow loris	*Nycticebus bengalensis*	NT-CT	Dr W. Enard, MPI, Leipzig, Germany
AY490609 AY490610	South African galago	*Galago moholi*	NT-CT	Dr W. Enard, MPI, Leipzig, Germany
AY490611 AY490612	brown-headed spider monkey	*Ateles fusciceps*	NT-CT	Dr W. Enard, MPI, Leipzig, Germany
AY490613	African green monkey	*Cercopithecus aethiops*	NT-CT	COS-7, ATCC – cell collection
Rodentia				
AY490614	European woodmouse	*Apodemus sylvaticus*	NT-CT	Dr G. Gerlach, University of Konstanz, Germany
AY490617 AY490618 AY490619 AY490620	wood mice	*Apodemus sp.*	NT-CT	Dr G. Gerlach, University of Konstanz, Germany V. Römpler, Erfurt, Germany
AY490615 AY490616	yellow-necked field mouse	*Apodemus flavicollis*	NT-CT	Dr H. Meisel, Humboldt University Berlin, Germany
AY490621	sibling vole	*Microtus epiroticus*	NT-CT	Dr G. Gerlach, University of Konstanz, Germany
AY490622 AY490623	common vole	*Microtus arvalis*	NT-CT	Dr H. Meisel, Humboldt University Berlin, Germany
AY490624 AY490625	bank vole	*Clethrionomys glareolus*	NT-CT	Dr G. Gerlach, University of Konstanz, Germany
AY490626	eastern European house mouse	*Mus musculus musculus*	NT-CT	Dr A. Orth, University of Montpellier, France
AY490627	Macedonian mouse	*Mus macedonicus*	NT-CT	Dr A. Orth, University of Montpellier, France
AY490628	southeastern Asian house mouse	*Mus musculus castaneus*	NT-CT	Dr A. Orth, University of Montpellier, France
AY490629	house mouse	*Mus musculus sp.*	NT-CT	Dr A. Orth, University of Montpellier, France
AY490630	Japanese mice	*Mus musculus molossinus*	NT-CT	Dr A. Orth, University of Montpellier, France
AY490631	western wild mouse	*Mus spretus*	NT-CT	Dr A. Orth, University of Montpellier, France
AY490632	steppe mouse	*Mus spicilegus*	NT-CT	Dr A. Orth, University of Montpellier, France
AY490633	Ryukyu mouse	*Mus caroli*	NT-CT	Dr A. Orth, University of Montpellier, France
AY490634	servant mouse	*Mus famulus*	NT-CT	Dr A. Orth, University of Montpellier, France
AY490635	western European house mouse	*Mus musculus domesticus*	NT-CT	Dipl. Biol. M. Buske, Federal Environmental Agency, Berlin, Germany
AY490636	house mice	*Mus musculus 129/SV*	NT-CT	Research Facility for Experimental Medicine, Berlin, Germany
AY490637	house mice	*Mus musculus Bal 6/3*	NT-CT	Charles River Laboratories, Wilmington, USA
AY490638	house mice	*Mus musculus OBA / 2*	NT-CT	Charles River Laboratories, Wilmington, USA
AY490639	house mice	*Mus musculus CBA / J*	NT-CT	Charles River Laboratories, Wilmington, USA
AY490640	house mice	*Mus musculus FvB*	NT-CT	Charles River Laboratories, Wilmington, USA
AY490641	midday gerbil	*Meriones meridianus*	IL1-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490642	Mongolian gerbil	*Meriones unguiculatus*	ECL2-CT	Research Facility for Experimental Medicine, Berlin, Germany

AY490643 AY490644 AY494001 AY494002	Norway rat	*Rattus norvegicus*	NT-CT	Research Facility for Experimental Medicine, Berlin, Germany Dipl. Biol. M. Buske, Federal Environmental Agency, Berlin, Germany Norbert Kelling, Woltersdorf, Germany Dr G. Jarrell, Alaska Frozen Tissue Collection, University of Alaska, Fairbanks, USA
AY490645 AY494003 AY494004	black rat	*Rattus rattus*	NT-CT	Dipl. Biol. M. Buske, Federal Environmental Agency, Berlin, Germany
	Dassie rat	*Petromus typicus*		Dr R. Adkins, University of TN, USA
	naked mole-rat	*Heterocephalus glaber*		Dr R. Adkins, University of TN, USA
	brushy-tailed rats	*Octodontomys gliroides*		Dr R. Adkins, University of TN, USA
AY490646 AY490647	guinea pig	*Cavia porcellus*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany Dipl. Biol. M. Buske, Federal Environmental Agency, Berlin, Germany
AY490648 AY490649	Chinese hamster	*Cricetulus griseus*	NT-CT	CHO-K1, ATCC – cell collection
AY528865	Syrian hamster	*Mesocricetus auratus*	NT-TMD6	Dipl. Biol. M. Buske, Federal Environmental Agency, Berlin, Germany
AY490650	muskrat	*Ondatra zibethicus*	NT-TMD7	Dr Lauenstein, Federal Centre for Agriculture and Forestry, Braunschweig, Germany
AY490651	spiny mouse	*Acomys sp.*	NT-TMD4	own source
AY490652	Barbary striped grass mouse	*Lemniscomys barbarus*	NT-CT	own source
AY502103 AY502104	Buffy Cuis	*Galea monastieri*	NT-CT	C. Kern, Zoo of Leipzig, Leipzig, Germany
Lagomorpha				
AY490653 AY490654	European hare	*Lepus europaeus*	NT-CT	Dipl. Med. J. Moeller, Erfurt, Germany
AY490655	rabbit	*Oryctolagus cuniculus*	NT-CT	own source
Cetacea				
AY490656 AY490657	minke whale	*Balaenoptera acutorostrata*	NT-CT	Dr P. J. Palsboll, University of Wales, U.K.
Artiodactyla				
AY490658 AY490659 AY490660 AY490661	domestic cow	*Bos taurus*	NT-CT	own source Dr J. Chae, University of Chonbuk, Korea
AY490662	red deer	*Cervus elaphus*	NT-CT	Dipl. Med. J. Moeller, Erfurt, Germany
AY490663	reindeer	*Rangifer tarandus*	NT-CT	K. Zierau, Berlin, Germany
AY490664	roe deer	*Capreolus capreolus*	NT-CT	Dipl. Med. J. Moeller, Erfurt, Germany

AY490665				
AY490666	moose	*Alces alces*	NT-CT	K. Zierau, Berlin, Germany
AY490667 AY490668	water buffalo	*Bubalus bubalis*	NT-CT	own source
AY490669	European bison	*Bison bonasus*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490670	domestic pig	*Sus scrofa domestica*	NT-CT	own source
AY490671	wild pig	*Sus scrofa*	NT-CT	Dipl. Med. J. Moeller, Erfurt, Germany
AY490672 AY490673	alpaca	*Lama pacos*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490674 AY490675	lama	*Lama glama*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490676 AY490677	Arabian camel	*Camelus dromedarius*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490678 AY490679	Bactrian camel	*Camelus bactrianus*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490680	pygmy hippopotamus	*Hexaprotodon liberiensis*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490681 AY490682	Mhorr gazelle	*Gazella dama mhorr*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490683 AY490684	Sabel antelope	*Hippotragus niger*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490685	giraffe	*Giraffa camelopardalis*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490686	muskox	*Ovibos moschatus moschatus*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490687 AY490688	sheep	*Ovis aries*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490689	mouflon	*Ovis aries musimon*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
Perissodactyla				
AY490690	horse	*Equus caballus*	NT-CT	own source
AY490691	Przewalski horse	*Equus przewalskii*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490692	kulan	*Equus hemionus kulan*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490693	tapirs	*Tapirus sp.*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490694	white rhinoceros	*Ceratotherium simum*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
Carnivora				
AY490695	cat	*Felis catus*	NT-CT	Dipl. Med. J. Moeller, Erfurt, Germany
AY490696	dog	*Canis familiaris*	NT-CT	own source
AY490697 AY490698	beach marten	*Martes foina*	NT-CT	own source
AY490699 AY490700	Eurasian river otter	*Lutra lutra*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany

AY490701	cheetah	*Acinonyx jubatus*	ECL1-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490702 AY490703	Baikal seal	*Phoca sibirica*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490704 AY490705	California sealion	*Zalophus californianus*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490706	brown bear	*Ursus arctos*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490707 AY490708	Malayan sun bear	*Helarctos malayanus*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490709	polar bear	*Ursus maritimus*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490710	American black bear	*Ursus americanus*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490711	ring-tailed coati	*Nasua nasua*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490712 AY490713	gray wolf	*Canis lupus*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490714	raccoon dog	*Nyctereutes procyonoides*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490715	red fox	*Vulpes vulpes*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490716	striped hyena	*Hyaena hyaena*	NT-IL3	Dr C. Pitra, IZW, Berlin, Germany
AY490717	jaguar	*Panthera onca*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490718	lion	*Panthera leo*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490719	puma	*Puma concolor*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490720 AY490721	snow leopard	*Uncia uncia*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490722 AY490723	Siberian tiger	*Panthera tigris altaica*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
Sirenia				
AY490724	Caribbean manatee	*Trichechus manatus*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
Proboscidea				
AY490725	Asiatic elephant	*Elephas maximus*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
AY490726	African elephant	*Loxodonta africana*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
Insectivora				
AY490733	European shrew	*Sorex araneus*	NT-CT	own source
AY490735	European mole	*Talpa europaea*	NT-TMD7	own source
AY490734	western European hedgehog	*Erinaceus europaeus*	IL1-CT	own source
Chiroptera				
AY490738	flying foxes	*Pteropus sp.*	NT-IL3	Dr C. Pitra, IZW, Berlin, Germany
AY490736	Central American yellow bat	*Rhogeessa io*	NT-IL3	Dr C. Pitra, IZW, Berlin, Germany
AY490737	Silver-haired Myotis	*Myotis albescens*	NT-IL3	Dr C. Pitra, IZW, Berlin, Germany

Scandentia				
AY490739 AY490740	northern tree shrew	*Tupaia belangeri*	NT-CT	Dr W. Enard, MPI, Leipzig, Germany
Dermoptera				
	Philippine flying lemur	*Cynocephalus volans*		Dr R. Adkins, University of TN, USA
Tubulidentata				
	aardvark	*Orycteropus afer*		Dr R. Adkins, University of TN, USA
Hyracoidea (Procavia)				
AY490741	cape rock hyrax	*Procavia capensis*	NT-IL3	Dr C. Pitra, IZW, Berlin, Germany
Xenarthra				
AY490742 AY490743	southern two-toed sloth	*Choloepus didactylus*	NT-CT	Dr C. Pitra, IZW, Berlin, Germany
Didelphimorphia				
AY490728	gray short-tailed opossum	*Monodelphis domestica*	ECL2-IL3	Dr A. Janke, Lund University, Sweden
Diprotodontia				
AY490732	silver-gray brushtail possum	*Trichosurus vulpecula*	IL1-IL3	Dr J. Demmer, AgResearch Ruakura, New Zealand
AY490727	common wombat	*Vombatus ursinus*	TMD4-IL3	Dr A. Janke, Lund University, Sweden
AY490729	long-nosed potoroo	*Potorous tridactylus*	ECL1-IL3	Dr A. Janke, Lund University, Sweden PtK2, ATCC – cell collection
AY490730 AY490731	wallaroo	*Macropus robustus*	IL1-IL3	Dr A. Janke, Lund University, Sweden
Monotremata				
	platypus	*Ornithorhynchus anatinus*	none	Dr W. Mayer, MPI Biology, Tübingen, Germany
	Australian echidna	*Tachyglossus aculeatus*	none	Dr A. Janke, Lund University, Sweden

Non-mammalian species analyzed in this study for presence of GPR33 orthologs are listed in the Supplementary Section of ref. (22).

Suppl. Table S2
Primers used for GPR33 ortholog amplification

primer number	sequence	orien-tation	domain
1	5'-TGYTTGCCAKTWTWRAAAYATTRC-3'	s	5'-UTR rodent
2	5'-GGTTGTGAGCCACCATGCAGTTGC-3'	s	5'-UTR rat
3	5'-CATTGCCATTTTGAATATTAATATG-3'	s	5'-UTR cattle
4	5'-TTGCCCTTCAGATTACTACTATAAG-3'	s	5'-UTR Apodemus
5	5'-AGATTSRTCTGATCTTAATTCCTCA-3'	s	5'-UTR non primate
6	5'-GCTTCTTTTTTKTTKTYATTYYCAGTTTCTCAYTCCAATT-3'	s	5'-UTR even-toed ungulate
7	5'-ATTTTTTTTTTTTTCCTYTNTTYTYCCTCCATAGGTAAYT-3'	s	5'-UTR rodents
8	5'-CTCCATAGGTMAYTATGGATYKGRTCAACT-3'	s	N-term mammalian
9	5'-CTCCATAGSYMAYTATGGAYYKGRYCARCTC-3'	s	N-term mammalian
10	5'-ATGGAYCTGACCAACTCTACTGRYKWYCTGCTCAAT-3'	s	N-term even-toed ungulate
11	5'-ATGGAYCTGACCAACTCTACTGRYKWYCTGCT-3'	s	N-term even-toed ungulate
12	5'-ATGGATTTGGTCAACTTTASTGATTTGC-3'	s	N-term carnivore
13	5'-ATGGATCTGATCAACTCTACTGATTAC-3'	s	N-term primate
14	5'-CCCACCATGGATTTGATCAACTCCTCCAC-3'	s	N-term mouse
15	5'-CCCACCATGGATCTGATCAACTCTACTG-3'	s	N-term human
16	5'-TCAGTAAGAAATAGCACTCATTTTCTAGCTCMYGC-3'	s	N-term odd-toed ungulate
17	5'-GATGCAGGAGCTAGAAACTGAGTGC-3'	as	N-term primate
18	5'-CAAACAGCACTGGTGTTCCAACTCC-3'	s	N-term murinae
19	5'-AGCYCCTRCATCAAAAATRATKATTGC-3'	s	TMD1 primate
20	5'-ATCASCAAYGGCCTCTAYCTRTGG-3'	s	TMD1 mammalian
21	5'-ATCASCAAYGGCCTNTAYCTNTGG-3'	s	TMD1 mammalian
22	5'-CAATGGTCTCTATCTATGG	s	TMD1 even-toed ungulate
23	5'-TATCTATGGGTGCTNNWWTTYAARATG-3'	s	TMD1&IL1
24	5'-TTTTCAYCTYATTYYTCHTAYTTTAT-3'	s	TMD2
25	5'-AATACNCTYTTRTWTTTYCATCTCATT-3'	s	TMD2
26	5'-AAATAAAATAGGAGAGAATGAGATG-3'	as	TMD2 mammalian
27	5'-TCMACATTGATTCTGCCMTTYATG-3'	s	TMD2 mammalian
28	5'-TAACTATGCTGTGTCYASTRACTGG-3'	s	EL2
29	5'-CCTTGCTTTCCCCAGTTAGTMSMCAC-3'	as	EL2
30	5'-CTGTTTTCTTCCTYTCNGCCATMRG-3'	s	TMD3
31	5'-CACTGGAGCTTTGGAACTGCCATGTG-3'	s	TMD3
32	5'-CCTTGCACATGGCAGTTCCAAAGCTCC-3'	as	TMD3
33	5'-GAAGGTAACGATCAAKACYKATRGC-3'	as	IL2
34	5'-CCYTAYCTGGTATTCCCGGGAGACC-3'	s	TMD4
35	5'-ATGACCATAAAGGAAGAGTGACCTG-3'	s	EL2
36	5'-GTCACCATAACCTTAAAGGGTTTGC-3'	as	IL3
37	5'-GAGTAGCCAACAAGATGAAAGAGAG-3'	s	IL3
38	5'-GATGACTTTGAAGGGYTTGYTGGA-3'	as	IL3
39	5'-GAAAGAGTAGCCAGAAAGATGAAAG-3'	s	IL3
40	5'-TCYTTCTTYSTGTGYTGGATGCC-3'	s	TMD6

41	5’-TCTCTTTCTTTGTGTGYTGGMTGCC-3’	s	TMD6 mammalian
42	5’-TARGYCATCCARCACASRAAGAARGA-3’	as	TMD6
43	5’-GGTRTAYRTGRTANGGCATCCARCA-3’	s	TMD6 mammalian
44	5’-GGTAGGGCATCCAACACASRAAGAA-3’	as	TMD6 mammalian
45	5’-GTGTAYRTGRTAGGGCATCCARCA-3’	as	TMD6 mammalian
46	5’-TCTTAGTGAGGACCAAGCCACTGTG-3’	as	TMD6 murinae
47	5’-GGGCTTAATGCTCACTAAGAACCGG-3’	s	TMD6 lemur
48	5’-GACTCTGATACTTACAGTGCTAAC-3’	s	TMD7
49	5’-GAARTTCTCCCCARYAAAYAARTARAG-3’	as	TMD7 mammalian
50	5’-TCCCCARYAAATAARTARAGWRYNGGAGA-3’	as	TMD7 mammalian
51	5’-CCATTCTTGCTCTGTTTGAGTC-3’	s	C-term primate
52	5’-GAGCAAGAATRGAYTTYTTRAARAC-3’	as	C-term mammalian
53	5’-TTCTGRRTTYAGTTTTGNGTYCTTTC-3’	as	C-term mammalian
54	5’-GAATTTAAAYTTMGRYTTCTGARTTNAG-3’	as	C-term- mammalian
55	5’-TTACATGTCGGCTTCTGAATTTAGG-3’	as	C-term cattle
56	5’-TTAAATGTCTGAGTTCAGGTTTG-3’	as	C-term Apodemus
57	5’-TCYAGAATTTAAATTTMRGYTTCTG-3’	as	C-term mammalian
58	5’-TCTGAGTTTAGGTTTGYGYCYTKTC-3’	as	C-term mammalian
59	5’-GGACTTCTTGAAGACSTTTTTGMARTTC-3’	as	C-term primate
60	5’-TCCCACCCAAAGGATCCAGAAT-3’	as	C-term mouse
61	5’-CCAAGTGCGTGTTTGCTTAAG-3’	as	C-term human
62	5’-GTGCRTGTTTGCTTAAGMATCTRGA-3’	as	3’-UTR primate
63	5’-TTGASCAAGTGCRTGTTTGCTTAAG-3’	as	3’-UTR primate
64	5’-CTAGTTATCTCTCTAATATAGGGAC-3’	as	3’-UTR
65	5’-GAATCACATTTGCAGCATGATCTTG-3’	as	3’-UTR
66	5’-GGTTATAGACAKATATGWGTKTTGG-3’	as	3’-UTR murinae
67	5’-TKCCTAGGRACTGGGGGTTATAGAC-3’	as	3’-UTR murinae

Suppl. Table S3
Human DNA samples

The DNA sample collection contained most samples of the CEPH panel, a panel representing all major linguistic groups (DNA panel from ref. 8), 45 individuals from Papua New Guinea (kindly provided by Mark Stoneking) and 21 Yoruba individuals (International HapMap Project) was used. # indicates a homozygote individual.

Number	sample group	Number of individuals	TGA allele	CGA allele
1	San	8	16	0
2	Bantu (south east)	5	10	0
3	Bantu (south west)	3	6	0
4	Biaka Pygmies	37	69	5
5	Mbuti Pygmies	19	38	0
6	Mbenzele Pygmies	2	4	0
7	Bamileke	1	2	0
8	Igbo	1	2	0
9	Yoruba	47	94	0
10	Effik	1	2	0
11	Hausa	1	2	0
12	Mandenka	24	48	0
13	Bantu (Kenya)	12	24	0
14	Lisombo	1	1	1
15	Ethiopian jew	2	4	0
16	Mozabite	30	55	5
17	Nubian	2	4	0
18	Kopte	1	2	0
19	Bedouin	49	98	0
20	Druze	50	98	2
21	Palestinian	53	105	1
22	Habani Syrian	2	4	0
23	Iranian	3	6	0
24	Yemenite jew	2	4	0
25	Brahui, Balochi, Hazara, Sindhi	100	192	8
26	Burusho, Makrani, Pathan, Kalash	100	200	0
27	Uzbek	2	4	0
28	Kirghiz	1	2	0
29	Tamile	1	2	0
30	Asian Indian	1	2	0
31	Thai	2	4	0
32	Cambodian	11	22	0
33	Filipino	2	2	2#
34	Papuan	66	125	7
35	NAN Melanesian	22	43	1
36	Nasioi	2	4	0
37	Samoan	2	3	1
38	Aborigine Australian	4	8	0
39	Japanese	35	69	1
40	Han	45	87	3
41	Tujia	10	19	1
42	Yizu	10	20	0
43	Miaozu	10	20	0
44	Oroqen	10	20	0
45	Daur	10	20	0
46	Mongola	10	20	0
47	Hezhen	10	19	1

48	Xibo	9	17	1
49	Uygur	10	20	0
50	Dai	10	20	0
51	Lahu	10	16	4
52	She	10	20	0
53	Naxi	10	19	1
54	Tu	10	19	1
55	Buryat	1	1	1
56	Korean	1	2	0
57	Yakut	26	51	1
58	Evenki	1	2	0
59	Chukchi	1	2	0
60	Siberian Inuit	1	2	0
61	Estonia	1	2	0
62	Saami	2	4	0
63	Swedish	1	2	0
64	Danish	1	2	0
65	German	10	20	0
66	Dutch	1	2	0
67	English	1	2	0
68	French	30	60	0
69	French Basque	24	48	0
70	Spain	1	2	0
71	Sardinian	28	56	0
72	North Italian	15	30	0
73	Tuscan	8	16	0
74	Orcadian	16	31	1
75	Islandic	1	2	0
76	Greek	1	2	0
77	Adygei	17	33	1
78	Russian	25	49	1
79	Pima	25	50	0
80	Maya	25	49	1
81	Warao Indian	2	4	0
82	Colombian	13	26	0
83	Karitiana	24	48	0
84	Surui	21	42	0
85	Kaingang	1	2	0
86	Guarani	1	2	0

Suppl. Table S4
Primers used for dating experiments (human, bonobo, chimpanzee)

primer number	sequence	orien-tation	domain
hu- 1	5'-TGAGGGGCTTAGCAGGAGA-3'	s	5'-UTR
hu- 2	5'-GGATAAAAACACCACAGAGAC-3'	s	5'-UTR
hu- 3	5'-TAGAAGGCAGAATTAACTTAGG-3'	s	5'-UTR
hu- 4	5'-ACAGAGTCATATCAGCAAATTG-3'	as	5'-UTR
hu- 5	5'-CACCATGTTGGCCAGGATG-3'	s	5'-UTR
hu- 6	5'-CATGCCTGTAATTCCAGCAC-3'	as	5'-UTR
hu- 7	5'-ACTTTAATTCCAGCACCTGGGAGGC-3'	as	5'-UTR
hu- 8	5'-CTGCGCCTGGCCAAATGTAG-3'	as	5'-UTR
hu- 9	5'-CTGAAGTCAGGAGTTTGAGACC-3'	s	5'-UTR
hu-10	5'-CCTACCTCTACCTCTGCACTTGGAA-3'	as	5'-UTR
hu-11	5'-GTGACAAAGCAAGACTCTGAC-3'	s	5'-UTR
hu-12	5'-AAGTACCTTCAGAATCATCAAC-3'	as	5'-UTR
hu-13	5'-TATTTGACTTGTGAGGGCAAAAGTAGC-3'	as	5'-UTR
hu-14	5'-ACTATGTTGCCCAGCCTGGTC-3'	as	5'-UTR
hu-15	5'-CATATATTTCTATTTTTGATGCTCAG-3'	s	5'-UTR
hu-16	5'-CCATAAATGGCAGAATCATTGTTG-3'	as	cod reg
hu-17	5'-GGAACTGCCTTGTGCAAGGTC-3'	s	cod reg
hu-18	5'-TGATCGTTACCTTCTCACTCT-3'	s	cod reg
hu-19	5'-CCCAGTTAGTAGACACAGCATAG-3'	as	cod reg
hu-20	5'-TCAAGAAGTCCATTCTTGCTC-3'	s	cod reg
hu-21	5'-ACCTGAGTAGCTGAGACTAC-3'	s	3'-UTR
hu-22	5'-GTGATCCGCCTGCCTCAACC-3'	s	3'-UTR
hu-23	5'-ACAAACTCTGTGCCCATTAAG-3'	s	3'-UTR
hu-24	5'-GAGTTACTACTTAATGGGCAC-3'	as	3'-UTR
hu-25	5'-CATGTTACATTCCCACCAGC-3'	s	3'-UTR
hu-26	5'-GTGGGAATGTAACATGGTGC-3'	as	3'-UTR
hu-27	5'-TACTCGAGAGGCTGAGGCAG-3'	as	3'-UTR
hu-28	5'-GTAGAACCACATAATATCACTCACC-3'	s	3'-UTR
hu-29	5'-ATTTAGAATTGACAGGACTACC-3'	as	3'-UTR
hu-30	5'-TCAGTGGGCTTGTCAGGACC-3'	as	3'-UTR
hu-31	5'-CAAAAGGCACTCTCTGTAGCTGC-3'	as	3'-UTR
hu-32	5'-CAATGTGCCTGTCAGCACTC-3'	s	3'-UTR
hu-33	5'-CACCCCACACTATCCCAGGTC-3'	s	3'-UTR
hu-34	5'-GGGCTTGGTCTCTCTTACTTGGAAT-3'	s	3'-UTR
hu-35	5'-TGGTGTCCTAGAAATCCTGCTAA-3'	s	3'-UTR
hu-36	5'-TATAGCCTACGCAGAAATCCCCAAA-3'	s	3'-UTR
hu-37	5'-ATGTTCAAGAATCTGCCTGTTGT-3'	s	3'-UTR
hu-38	5'-GAGCTCCCTCATTATTTAGTAACATC-3'	s	3'-UTR
hu-39	5'-ATAAAGCTTTTCAATTTGCACCA-3'	s	3'-UTR
hu-40	5'-TCATCATTCCAAGGACAAACTCT-3'	s	3'-UTR

hu-41	5’-CCAAAATTTTTCATCATTCCAAGGAC-3’	s	3’-UTR
hu-42	5’-GATTCTTAAGCAAACACGCACTT-3’	s	3’-UTR
hu-43	5’-GGAAAGCAAGGAGATGCAAG-3’	s	3’-UTR
hu-44	5’-AAGACAATCACTGGAACTTTGGA-3’	s	3’-UTR
hu-45	5’-GCACTCAGTTTCTAGCTCCTGCATC-3’	s	3’-UTR
hu-46	5’-TTCACTGTTTCTCACTCCACAGGTC-3’	s	3’-UTR
hu-47	5’-GGTGATTCATTTGCATATTGAGG-3’	as	3’-UTR
hu-48	5’-AACCACTGCACTAGACTTCAAGC-3’	as	3’-UTR
hu-49	5’-TCCCCTAATTTCTCAGGGATATGGA-3’	as	3’-UTR
hu-50	5’-TGAGGAAAGAACTTACACTGGCATT-3’	as	3’-UTR
hu-51	5’-AAATACTTCAAAAAGCCCTGACC-3’	as	3’-UTR
hu-53	5’-AGCTTCAAAACCTTGAAATAGGG-3’	as	3’-UTR
hu-54	5’-GCTGGTGTATTCAGGGACTGAGTGT-3’	as	3’-UTR
hu-55	5’-AACAAGATGAGTGCCTTGTGG-3’	as	3’-UTR
hu-56	5’-GGTGGTGATGGTTGTACAATTTCGT-3’	as	3’-UTR
hu-57	5’-TGGGCAAAATCCCATCTCTA-3’	as	3’-UTR

Suppl. Table S5
Primers used for dating experiments (rats)

primer number	sequence	orien-tation	domain
rat- 1	5’-TCATCAGGCACACATGTGGTG-3’	s	5’-UTR
rat- 2	5’-CAGGGTAATCCAAATCAAAGAGC-3’	s	5’-UTR
rat- 3	5’-GTGAAATGCTGTCTAACAGGCTC-3’	as	5’-UTR
rat- 4	5’-AACCGAGGTACAGATCTTCTTG-3’	as	5’-UTR
rat- 5	5’-TCCAACCACTGCCCAGATGC-3’	s	5’-UTR
rat- 6	5’-GACAACAATCTTTGGCCTGTCAC-3’	s	5’-UTR
rat- 7	5’-ATTTCATTGGCTTATGGGCCATC-3’	as	5’-UTR
rat- 8	5’-CTAGTCTGCGGTTCCTGAGG-3’	s	5’-UTR
rat- 9	5’-CTTACAGCTACAGACTCTATTTC-3’	as	5’-UTR
rat-10	5’-ATCTAAGGAGCTAAGTCCACTG-3’	s	5’-UTR
rat-11	5’-GCTGCATGACAAGTTTGAG-3’	s	5’-UTR
rat-12	5’-CTAAAATCTTTTGATGTGTAATTCG-3’	s	5’-UTR
rat-13	5’-GTAAGGACTGACCTTCAAGAG-3’	as	5’-UTR
rat-14	5’-TGAACAACAATAAAGATTAAATG-3’	as	5’-UTR
rat-15	5’-GGAGTTGGAACACCAGTGCTG-3’	as	cod reg
rat-16	5’-CCATGAAGGGCAGAATCAATGTGG-3’	as	cod reg
rat-17	5’-CCTCCTTCCTTCAGGACAATCAC-3’	s	cod reg
rat-18	5’-CGAAACAGGCTGCATGAATCC-3’	as	cod reg
rat-19	5’-CAGCAAGCCCTTCAAAGTCATG-3’	s	cod reg
rat-20	5’-AATACAAAATCAGAGGATTTAACTC-3’	as	3’-UTR
rat-21	5’-AGCATAAACTCGGCATTTGCTC-3’	s	3’-UTR
rat-22	5’-CTGCAAAGAGAAACTCTGTATTC-3’	s	3’-UTR
rat-23	5’-GCCTGAAATACTTTATTATGTAGATG-3’	as	3’-UTR
rat-24	5’-CACATCTTCATCACCACTTGGTG-3’	s	3’-UTR
rat-25	5’-CCAAGGCAACTCTTATAAAGGAC-3’	as	3’-UTR
rat-26	5’-TATTCACAGAATTTAACCTAAAAAGC-3’	s	3’-UTR
rat-27	5’-GCACTCCTGAGGCTGCAGAG-3’	s	3’-UTR
rat-28	5’-CACAGAGTGCAGTCTCTTCTG-3’	as	3’-UTR
rat-29	5’-ACTCAAGACACAGGCCCACAG-3’	s	3’-UTR
rat-30	5’-GTTATGGGGACAGAGTGTTCTGC-3’	as	3’-UTR
rat-31	5’-GTGTTTGGATATTCCGTGTTTGC-3’	as	3’-UTR

Suppl. Table S6

Determination of mutation rates in the GPR33 locus of *Hominoidea* and rats

The number of nucleotide differences in the genomic non-coding and coding regions between the indicated GPR33 alleles of human, chimpanzee and bonobo (upper panel; 5,447 bp total) and *R. norvegicus* and *R. rattus* (lower panel; 5,293 bp total) was counted. Data are given as nucleotide differences per 1,000 bp.

5,447 bp total	Human **Stop140**	Human Arg140	Bonobo	Chimpanzee Ser39/Arg140	Chimpanzee **Stop39**/Arg140
Human Arg140	2.2	-			
Bonobo	9.4	9.4	-		
Chimpanzee Ser39/Arg140	9.6	9.4	2.8	-	
Chimpanzee **Stop39**/Arg140	9.4	9.2	2.6	1.3	-
Chimpanzee Ser39/**Stop140**	9.9	10.1	3.5	1.1	2.0

5,293 bp total	*R. norvegicus* A	*R. norvegicus* B
R. norvegicus B	3.4	-
R. rattus	30.2	29.5

Suppl. Table S7
Putative Yoruba haplotypes

The GPR33 coding region (~1 kb) and 3.8 kb of the-3' untranslated region from 21 Yoruba individuals were directly sequenced. The most likeliest haplotypes for each individual were calculated with PHASE.

Nucleotide position in Chr14 (29938165-29943210)		29942793	29942783	29942296	29942001	29941596	29941588	29941369	29941186	29941016	29940934	29940530	29940491	29940382	29940167	29940164	29940148	29940142	29940122	29940118	29940104	29940099	29940090	29940089	29940079	29939560	29939485	29939424	29939310	29939068	29939013	29938741
position starting with Start-ATG in GPR33		418	428	915	1215	1615	1623	1842	2025	2195	2277	2681	2720	2829	3044	3047	3063	3069	3089	3093	3107	3112	3121	3122	3132	3651	3726	3787	3901	4143	4198	4470
	Phase haplotype	Stop																														
ANCESTRAL STATE		C	G	C	C	C	C	C	A	A	T	A	T	A	A	T	T	T	A	T	A	A	C	A	G	G	G	T	G	C	T	A
Chimpanzee		C	G	C	C	C	C	C	A	A	T	A	T	A	A	T	T	T	A	T	A	A	C	A	G	G	G	T	G	C	T	A
Human CGA Filipino		C	G	C	C	C	C	C	A	A	T	A	T	A	A	T	T	T	A	T	A	A	C	A	G	G	C	T	G	C	T	W
Yoruba-21	1	T	-	T	-	-	-	-	-	-	-	-	-	-	-	-	C	C	-	-	-	-	-	-	-	A	-	-	-	T	-	-
	2	T	-	T	-	T	-	-	-	-	C	-	-	-	C	-	-	-	-	-	G	G	-	-	-	A	-	-	-	T	-	-
Yoruba-20	3	T	-	T	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	G	-	-	C	-	-	-	-	T
	4	T	-	T	-	-	-	-	-	-	-	-	-	G	C	-	-	-	-	-	-	-	-	-	-	A	-	C	-	T	-	-
Yoruba-19	5	T	-	T	-	T	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	C	-
	6	T	-	T	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	A	-	C	-	T	-	-
Yoruba-18	7	T	-	T	-	-	-	-	-	-	-	-	-	-	C	-	-	-	G	C	-	-	-	-	-	A	-	-	-	-	-	T
	8	T	-	T	T	-	-	-	-	-	C	G	-	-	C	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-	-
Yoruba-17	9	T	-	T	-	T	-	-	-	-	-	G	C	-	-	C	C	C	-	-	-	-	-	-	-	-	-	-	-	-	-	T
	10	T	-	T	-	T	-	-	-	-	-	G	C	-	-	-	-	C	-	-	-	-	-	-	-	-	C	-	-	-	-	T
Yoruba-16	11	T	-	T	-	-	-	-	-	-	C	G	-	-	C	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-	-
	12	T	-	T	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	G	-	A	-	C	-	T	-	-

Yoruba-15	13	T	-	T	-	T	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-	-
	1	T	-	T	-	-	-	-	-	-	-	-	-	-	-	-	C	C	-	-	-	-	-	-	-	A	-	-	-	T	-	-
Yoruba-14	11	T	-	T	-	-	-	-	-	-	C	G	-	-	C	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-	-
	14	T	-	T	T	-	-	-	-	-	-	G	-	-	C	-	-	-	-	C	-	-	-	-	-	A	-	-	-	T	C	-
Yoruba-13	15	T	-	T	-	T	-	-	G	-	-	-	-	-	C	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T
	16	T	-	T	-	-	-	-	-	-	-	-	-	G	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T
Yoruba-12	17	T	-	T	-	T	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-	-
	18	T	A	T	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	G	-	-	C	-	-	-	-	T
Yoruba-11	19	T	A	T	-	-	-	T	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-	-
	20	T	-	T	-	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	G	G	-	-	-	A	-	-	-	T	-	-
Yoruba-10	21	T	-	T	-	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-	T
	1	T	-	T	-	-	-	-	-	-	-	-	-	-	-	-	C	C	-	-	-	-	-	-	-	A	-	-	-	T	-	-
Yoruba-9	22	T	-	T	-	-	-	-	-	-	-	-	-	G	C	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-	-
	11	T	-	T	-	-	-	-	-	-	C	G	-	-	C	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-	-
Yoruba-8	1	T	-	T	-	-	-	-	-	-	-	-	-	-	-	-	C	C	-	-	-	-	-	-	-	A	-	-	-	T	-	-
	23	T	-	T	-	T	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	T	G	-	-	C	-	-	-	-	T
Yoruba-7	24	T	A	T	-	T	-	T	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	C	-
	25	T	-	T	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	A	-	C	-	-	-	-	T
Yoruba-6	26	T	-	T	-	-	-	-	-	-	C	G	-	-	-	-	-	-	-	-	-	-	-	-	-	A	C	-	-	T	-	-
	27	T	-	T	-	T	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	T
Yoruba-5	22	T	-	T	-	-	-	-	-	-	-	-	-	G	C	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-	-
	1	T	-	T	-	-	-	-	-	-	-	-	-	-	-	-	C	C	-	-	-	-	-	-	-	A	-	-	-	T	-	-
Yoruba-4	28	T	-	T	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	C	-	A	-	-	T
	29	T	-	T	-	-	-	-	-	-	-	-	-	G	C	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	T
Yoruba-3	30	T	A	T	T	-	-	-	G	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-	-
	27	T	-	T	-	T	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	T
Yoruba-2	1	T	-	T	-	-	-	-	-	-	-	-	-	-	-	-	C	C	-	-	-	-	-	-	-	A	-	-	-	T	-	-
	7	T	-	T	-	-	-	-	-	-	-	-	-	-	C	-	-	-	G	C	-	-	-	-	-	A	-	-	-	-	-	T
Yoruba-1	31	T	-	T	-	T	-	-	-	-	-	G	C	-	-	C	C	C	-	-	-	-	-	-	-	-	C	-	-	-	-	T
	32	T	-	T	-	-	-	-	-	-	C	-	-	-	C	-	C	T	A	-	-	-	-	-	-	A	-	-	-	T	-	-