GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Uneven evolutionary rate of the melatonin-related receptor gene (*GPR50*) in primates

Y. Zhang[1,2]*, H.Q. Li[3]*, Y.F. Yao[1]*, W. Liu[1,2], Q.Y. Ni[1], M.W. Zhang[1] and H.L. Xu[1,4]

[1]College of Animal Science and Technology, Sichuan Agricultural University, Ya'an, China
[2]Forestry College, Sichuan Agricultural University, Ya'an, China
[3]College of Life Science and Technology, Yangtze Normal University, Chongqing, China
[4]Key Laboratory of Animal Disease and Human Health of Sichuan Province, Sichuan Agricultural University, Ya'an, China

*These authors contributed equally to this study.
Corresponding author: H.L. Xu
E-mail: huailxu@yahoo.com

Genet. Mol. Res. 14 (1): 680-690 (2015)
Received November 11, 2013
Accepted March 17, 2014
Published January 30, 2015
DOI http://dx.doi.org/10.4238/2015.January.30.11

ABSTRACT. The melatonin-related receptor GPR50 plays an important role in mammalian adaptive thermogenesis in response to calorie intake. The evolutionary history of the *GPR50* gene is poorly understood in primates; however, it has been reported that GPR50 is the mammalian ortholog of Mel1c, which has been well characterized. In this study, the complete coding sequences of the *GPR50* gene in the Sichuan snub-nosed monkeys (*Rhinopithecus roxellana*) and Tibetan macaques (*Macaca thibetana*) were sequenced, and the orthologous nucleotide acid sequences of the *GPR50* gene in 11 other primate species were downloaded from GenBank. Thirteen species representing 6 major primate lineages (human, great ape, lesser ape, Old World monkey, New World monkey, and prosimian monkey) were subjected to statistical analyses. A selective test showed that the entire *GPR50* gene sequence is under strong purifying selection in these primates but has a significantly different evolutionary rate among the 6 major primate

lineages. Notably, both the *Homo* and *Pan* branches exhibited an ω ratio >1, indicating accelerated evolution of the two lineages. Further analysis of different domains revealed that the acceleration trend was more significant in the C-terminal domain (CTD). Interestingly, in the alignment of 13 primate *GPR50* nucleotide acid sequences, numerous insertions or deletions were only found in the CTD region, implying that this region may play a key role in the process of primate *GPR50* evolution. The results provide deeper insight into the functional evolution of *GPR50* in mammals at the molecular level.

Key words: *GPR50* gene; Melatonin-related receptor; Primate; Molecular evolution; Sichuan snub-nosed monkey; Tibetan macaque

INTRODUCTION

The orphan receptor GPR50 belongs to the G protein-coupled receptor (GPCR) superfamily and has the typical 7-transmembrane structure (Marinissen and Gutkind, 2001; Lundstrom, 2006). The original *GPR50* gene was cloned from the human pituitary complementary DNA (cDNA) library (Reppert et al., 1996). The gene is located on chromosome Xq28 in humans (Gubitz and Reppert, 1999) and consists of 2 exons separated by a single intron. Generally, *GPR50* in humans is a splice variant of 617 amino acids. GPR50 has 45% identity with the melatonin receptor family in eutherian mammals; thus, it was identified as a melatonin-related receptor (Reppert et al., 1996). Three melatonin receptor subtypes have been characterized, including MT1, MT2, and Mel1c. Mel1c was recently identified in amphibians, fish, and birds; the others are expressed in mammalian and non-mammalian species (Dubocovich et al., 2010). However, despite this close structural relationship, GPR50 does not bind melatonin (Drew et al., 1998), and it has been shown to interact with the melatonin receptors in a ligand-independent manner when overexpressed (Levoye et al., 2006). *GPR50* is most highly expressed in the brain, especially in the dorsal medial hypothalamus, but it is also found in the adrenal glands, testes, ovaries, eyes, lungs, kidneys, intestines, and hearts of mice and rats (Drew et al., 2001; Izzo et al., 2010). Recently, *GPR50* was identified as a gene indicating risk of major mental illness in women through a case-control association (Thomson et al., 2005; Macintyre, 2010). A previous study (Bechtold et al., 2012) established the orphan receptor GPR50 as a critical player that significantly advances the field toward understanding the molecular signaling mechanisms involved in achieving a torpid state. Remarkably, *GPR50*-knockout mouse models have been developed. Comparatively, the *GPR50*$^{-/-}$ mice exhibited an increase in metabolic rate and a decrease in the amount of fat stores when fed a normal chow diet (Ivanova et al., 2008). Hence, the evidence further lends support to the potential role of the GPR50 receptor in metabolic regulation. GPR50 is a GPCR recently demonstrated to play an important role in adaptive thermogenesis in response to calorie intake (Levoye et al., 2006). Interestingly, a recent phylogenetic study of GPR50 and the melatonin receptors suggested that GPR50 was the mammalian ortholog of Mel1c (Dufourny et al., 2008); however, the receptor has experienced a number of changes throughout its evolution, including the mutation of numerous amino acids and the addition of a long C-terminal domain (CTD).

Nonhuman primates (NHPs) are a diverse group of animals. Due to the close genetic relationship between NHPs and humans, NHPs have been extensively used in biomedical re-

search as experimental models. Comparative genomics and evolutionary analyses have been powerful tools in the exploration of gene function. Although the roles of GPR50 have been well clarified in other mammals, including mice and rats, little is known on the functions of these proteins in NHPs. In particular, earlier studies have rarely referred to the evolution of the primate *GPR50* gene. This study is expected to provide insight into the functional significance of the gene throughout the history of primate evolution.

MATERIAL AND METHODS

Samples

DNA samples were isolated from the muscle tissue of Sichuan snub-nosed monkeys (*Rhinopithecus roxellana*) and Tibetan macaques (*Macaca thibetana*) at the Wildlife Conservation Laboratory, Sichuan Agricultural University, China. All the samples used in this study were collected according to the law of the PR China on the protection of Wildlife established by the Ministry of Forestry of China. We collected known and predicted *GPR50* gene sequences from humans (accession No. NM004224) as query sequences and conducted BLAST searches to identify *GPR50* gene sequences against currently available genome sequences in the NCBI (http://www.ncbi.nlm.nih.gov) database. Eleven primate species with the *GPR50* gene coding sequence were obtained, covering the major lineages of primates. The data set included the common marmoset (*Callithrix jacchus*, XM002763367), western gorilla (*Gorilla gorilla*, XM004065017), rhesus macaque (*Macaca mulatta*, XM001092026), northern white-cheeked gibbon (*Nomascus leucogenys*, XM003271842), northern greater galago (*Otolemur garnettii*, XM003800937), bonobo (*Pan paniscus*, XM003804526), chimpanzee (*Pan troglodytes*, XM001136005), olive baboon (*Papio anubis*, XM003918414), Sumatran orangutan (*Pongo abelii*, XM002832239), common squirrel monkey (*Saimiri boliviensis*, XM003943417), and modern human (*Homo sapiens*). The secondary structure of GPR50 in humans was predicted using the Universal Protein Resource (www.uniprot.org, query No. Q13585).

Polymerase chain reaction (PCR) and DNA sequencing

The PCR primers were designed to cover the coding region of *GPR50* by using the Primer 5.0 software based on the sequence alignment of the Rhesus macaque and synthesized by BGI-Shenzhen (Shenzhen, China). Nucleotide sequences of the primers were as follows: GPR1-F: 5'-ATTGGTCGGGGGGAGGGATTTG-3' (forward) and GPR1-R: 5'-TGAGGGCACTGT CCATAGATTC-3' (reverse); GPR2-1F: 5'-TGTCTTTCCCCTTTCTCCCTCC-3' (forward) and GPR2-1R: 5'-CATCAGGGTTTCCTCCACAGCA-3' (reverse); and GPR2-2F: 5'-CTGTC AGTCCGAAGGAGATGGC-3' (forward) and GPR2-2R: 5'-GCAAAACAAACTGGAAAGG CGA-3' (reverse). The primer pair GPR1-F and GPR1-R was used to amplify the first exon, and the others were used to amplify the second exon. PCR was performed in a thermal cycler (Mastercycler gradient, Eppendorf, Germany) with a total reaction volume of ~10 µL, which was comprised of 1 µL genomic DNA (10 ng/µL), 0.3 µL of each primer, 5 µL Premix LA Taq (loading dye mix; TaKaRa Biotechnology Co., Ltd., Dalian, China), 3.5 µL double-distilled water, and 1 µL mineral oil. PCR conditions were initial denaturation at 95°C for 5 min; 32 cycles of denaturation at 94°C for 35 s, annealing at 50°C for 50 s, and an extension at 72°C for 1 min; followed by a final extension at 72°C for 10 min. The amplified GPR50 fragment was purified using a PCR gel extraction kit (Sangon Biotech Co., Ltd., Shanghai, China).

Data analysis

Nucleotide sequence variation was confirmed through artificial check and correction using the Chromas 1.45 (McCarthy, 1998) and SeqMan programs (DNASTAR LaserGene, Madison, WI, USA). To identify sequence characterization, we completed a comprehensive analysis of the sequence, including sequence homology, GC content of various codons, ratio of transition/transversion (*ti/tv*), and information of variable sites, in the MegAlign program (DNAStar LaserGene) and MEGA 5.1 (Tamura et al., 2011). We utilized the Bayesian method to infer ancestral sequences and then calculated the evolutionary rates using the following equation:

$$r = K/T$$

where r denotes the evolutionary rate, K denotes the number of substitutions per site between 2 sequences, and T denotes the divergence times between the putative ancestors and the living primate species (Graur and Li, 2000). The relative rate test was conducted to detect variations in the substitution rate (Tajima, 1993). Phylogenetic reconstructions were performed using MEGA 5.1 (Tamura et al., 2011) for neighbor-joining (NJ) analyses. The non-synonymous to synonymous rate ratio ω (d_N/d_S) provided an indication of the change in selective pressures. The d_N/d_S ratios of 1, <1, and >1 are indicative of neutral evolution, purifying selection, and positive selection, respectively. We applied the codon substitution models implemented in the CODEML program in the PAML package (Yang, 2007). Two models were used in the PAML analysis: 1) the branch model, which accommodated variable ω ratios among particular lineages of interest and 5 different classifications; and 2) the site model, which allowed the ω ratio to vary among sites (Yang et al., 2000) and incorporated several null and alternative models, including M0 (null), M1a (nearly neutral), M2a (positive selection) (Wong et al., 2004), M3 (discrete), M7 (beta), and M8 (beta and ω) (Yang et al., 2000). The likelihood ratio tests (LRT) of 3 pairwise comparisons (i.e., comparisons of M1a *vs* M2a, M0 *vs* M3, and M7 *vs* M8) determined whether particular models would provide a significantly better fit. When the LRT suggested positive selection, the Bayes empirical Bayes method was used to detect the sites subjected to positive selection (Yang et al., 2005).

RESULTS

Sequence characterization of the primate *GPR50* genes

We first obtained 2 nucleotide sequences of GPR50 (complete coding sequence, CDS) from the Tibetan macaque (1887 bp) and Sichuan snub-nosed monkey (1866 bp) by PCR, cloning, and sequencing (accession Nos. KF207916 and KF220655). Moreover, we obtained 11 corresponding sequences from other primate species from GenBank. In total, 13 major representative primates, including the prosimian monkey (PM), New World monkeys (NWM), Old World monkeys (OWM), lesser apes, great apes, and humans, were compared. Two hundred and forty-five variable sites were detected, of which 61 were parsim-informative and 183 were singleton sites. All had 4 highly conserved regions. We noticed a repeat (SVHFK) between the amino acids 460 and 492 by sequence alignment analysis. Another remarkable repeated (RA) insert arose in the common marmoset, and a 12-bp deletion occurred in the modern human between amino acids 1074 and 1083. There was no insert or deletion in the transmembrane (TM) domain (Figure 1). Mutations were not identified in the TM6 region. Histidine was conserved

in the TM5 region of all melatonin receptors (Dubocovich et al., 2010) and retained in the TM5 of GPR50, although it did not bind to melatonin. At the amino acid Asp22 position, aspartic acid was replaced by Glu in humans. Interestingly, at position 621 of the amino acid sequence, alanine was substituted by asparagine in the OWM, great apes, and lesser apes but replaced by serine in humans. The most variable sites were identified in the long CTD.

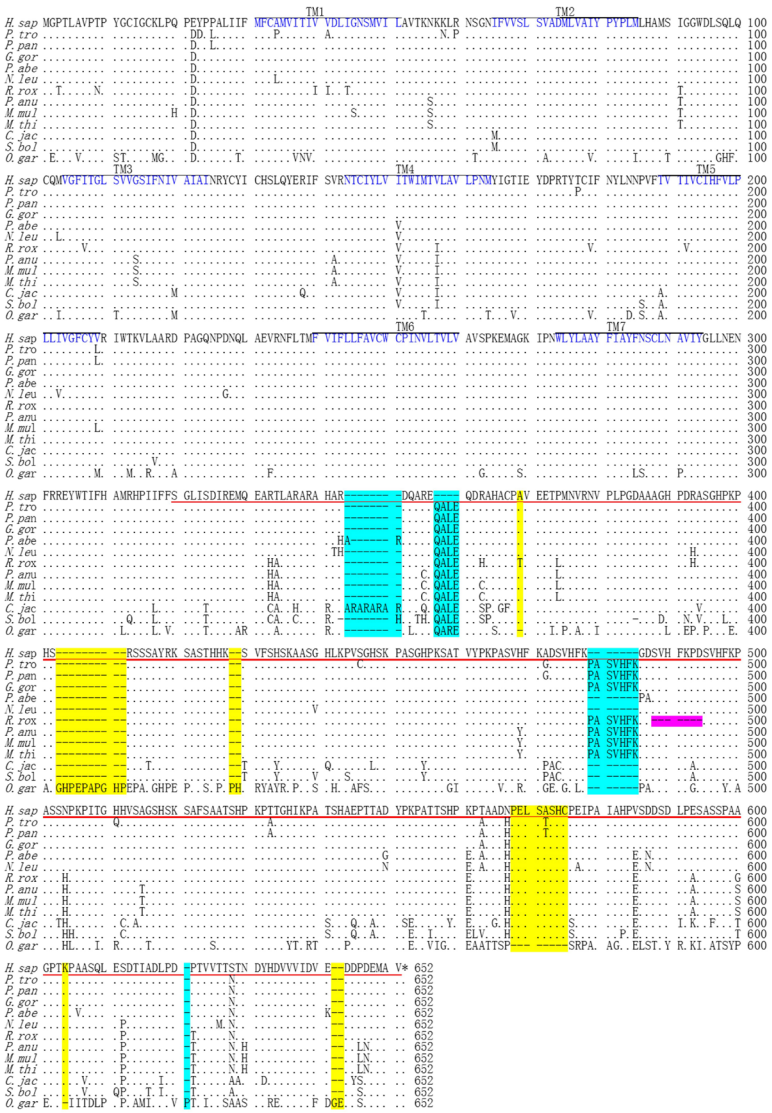

Figure 1. Alignment of the amino acid sequences of *GPR50* in 13 primates. Seven transmembrane regions are shown by transmembrane (TM) and blue font (TM ~17); C-terminal domin (CTD) is marked by the red underline. Insert or deletions (indel) in several species are highlighted by cyan and *O. gar* species-specific indels are highlighted by yellow. *H. sap*, *Homo sapiens*; *P. tro*, *Pan troglodytes*; *P. pan*, *Pan paniscus*; *G. gor*, *Gorilla gorilla*; *P. abe*, *Pongo abelii*; *N. leu*, *Nomascus leucogenys*; *M. mul*, *Macaca mulatta*; *M. thi*, *Macaca thibetana*; *P. anu*, *Papio anubis*; *R. rox*, *Rhinopithecus roxellana*; *S. bol*, *Saimiri boliviensis*; *C. jac*, *Callithrix jacchus*; *O. gar*, *Otolemur garnettii*.

We explored the signs of sequence characterization. The *ti/tv* rate ratio was 2.39 under the Hasegawa et al. (1985) model (+G) with pairwise sequence comparison, indicating a low *ti/tv* preference. Most disparity indices between sequences were near 0, suggesting relatively small differences in base composition biases. The sequence homology generated by pairwise sequence alignments revealed that the identity levels of amino acid sequences and nucleotides among the major primates ranged from 81.2 to 99.4% and from 72.3 to 99.8%, respectively (Table 1).

Table 1. Identity of nucleotide and amino acid sequence between 13 primate *GPR50* genes.

	H. sap	*P. tro*	*P. pan*	*G. gor*	*P. abe*	*N. leu*	*M. mul*	*M. thi*	*P. anu*	*R. rox*	*S. bol*	*C. jac*	*O. gar*
H. sap	-	98.9	99.4	99.4	98.6	97.8	96.6	96.9	97.0	96.5	93.6	93.8	82.6
P. tro	97.3	-	99.4	98.9	98.0	97.3	96.0	96.2	96.3	95.9	93.0	93.1	82.3
P. pan	98.6	98.6	-	99.4	98.5	97.7	96.5	96.7	96.8	96.3	93.5	93.6	82.8
G. gor	99.3	97.9	99.3	-	98.6	97.8	96.5	96.8	96.9	96.4	93.8	93.9	82.8
P. abe	97.9	96.6	97.9	98.6	-	97.9	96.3	96.5	96.7	96.2	93.5	93.5	82.9
N. leu	96.8	95.6	96.8	97.4	97.6	-	95.7	96.0	96.1	95.9	93.3	93.2	82.4
M. mul	95.1	93.8	95.1	95.4	95.1	94.3	-	99.6	99.2	97.4	92.7	92.7	81.7
M. thi	95.6	93.9	95.3	95.9	95.6	94.8	99.4	-	99.5	97.7	92.9	93.0	81.9
P. anu	95.8	94.1	95.4	96.1	95.8	94.9	99.3	99.8	-	97.8	93.0	93.1	82.0
R. rox	95.6	94.1	95.3	95.9	95.6	95.1	96.1	96.6	96.6	-	92.5	92.7	82.1
S. bol	90.9	89.4	90.7	91.2	90.9	90.4	90.4	90.9	90.9	90.6	-	95.8	81.2
C. jac	90.6	88.9	90.2	90.7	90.4	89.6	90.2	90.7	90.7	90.4	93.9	-	81.9
O. gar	74.0	73.1	74.2	74.3	74.5	73.7	72.8	73.1	73.3	73.1	72.6	72.3	-

Deleted or inserted nucleotides were excluded from the calculation of divergence. The homology of amino acids is shown above diagonal and that of nucleotide sequence below diagonal. For abbreviations, see legend to Figure 1.

Selective pressure analysis of the *GPR50* gene in primates

The NJ tree reconstructed using the CDS of 13 primate *GPR50* genes displayed only subtle differences in misplacement between Rhesus and Tibetan macaques when compared to the previously established primate species tree (Goodman et al., 1998), suggesting that the sequences obtained are orthologous. Thus, we utilized the established species tree in all subsequent analyses (Figure 2).

In order to investigate the evolution of the primate *GPR50* gene, we estimated the ω ratio of non-synonymous to synonymous substitution rates of the genes in 13 species using a likelihood method. We first assumed a uniform ω for all branches of the 13 primates (model A in Table 2); here, ω was estimated to be 0.4033, which is significantly <1 ($P < 0.01$; see the comparison with model B in Table 2). This result suggests that, overall, *GPR50* is under strong purifying selection in these primates. Subsequently, we tested a model that allowed for different values of ω for the 6 major primate lineages (human, great ape, lesser ape, OWM, NWM, and PM; model D in Table 2) compared to a model that assumed they have the same ω values (model C in Table 2); model D fit the data significantly better than model C ($P = 0.039$), indicating different levels of selective pressure on *GPR50* among the 6 major primate lineages. We then examined a model in which every branch had its own ω (model H). This model was significantly better than the uniform ω model A ($P = 0.0014$), suggesting that the variation of ω among the 13 different species lineages was significant.

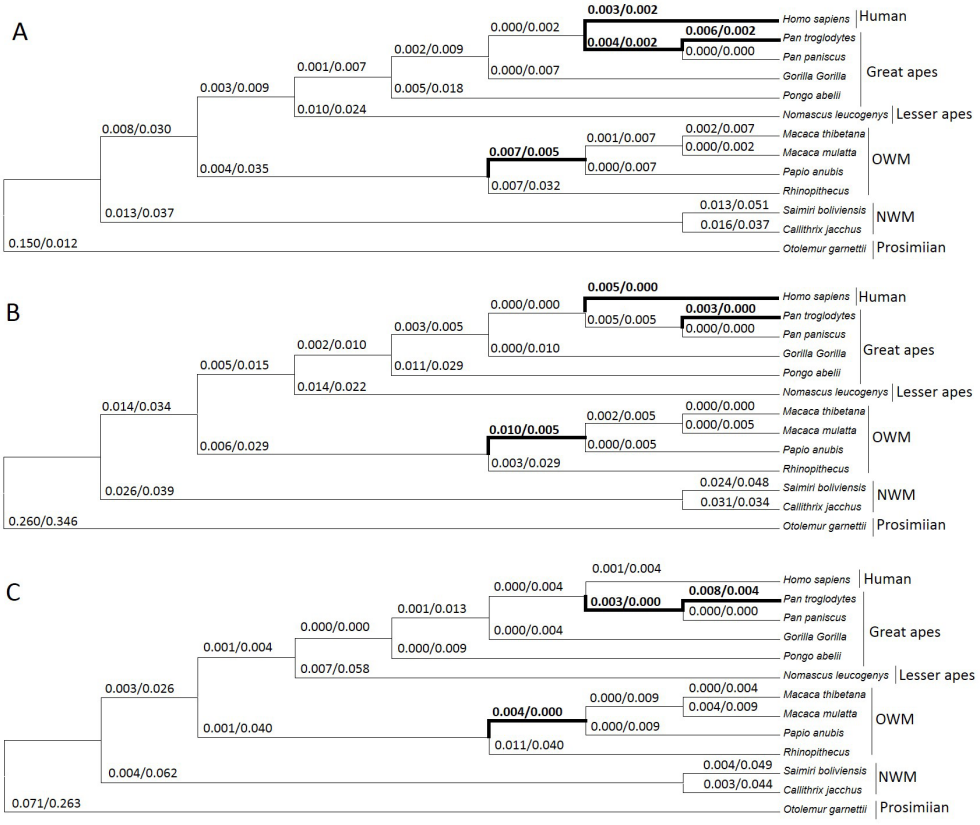

Figure 2. Evolutionary rates of the *GPR50* gene in major primate lineages. Tree topology followed Goodman et al. (1998). The numbers shown along each branch correspond to maximum-likelihood estimates of the numbers of non-synonymous or synonymous substitution along that branch. Omegas greater than 1 are in bold. **A.** d_N/d_S ratios for the entire coding sequences. **B.** d_N/d_S ratios for the CTD. **C.** d_N/d_S ratios for the non-CTD. OWM = Old World monkeys; NWM = New World monkeys.

Table 2. Likelihood ratio test of selective pressure on *GPR50* in primates using branch models in PAML.

Models	Parameter estimates	lnL[a]	np[b]	Models compared	2Δ(lnL)[c]	P value
A: All branches have the same ω	$\omega = 0.4033$	-5456.66	25			
B: All branches have one $\omega = 1$	$\omega = 1$	-5505.40	24	B *vs* A	97.48	P < 0.01
C: Braches in human, great apes, lesser apes, old world monkeys, new world monkeys, and prosimian monkey have the same ω. Other branches have their own omega.	$\omega = 0.4181$	-5455.96	26			
D: Braches in human, great apes, lesser apes, Old World monkeys, New World monkeys, and prosimian monkey have $\omega1$, $\omega2$, $\omega3$, $\omega4$, $\omega5$, and $\omega6$, respectively. Other branches have their own omega	$\omega1 = 999.00$ $\omega2 = 0.3569$ $\omega3 = 0.4047$ $\omega4 = 0.2579$ $\omega5 = 0.3347$ $\omega6 = 0.4761$	-5450.11	31	C *vs* D	11.7	P = 0.039
H: Each branch has its own ω	Variable ω in each branch	-5433.10	47	A *vs* H	47.12	P = 0.0014

[a]Natural logarithm of the likelihood value; [b]number of parameters; [c]twice the difference in lnL between the two models compared.

To further explore whether some key local regions presented different evolutionary patterns, we divided the complete sequence into 3 regions (i.e., extracellular, transmembrane, and cytoplasmic). The results for the ω value were <1 (ω = 0.30273, 0.13709, and 0.59088, respectively; Table 3), indicating that the three local regions are also under strong purifying selection. We calculated the ω values of the CTD and non-CTD. For all human sequences, we found that $d_N > d_S$; however, further calculation indicated $d_N > d_S$ only in the CTD. This indicates that relaxed selection pressure is primarily acting on the C-terminal sites of the human evolutionary processes. The one-tailed *Z*-tests (positive hypothesis: $d_N > d_S$) of the CTD showed that the human lineage has significantly deviated from the expectation of adaptive evolution ($P < 0.05$).

Table 3. Likelihoods (lnL) and parameter estimates for structure-specific regions under $\omega = 1$ and one estimated ω.

Region	lnL ($\omega = 1$)	lnL (one ω)	2dL	P (d.f. = 1)
Extracellular	-658.35	-648.98 (0.30273)	18.74	P < 0.01
Transmembrane	-1083.41	-1041.93 (0.13709)	82.96	P < 0.01
Cytoplasmic	-3622.07	-3611.59 (0.59088)	20.96	P < 0.01

2dL = twice the difference in lnL between the two models compared; d.f. = degrees of freedom.

To further investigate selection pressure on the primate *GPR50* gene, we used the site model in the PAML package. Six non-synonymous substitution site models were run in CODEML (Table 4). The first comparison model of M0 and M3 was used to test whether the ω value among the sites was variable. LRTs showed that M3 was a significantly better fit than M0 ($P < 0.01$). The result showed that the selective pressure was highly variable among sites. The LRTs estimated a similar value for a comparison of models M1a and M2a (-5136.16). The LRT was $2\Delta(\text{lnL}) = 0$, which refused the M2a and allowed the M1a model. Therefore, we can infer that 68.8 and 31.2% of the sites of the *GPR50* gene are subjected to purifying and neutral selections, respectively; there were no sites subjected to positive selection. In addition, an analogous result was found in the comparison of the M7 and M8 models. The test results were not significant ($P > 0.05$). Thus, the model did not support positive selection.

Table 4. Likelihood ratio test of selective pressure on *GPR50* in primates using site models in PAML.

Models	Estimate of parameters	lnL[a]	Models compared	2Δ(lnL)[b]	Positively selected sites
M0	$\omega = 0.38533$	-5148.86			
M1a	$p0 = 0.68763$, $p1 = 0.31237$ $\omega 0 = 0.19178$, $\omega 1 = 1.00000$	-5136.16			Not allowed
M2a	$p0 = 0.68763$, $p1 = 0.18389$, $p2 = 0.12848$ $\omega 0 = 0.19178$, $\omega 1 = 1.00000$, $\omega 2 = 1.00000$	-5136.16	M1a *vs* M2a	0	
M3	$p0 = 0.26703$, $p1 = 0.70718$, $p2 = 0.02579$ $\omega 0 = 0.00000$, $\omega 1 = 0.53385$, $\omega 2 = 2.77055$	-5135.77	M0 *vs* M3	26.18 ($P = 2.911 \times 10^{-5}$)	
M7	$p = 0.51904$, $q = 0.67306$	-5136.35			Not allowed
M8	$p0 = 0.98768$, $p = 0.67702$, $q = 0.96036$ ($p1 = 0.01232$), $\omega = 3.36944$	-5135.81	M7 *vs* M8	1.08 (P = 0.58275)	

[a]Natural logarithm of the likelihood value; [b]twice the difference in lnL between the two models compared.

Evolutionary rate of the *GPR50* gene

The evolutionary rates of the major primate lineages were calculated based on the well-established divergence times for primates (Goodman et al., 1998) (Table 5). The non-

synonymous rates ranged from 0 to 3.02 in the different primate lineages, with an average of 0.711. The average synonymous rate (1.619) was much higher than the average non-synonymous rate, which is concordant with the conservation of *GPR50* in primates. We conducted the relative rate test developed by Tajima (1993) and observed significant rate differences among the major primate lineages ($P < 0.05$). However, significant rate differences were not shown in the great-ape primate lineage, excluding *P. tro* and *P. pan*. The equality of the evolutionary rate was conducted between sequences A (*P. tro*) and B (*P. pan*), with sequence C (*H. sap*) used as an outgroup in Tajima's relative rate test (Tajima, 1993). The χ^2 test result was 9.00 ($P = 0.00270$ with 1 d.f.) with a P value <0.05. Significant differences were identified, implying that *P. tro* and *P. pan* have different evolutionary rates.

Table 5. Evolutionary rates of the *GPR50* gene in seven major primate lineages.

Primate lineage	Non-synonymous rate	Synonymous rate
Homo (6)	0.500 ± 0.167	0.333 ± 0.333
Pan (6)	0.617 ± 0.257	0.377 ± 0.371
Gorilla (7)	0.000 ± 0.000	1.286 ± 0.571
Pongo (14)	0.357 ± 0.143	1.286 ± 0.429
OWM (25)	0.160 ± 0.080	1.480 ± 0.360
NWM (40)	0.325 ± 0.075	0.850 ± 0.200
PM (50)	3.020 ± 0.240	5.720 ± 0.540
Average value	0.711 ± 0.137	1.619 ± 0.401

Calculation was based on the sequence divergence in the primate lineages. For example, the evolutionary rate of the homo lineage means that average substitutions between human and the putative ancestor divided by the known divergence time and the rate unit is substitutions per site per 10^9 years. OWM = Old World monkey; NWM = New World monkey; PM = prosimian monkey.

DISCUSSION

Molecular evolution of the *GPR50* gene in primates has been poorly studied. For the first time, we conducted a molecular evolutionary analysis of the *GPR50* gene in major primate lineages. R (*ti*/*tv*) and disparity index per site of nucleotide sequences revealed low *ti*/*tv* preference and base composition biases. A relatively high similarity between nucleotide and amino acid sequences of the *GPR50* gene was also found among the major primate groups. The results showed that *GPR50* is highly conserved at the molecular level. Rich Ser and 2 repeats (RA and SVHFK) were also found in the CTD (Figure 2). Dufourny et al. (2008) analyzed the CTD by using a bidimensional Hydrophobic Cluster Analysis. They first aligned the repeated heptapeptide and found a heptad in the C-terminal extension of human *GPR50*. Then, they compared it to the repeated heptad observed in the CTD of RNA polymerase II (RNAPII). The result suggested that the repeats might constitute a flexible scaffold for the binding of partner(s) to recognize specific phosphorylation sites, and the structure includes the SVHFK repeats. However, in the present study, the repeats were revealed as inserts or deletions in primate evolutionary processes; another obvious repeat was the four-RA insert, which only exists in *C. jacchus*.

The ω values for the sequences of entire coding and local regions in major primates were calculated to test whether the conservation of *GPR50* resulted from purifying selection according to the maximum likelihood method (Yang, 2007). We found that the *GPR50* genes in major primate lineages have undergone strong purifying selection. It is important to note

that selective tests were performed among 6 major primate lineages, which yielded independent ω values for each lineage. When *Pan* and *Homo* were classified as hominine, they yielded the same ω value, and the 6 lineages presented different levels of selective pressure. At the same time, the value $\omega > 1$ was found in the Hominini lineages, implying that accelerated evolution may have occurred in these lineages. Thus, they have undergone special adaptive evolution. The tribe of Hominini included *Pan* and *Homo* (Mann and Weiss, 1996). After the branch-model comparison, the models indicate that each major primate lineage has its own best-fit ω value. The relative rate test showed similar results; it also observed that $d_N > d_S$ in the *Pan* and *Homo* branches from the evolutionary rates of GPR50 in the major primate lineages (see Table 5). Remarkably, the results showed significant rate differences between the *P. troglodytes* and *P. paniscus* branch lineages ($P > 0.05$) by the relative rate test, which may be attributed to the differential availability of food sources (Macdonald, 2001). The results of the ML pairwise comparison in PAML showed that the value of $\omega > 1$ was significant in the Hominini, although the ω value of any two members from the human, bonobo, and chimpanzee lineages were also >1. This result indicates that relaxed selection pressure may have occurred in the Hominini lineage. *GPR50* belongs to the G protein-coupled melatonin receptor family, and a previous study provides a potential molecular link between the circadian/seasonal hormone melatonin and seasonal affective disorder (Delavest et al., 2011). Previous studies have discussed seasonality in primates because of their different intake methods (Brockman and Van Schaik, 2005). To achieve energy balance, all species of primates adjust their energy metabolism to adapt to changing intake conditions, and *GPR50* plays an important role in regulating entry into torpor. For this reason, there may be different selective pressures acting on the different lineages of *GPR50*.

Additionally, we did not identify any sites that were under positive selection using the Bayes empirical Bayes method. A similar result was found in a previous study. Parameter estimates suggest that 71% of the sites evolved under purifying selection ($\omega_0 = 0$), whereas 29% of the sites were identified under the neutrality assumption for the *GPR50* gene (Dufourny et al., 2008). A recent study reported that the *MT1* gene is primarily under purifying selection when compared to the *MT2* gene, which is under positive selection (He et al, 2012). Although the association between *GPR50* and *MT2* did not modify the function of *MT2*, *GPR50* abolished high-affinity agonist binding and G protein coupling to the MT1 protomer in the heterodimer (Levoye et al., 2006).

In summary, our study show that *GPR50* genes still undergo strong purifying selection in major primate lineages, no positive sites have been found in this study, the result implied that the GPR50 is extremely conserved during primate evolution. We also demonstrated that the purifying selection acted on the CTD domains of GPR50. In the future, analyses of additional species should contribute to better understanding of the ecology and the selective agent(s) acting on the evolution of *GPR50* genes.

ACKNOWLEDGMENTS

Research supported by the National Natural Science Foundation of China (#31370407, #30970383), the Program for Changjiang Scholars and Innovative Research Team in Sichuan Agricultural University (#IRT0848), and the Innovative Research Team in University of Sichuan Bureau of Education.

REFERENCES

Bechtold DA, Sidibe A, Saer BR, Li J, et al. (2012). A role for the melatonin-related receptor GPR50 in leptin signaling, adaptive thermogenesis, and torpor. *Curr. Biol.* 22: 70-77.

Brockman DK and Van Schaik CP (2005). Seasonality in primates: studies of living and extinct human and non-human primates. Cambridge University Press, New York.

Delavest M, Even C, Benjemaa N, Poirier MF, et al. (2011). Association of the intronic rs2072621 polymorphism of the X-linked *GPR50* gene with affective disorder with seasonal pattern. *Eur. Psychiatry* 27: 369-371.

Drew JE, Barrett P, Williams LM, Conway S, et al. (1998). The ovine melatonin-related receptor: cloning and preliminary distribution and binding studies. *J. Neuroendocrinol.* 10: 651-661.

Drew JE, Barrett P, Mercer JG, Moar KM, et al. (2001). Localization of the melatonin-related receptor in the rodent brain and peripheral tissues. *J. Neuroendocrinol.* 13: 453-458.

Dubocovich ML, Delagrange P, Krause DN, Sugden D, et al. (2010). International union of basic and clinical pharmacology. LXXV. Nomenclature, classification, and pharmacology of G protein-coupled melatonin receptors. *Pharmacol. Rev.* 62: 343-380.

Dufourny L, Levasseur A, Migaud M, Callebaut I, et al. (2008). GPR50 is the mammalian ortholog of Mel1c: evidence of rapid evolution in mammals. *BMC Evol. Biol.* 8: 105.

Goodman M, Porter CA, Czelusniak J, Page SL, et al. (1998). Toward a phylogenetic classification of Primates based on DNA evidence complemented by fossil evidence. *Mol. Phylogenet. Evol.* 9: 585-598.

Graur D and Li WH (2000). Fundamentals of Molecular Evolution. 2nd edn. Sinauer Associates, Sunderland.

Gubitz AK and Reppert SM (1999). Assignment of the melatonin-related receptor to human chromosome X (*GPR50*) and mouse chromosome X (*GPR50*). *Genomics* 55: 248-251.

Hasegawa M, Kishino H and Yano T (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol*. 22: 160-174.

He CB, Jiang, XP and Liu GQ (2012). Positive selection driven the differentiation of melatonin receptor *MT1* and *MT2* genes. *China Agric. Sci.* 45: 4084-4091.

Ivanova EA, Bechtold DA, Dupre SM, Brennand J, et al. (2008). Altered metabolism in the melatonin-related receptor (GPR50) knockout mouse. *Am. J. Physiol. Endocrinol. Metab.* 294: E176-182.

Izzo G, Francesco A, Ferrara D, Campitiello MR, et al. (2010). Expression of melatonin (MT1, MT2) and melatonin-related receptors in the adult rat testes and during development. *Zygote* 18: 257-264.

Levoye A, Dam J, Ayoub MA, Guillaume JL, et al. (2006). The orphan GPR50 receptor specifically inhibits MT1 melatonin receptor function through heterodimerization. *EMBO J.* 25: 3012-3023.

Lundstrom K (2006). Latest development in drug discovery on G protein-coupled receptors. *Curr. Protein Pept. Sci.* 7: 465-470.

Macdonald D (2001). The new encyclopedia of mammals. Oxford University Press, New York.

Macintyre DJ, McGhee KA, Maclean AW, Afzal M, et al. (2010). Association of GPR50, an X-linked orphan G protein-coupled receptor, and affective disorder in an independent sample of the Scottish population. *Neurosci. Lett.* 475: 169-173.

Mann A and Weiss M (1996). Hominoid phylogeny and taxonomy: a consideration of the molecular and fossil evidence in an historical perspective. *Mol. Phylogenet. Evol.* 5: 169-181.

Marinissen MJ and Gutkind JS (2001). G-protein-coupled receptors and signaling networks: emerging paradigms. *Trends Pharmacol. Sci.* 22: 368-376.

McCarthy C (1998). Chromas Version 1.45. School of Health Sciences, Griffith University Gold Coast Campus, Southport, Australia.

Reppert SM, Weaver DR, Ebisawa T, Mahle CD, et al. (1996). Cloning of a melatonin-related receptor from human pituitary. *FEBS Lett.* 386: 219-224.

Tajima F (1993). Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics* 135: 599-607.

Tamura K, Peterson D, Peterson N, Stecher G, et al. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731-2739.

Thomson PA, Wray NR, Thomson AM, Dunbar DR, et al. (2005). Sex-specific association between bipolar affective disorder in women and GPR50, an X-linked orphan G protein-coupled receptor. *Mol. Psychiatry* 10: 470-478.

Wong WS, Yang Z, Goldman N and Nielsen R (2004). Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. *Genetics* 168: 1041-1051.

Yang Z (2007). Paml 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24: 1586-1591.

Yang Z, Nielsen R, Goldman N and Pedersen AMK (2000). Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155: 431-449.

Yang Z, Wong WS and Nielsen R (2005). Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol. Biol. Evol.* 22: 1107-1118.