Molecular Cloning and Evolutionary Analysis of *GJB6* in Mammals

Binghua Ru · Naijian Han · Guimei He ·
Kathryn Brayer · Shuyi Zhang · Zhe Wang

Abstract *GJB6* plays a crucial role in hearing. In mammals, bats use ultrasonic echolocation for orientation and locating prey. To investigate the evolution of *GJB6* in mammals, we cloned the full-length coding region of *GJB6* from 16 species of bats and 4 other mammal species and compared them with orthologous sequences in 11 other mammals. The results show purifying selection on *GJB6* in mammals, as well as in the bat lineage, which indicates an important role for *GJB6* in mammal hearing. We also found one unique amino acid substitution shared by 16 species of bats and 10 shared by two species of artiodactyls. This positioned the artiodactyls at an abnormal location in the gene tree. In addition, the cytoplasmic loop and carboxy terminus were more variable than other domains in all the mammals. These results demonstrate that *GJB6* is basically conserved in mammals but has undergone relatively rapid evolution in particular lineages and domains.

B. Ru · G. He · S. Zhang · Z. Wang (✉)
Institute of Molecular Ecology and Evolution, Institutes for Advanced Interdisciplinary Research,
East China Normal University, Shanghai 200062, China
e-mail: zhe.email@gmail.com

B. Ru
e-mail: waoe0.1@163.com

G. He
e-mail: guimeihe@163.com

S. Zhang
e-mail: syzhang@bio.ecnu.edu.cn

N. Han
Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China
e-mail: hannaijian@163.com

K. Brayer
Yale Systems Biology Institute, Department of Ecology and Evolutionary Biology, Yale University,
West Haven, CT 06516, USA
e-mail: kathryn.brayer@yale.edu

Introduction

In vertebrates, gap junctions facilitate gap junctional communication, a process that allows the transport of ions, nutritive material, secondary messengers, and small molecules of up to 1 kDa in size directly between adjacent cells (Goodenough et al. 1996). Gap junctions are dynamic multifunctional complexes that are essential for normal vertebrate development and physiology. In the inner ear, gap junctions are divided into the epithelial cell gap junction system and the connective tissue gap junction system; these gap junctions play a crucial role in generation and maturation of endocochlear potential (Kikuchi et al. 1995).

Gap junctions are normally formed when six connexin proteins assemble to form a hemichannel (connexon), and two hemichannels form a complete gap junction (Kumar and Gilula 1996). One of the predominant connexins in the cochlea, Connexin30, is found in the supporting cells of the organ of Corti, the basal cells and the intermediate cells in the stria vascularis, and the fibrocytes in the spiral ligament (Lautermann et al. 1998, 1999; Ahmad et al. 2003; Forge et al. 2003). Gap junction networks involving Connexin30 are very important for hearing because they facilitate the recycling of potassium ions from the hair cells back into the cochlear endolymph during the auditory transduction processes (Kikuchi et al. 1995; Zhao et al. 2006). Connexin 30 is also broadly found in many other tissues and organs, such as the brain, epidermis, uterus, testicles, esophagus, and kidney in humans, rats, and mice (Dahl et al. 1996; Kelley et al. 1999; Nagy et al. 1999).

GJB6, the gene encoding the protein Connexin30, is located at 13q11–q12.1 in humans (Essenfelder et al. 2005). Mutations in *GJB6* are associated with hearing loss and skin diseases. For example, ablation of the Connexin30 protein in cochlea leads to hearing loss (Cohen-Salmon et al. 2002; Teubner et al. 2003). Similarly, nonsyndromic deafness is observed in patients carrying mutations in their *GJB6* coding regions (Rabionet et al. 2002). Grifa et al. (1999) found a threonine-to-methionine change at position 5 (T5M) in an Italian family affected by bilateral middle/high-frequency hearing loss. Other amino acid mutations in *GJB6* are found in Clouston syndrome (hidrotic ectodermal dysplasia), a disorder affecting mainly palmoplantar skin, nails, and hair (Lamartine et al. 2000; Smith et al. 2002) and causing abnormal connexin hemichannel activity that leads to leakage of ATP into the extracellular medium (Essenfelder et al. 2004). KID syndrome, which is characterized by keratitis, ichthyosis, and deafness, is also genetically heterogeneous and caused by mutations in *GJB6* (Jan et al. 2004). Together, these studies support the idea that *GJB6* plays an important role in both auditory transduction and epidermal differentiation.

Owing to their capacity for echolocation and powered flight, bats are an unusual and specialized group of mammals. When they are flying and foraging, the specialized hearing system of bats plays an important role in receiving and processing echo information (Jones and Teeling 2006). To aid in our understanding of the origin and evolution of echolocation in bats, we performed an evolutionary

analysis of *GJB6*. In this study, we cloned the full-length coding region of the *GJB6* gene from 16 bat species and 4 other mammals, and compared them with 11 mammalian orthologs obtained from public databases to test the evolutionary selection pressure on this gene. Using two methods, we constructed molecular phylogenetic trees of *GJB6* in mammals. We also used molecular evolution approaches to determine whether key evolutionary transitions in *GJB6* occurred in mammals and to identify signatures of natural selection.

Materials and Methods

Taxonomic Coverage

Tissue samples for 20 mammal species were obtained from the School of Life Science, East China Normal University, Shanghai, China. These species included 16 bats (*Myotis ricketti*, *Rhinolophus ferrumequinum*, *R. pearsoni*, *R. pusillus*, *R. macrotis*, *R. rex*, *Taphozous melanopogon*, *Chaerephon plicata*, *Miniopterus magnater*, *Pteronotus parnellii*, *P. macleayii*, *P. quadridens*, *Hipposideros armiger*, *Carollia perspicillata*, *Artibeus jamaicesis*, and *A. lituratus*), two carnivores (*Felis catus* and *Arctonyx collaris*), one perissodactyl (*Equus asinus*), and one artiodactyl (*Capra hircus*).

From GenBank, we obtained *GJB6* sequences for four rodents (*Mus musculus* NM_001010937; *Rattus norvegicus* NM_053388; *Cricetulus griseus* AY2826256; and *Mesocricetus auratus* AY570781), one primate (*Homo sapiens* NM_006783), and one artiodactyl (*Bos taurus* NM_001015546). We also obtained Ensembl predicted *GJB6* sequences from published genomes (http://genome.cusc.edu/ and http://www.ensembl.org) for two primates (*Pan troglogytes* and *Otolemur garnettii*), one carnivore (*Canis familiaris*), one perissodactyl (*Equus caballus*), and one rodent (*Spermophilus tridecemlineatus*).

Cloning and Sequencing the Whole Coding Sequence of *GJB6*

DNA was extracted from six species of bats (*Carollia perspicillata*, *Pteronotus parnellii*, *P. macleayii*, *P. quadridens*, *Artibeus jamaicesis*, and *A. lituratus*) using a DNeasy Blood and Tissue Kit (Qiagen). Total RNA was prepared from the whole brain from the other 14 species of mammals using the RNAiso kit (Takara D312, Japan) and then was reverse synthesized using SuperScript III reverse transcriptase (Invitrogen 18080-051, USA) to make cDNA. We designed one pair of primers (F1 and R1) to obtain part of the coding sequence from *Myotis ricketti*, *Rhinolophus ferrumequinum*, and *Rousettus leschenaulti* (Table 1). Then we performed 5′ and 3′ rapid amplification of cDNA ends (RACE) using a Smart RACE cDNA amplification kit (Clontech 634914, USA) to obtain the whole coding sequence of *GJB6* and the untranslated regions (UTRs). Based on the partial coding sequence obtained, we designed gene-specific primers R2 and F2 for 5′ and 3′ cDNA ends of *Myotis ricketti*, R7 and F2 for 5′ and 3′ cDNA ends of *Rhinolophus ferrumequinum*, and R3 and F3 for 5′ and 3′ cDNA ends of *Rousettus leschenaulti*. Finally, based on the conserved

Table 1 Primer sequences used in this study

Primer	Sequence
F1	5′-CAGCAT(C/A)GGGAAGGTGTGGA-3′
R1	5′-GCAAATCACAGA(C/T)GC(A/G)GAAATCATA-3′
R2	5′-GGGACACGGGGAAGAAGTGGTCATAGCA-3′
F2	5′-GCCGCCCGCAAATTCAGACGAGGAGAGA-3′
R3	5′-TTGATGTCCTCCAGGTCCTTGAACTCGC-3′
F3	5′-CACTTCTTCCCCGTGTCCCACATCCG-3′
F4	5′-CG(G/A)(A/C)(G/A)CAGGAATAA(G/A)C(G/C)(G/A)A(C/T)ACAATG-3′
R4	5′-AGTT(C/T)TGTTAGGAGC(C/T)(A/G)GCTAC(C/A)T(C/A)G(C/A)AC-3′
F5	5′-AGG(G/A)(C/A)(C/T)(G/A)A(G/A)C(C/G)(G/A)(G/A)(C/T)(G/A)C(C/A)ATGGA (C/T)TGGG-3′
R5	5′-AGC(T/C)(G/A)GCTC(G/A)(C/T)CGC(G/A)CT(C/T)(C/T)G(G/A)A(G/A)(C/T)GTCTA-3′
F6	5′-G(G/A)(G/A)T(G/A)A(A/G)(G/C)(C/G)(A/G)(G/A)(C/T)(A/G)(G/C)(C/A)ATGGA (C/T)TGGG-3′
R6	5′-T(A/G)T(T/C)CTACA(G/A)(G/A)A(A/C)(G/T)T(T/G)AACTCA(G/A)GT(G/A)TCTA-3′
R7	5′-GACAAAGTCTTCCTGCTCGTCACCCCAAAC-3′

regions of UTRs in the three bat species and the mammal species for which *GJB6* sequences were available on public databases, we designed three pairs of primers (F4 and R4 for *Taphozous melanopogon*, *Chaerephon plicata*, *Miniopterus magnater*, *Pteronotus parnellii*, *P. macleayii*, *P. quadridens*, *Hipposideros armiger*, *Rhinolophus pearsoni*, *R. pusillus*, *R. macrotis*, and *R. rex*; F5 and R5 for *Carollia perspicillata*, *Artibeus jamaicesis*, *A. lituratus*, *Eonycteris spelaea*, and *Felis catus*; and F6 and R6 for *Arctonyx collaris*, *Equus asinus*, and *Capra hircus*) to obtain the whole coding regions of *GJB6* from DNA or cDNA templates (Table 1).

Amino Acid Alignment

We used ClustalX 1.81 software (Thompson et al. 1997) to align the nucleotide and amino acid sequences. Mega4 (Tamura et al. 2007) was used to analyze the divergence of nucleotides and amino acids among the mammalian lineages with the Kimura 2-parameter model, as well as to calculate pairwise comparisons of genetic distances and the Poisson correction model for amino acid data. The Nei–Gojobori method (Nei and Gojobori 1986) was employed to calculate the synonymous and nonsynonymous substitutions per synonymous (dS) and nonsynonymous (dN) sites among lineages using Mega4 and 500 replications for bootstrap testing. We also used Mega4 to estimate the ratio and the saturation of transition and transversion on three codon sites and the third codon sites.

Phylogenetic Construction of *GJB6*

Before phylogenetic construction, the program Model Test 3.6 (Posada and Crandall 1998) was used to estimate the most appropriate nucleotide substitution model and

parameters for the maximum likelihood method. Then we applied the Bayesian method using MrBayes 3.1 software (Huelsenbeck and Ronquist 2001) to construct the phylogenetic tree of *GJB6* for 31 species of mammals. In the control block of the input file for MrBayes 3.1, we set the generations equal to 1,000,000 and 6 Markov chains for simulation, discarding 500,000 generations before the simulation was predicted to reach a steady condition. We also applied the maximum likelihood method to construct another phylogenetic tree using PhyML 3.0 (Guindon and Gascuel 2003).

Evolutionary Analysis

A mammalian species tree used for the evolutionary analysis was constructed according to the topology of previous publications (Murphy et al. 2001; Teeling et al. 2005). We first used the branch-specific models to test for different selection pressures among the lineages of bats and other mammals using the CodeML program in PAML4 (Yang 2007). We then tested the free-ratio model, which assumes an independent ω ratio for each branch in the tree. Next, we used the site-specific models to test the selection pressure on the codon sites of the *GJB6* gene during the divergence and evolution of the mammal lineages. The site-specific models involved model M0 (one-ratio model), which assumes an equal dN/dS ratio among all the amino acid sites. This model can be compared with model M3 (discrete), which assumes a variable dN/dS ratio among different parts of the amino acid sites, to test the possibility of positive natural selection affecting the protein evolution. Comparisons between the model M7 (beta) and model M8 (beta and ω) and between model M1 and model M2 were used to identify positively selected sites. Finally, the likelihood ratio test (LRT) was used to appraise the null hypothesis and alternative hypothesis.

Sliding Window

A sliding window was used to highlight where the rates of dN exceeded the rates of dS in specific functional regions in the gene and, hence, to identify regions of positive selection. It was also used to compare variability of dN, dS, and ω ratios along different clades or different functional regions in the gene. We performed the sliding window analysis (window size 30 bp, step size 6 bp) for *GJB6* using Nei–Gojobori estimates of synonymous and nonsynonymous substitutions per site and the dN/dS (ω) in Swaap 1.0.2 (Pride 2000).

Results

Sequence Analysis

To compare the *GJB6* gene sequences among lineages of mammals, we obtained the whole coding sequence (786 bp) of the *GJB6* genes from 16 species of bats, two

carnivores, one perissodactyl, and one artiodactyl. These sequences were submitted to GenBank under accession nos. GU326363–GU326382.

In general, alignment of the amino acid sequences of *GJB6* from 31 species of mammals showed the highly conserved character of this gene. We found no variation in the length of the amino acid sequences among the six orders of mammals tested here, except for an amino acid deletion in five species of bats (*Artibeus jamaicesis*, *A. lituratus*, *Carollia perspicillata*, *Taphozous melanopogon*, and *Chaerephon plicata*). *GJB6* amino acid sequences are also conserved over the six orders of mammals tested here. The analysis showed that 205 of 261 amino acid sites were conserved, encompassing more than 78.5% of the whole gene. Among bats, 238 amino acid sites were identified as conserved, and less than 8.8% of amino acids were variable sites. The calculation of the pairwise amino acid distance, using the Poisson correction model in Mega4, showed that the highest amino acid difference among bats was 0.055.

At the nucleotide level, 524 of 786 nucleotide sites (66.7%) were identical, and 262 nucleotide sites were variable (33.3%), along the whole sequence length for all mammals in this study. Among the bats, 176 nucleotide sites were identified as variable. The overall average of transition/transversion based on the *GJB6* gene sequences from the various lineages of mammals was 2.294, indicating that these sequences could be used for phylogenetic construction. The evaluation of the degree of saturation showed a linear increase in transition and transversion changes, with increasing sequence divergence for complete sequences. In addition, the data indicate that transition and transversion substitutions are not saturated at third codon sites.

We found a T5A mutation that occurred only in *Arctonyx collaris*. We found only one bat-specific amino acid substitution, Arg131, located on the inner loop between the second and third transform membrane domains. The number of artiodactyl-specific amino acid substitutions (10) was striking (I10V, I25V, I28L, I30V, K111T, N113S, I145V, N230A, H231P, and D250E). Three amino acid substitutions located in the first transform membrane domain were caused by a nucleotide mutation (I25V and I30V by A to G and I28L by A to C) in the first codon position. In the third transform membrane domain, a I145V substitution was also caused by a nucleotide mutation (A to G) at the first codon position. Both of these domains are considered conservative (Essenfelder et al. 2005). The remaining substitutions were located in the highly variable cytoplasmic loop (I10V, K111T, N113S) and the carboxy terminus (N230A, H231P, D250E).

Phylogenetic Construction

The Model Test selected the best fit model TIM3ef with a gamma distribution (G). According to the topologies of the trees constructed by two programs, MrBayes (Fig. 1) and PhyML (Fig. 2), the evolution of the *GJB6* gene is highly conserved in mammals. The gene trees show high identity with the species tree within the order of bats, but not for the out-groups. Artiodactyls and rodents show evidence of long-branch attraction (a phenomenon in which rapidly evolving lineages are inferred to be closely related, regardless of their true evolutionary relationships). In addition,
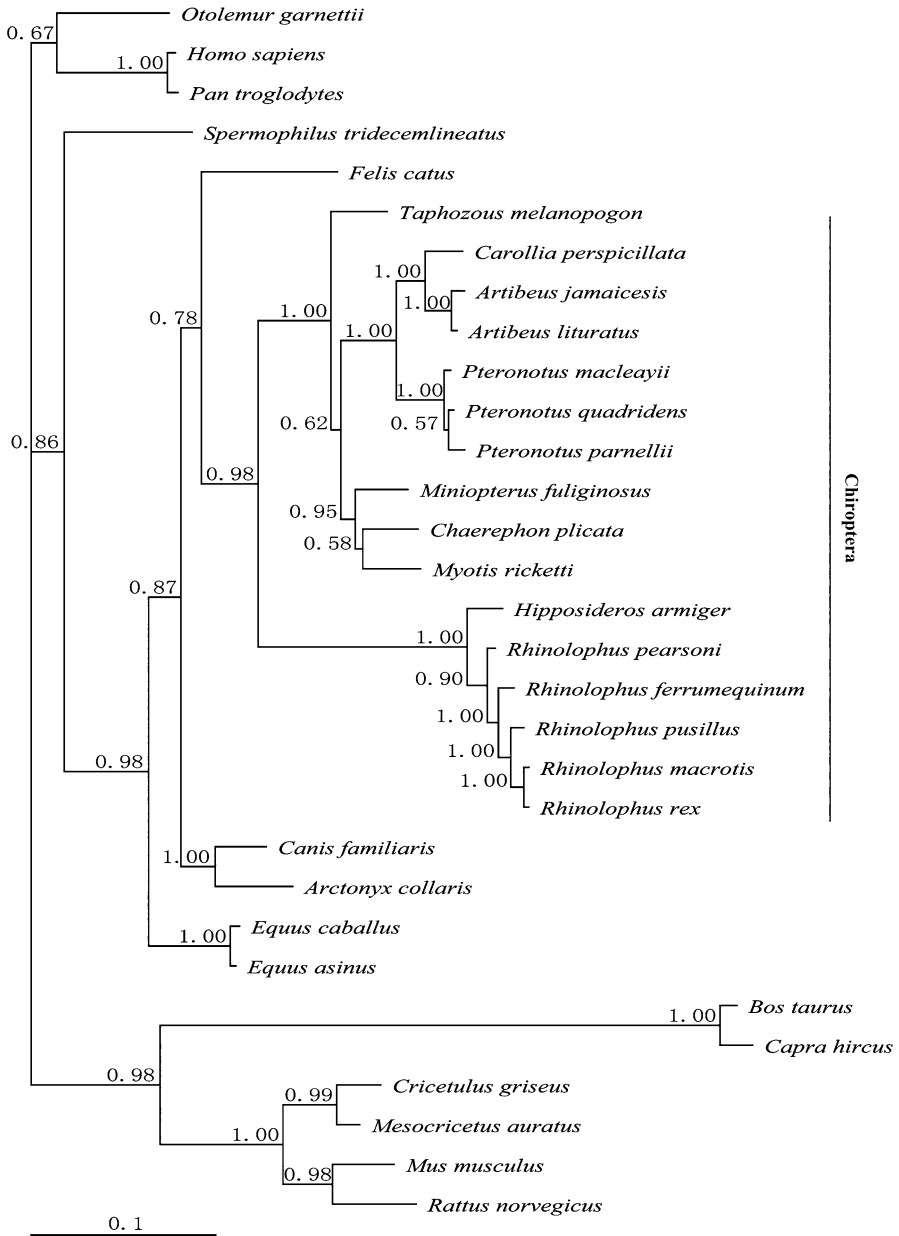
Fig. 1 Unrooted maximum posterior probability Bayesian tree based on complete sequences of *GJB6* genes from various mammal lineages, using MrBayes 3.1. The *line on the right* indicates the 16 bat species. The *values on the nodes* indicate the posterior probability

Felis catus and *Spermophilus tridecemlineatus* are present in the wrong positions according to the advanced molecular phylogenetic reconstruction of the lineages of mammals (Murphy et al. 2001), but the posterior probabilities are relatively low.
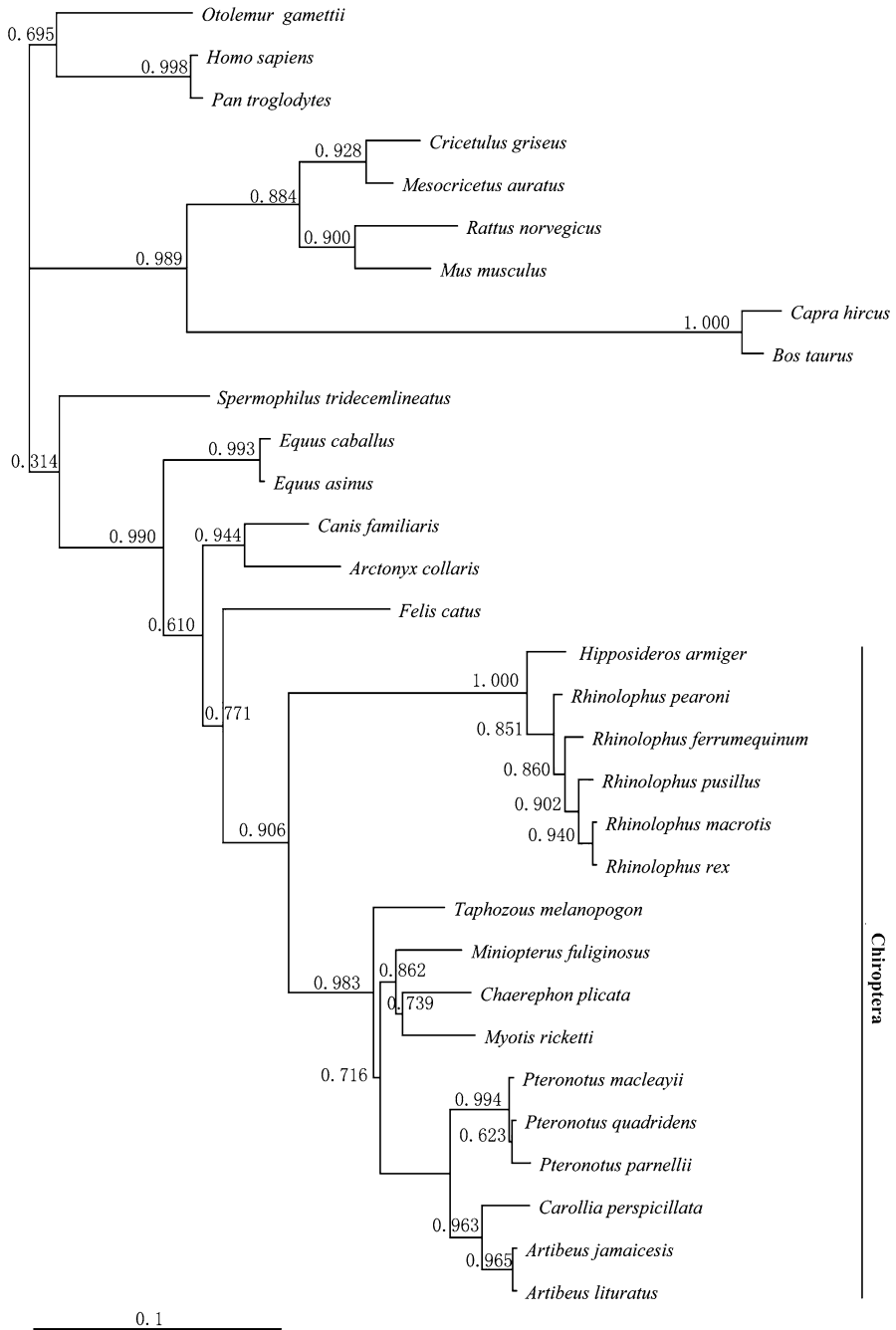
Fig. 2 Unrooted maximum likelihood tree based on complete sequences of *GJB6* genes from mammal lineages, using PhyML 3.0. The *line on the right* indicates the 16 bat species. *The values on the nodes* indicate the bootstrap support

Evolutionary Analysis

We used the maximum likelihood method to estimate the ω ratio of *GJB6* among lineages of mammals, which was regarded as the branch-specific model. Log-likelihood values and relative parameters were estimated by maximum likelihood under various models (Table 2). The one-ratio model (M0), which constrains all lineages in the phylogeny to the same dN/dS ratio, estimated dN/dS at 0.026, showing that the average synonymous substitution rate is much higher than the rate of nonsynonymous substitution and that the *GJB6* gene has undergone elementary purifying selection in general during mammal evolution. This model gave a log maximum-likelihood value of −4552.98. In contrast, the free-ratio model assumes an independent ω ratio for each branch in the tree. In the species tree, 59 branches were assumed to have their own ω ratios under the free-ratio estimation, which leads to a log maximum-likelihood value of −4515.105220. In comparison with the one-ratio model, the free-ratio model involved 58 extra parameters. Therefore, we

Table 2 Models of variation of the *GJB6* sequence in 31 mammals

Model	Parameters in the ω distribution	ln L^{a}	Parameter[b]	Positively selected sites
Free-ratio	119	−4515.10522	$\kappa = 3.29684$	Not allowed
M0 (one-ratio)	61	−4552.980907	$\kappa = 3.23346$, $\omega = 0.02646$	Not allowed
M1 (nearly neutral)	62	−4548.907685	$\kappa = 3.34190$	Not allowed
			$P_0 = 0.98549$ ($P_1 = 0.01451$)	
			$\omega_0 = 0.02295$ ($\omega_1 = 1$)	
M2 (positive selection)	64	−4548.907685	$\kappa = 3.34190$	None
			$P_0 = 0.98549$, $P_1 = 0.01451$ ($P_2 = 0.00000$)	
			$\omega_0 = 0.02295$ ($\omega_1 = 1$), $\omega_2 = 4.33830$	
M3 (discrete)	65	−4528.813227	$\kappa = 3.23856$	None
			$P_0 = 0.25050$, $P_1 = 0.62485$ ($P_2 = 0.12466$)	
			$\omega_0 = 0.00995$, $\omega_1 = 0.00995$, $\omega_2 = 0.16748$	
M7 (beta)	62	−4529.541567	$\kappa = 3.23448$	Not allowed
			$P = 0.23268$, $q = 6.97593$	
M8 (beta and ω)	64	−4529.541827	$\kappa = 3.23448$	None
			$P_0 = 1.00000$ ($P_1 = 0.00000$)	
			$P = 0.23268$, $q = 6.97593$, $\omega = 2.46539$	

[a] Log-likelihood value

[b] k estimate of transition/transversion rate ratio

used LRT to compare these two models with a chi-square distribution with 58 degrees of freedom. We found that the one-ratio model fits significantly better than the free-ratio model ($2\Delta\ell = 75.74$, df = 58, $P > 0.05$), which indicates consistent dN/dS ratios among different lineages of mammals. The branch-specific model, however, does not support the hypothesis that there is positive selection to accelerate the evolutionary rate of the *GJB6* gene in any specific mammal lineage.

We also performed site-specific models to test if there is positive selection on particular amino acid sites. This survey was achieved by three comparisons among six models: M3 vs. M0, M2 vs. M1, and M8 vs. M7 (Yang 2007). The results show that the LRT between model M3 and model M0 is significant ($2\Delta\ell = 48.34$, df = 4, $P < 0.001$), suggesting that discrete selection pressures act on the different sites of the *GJB6* gene, but none of the sites was identified as being under positive selection. The LRT between models M1 and M2 did not support the hypothesis that model M2 was better than model M1 ($2\Delta\ell = 0$, df = 2, $P > 0.05$), and none of the sites was identified as being under positive selection. The LRT between model M8 and M7 did not support the hypothesis that model M8 was better than model M7 ($2\Delta\ell = 0$, df = 2, $P > 0.05$), none of the sites was identified as being under positive selection. Altogether, these models found no positively selected sites.

Sliding Window

Sliding window analysis is another way to identify regions under positive selection. A sliding window can highlight where the rates of nonsynonymous substitution exceed the rates of synonymous substitution in specific functional regions and the variability of dN, dS, and ω ratios along different clades or different functional regions of the gene. The results of the sliding window analysis showed purifying selection on the whole sequence of *GJB6* in mammals (Fig. 3). The cytoplasmic loop and carboxy terminus, however, were more variable than other domains that are so conserved that the ω ratios were all nearly zero (Fig. 3).

Discussion

Among mammals, bats are faced with extreme demands imposed by echolocation. In this study, we compared the whole coding sequence of *GJB6* from 16 bat species with those of 15 other mammal species. Similar to other mammals, the bat sequences showed high conservation, except in the cytoplasmic loop and carboxy terminus. Our results indicate that this gene is also very conserved among all the mammals according to the evolutionary analysis. The average synonymous mutation rate is far higher than the nonsynonymous mutation rate, which means that natural selection has tended to stabilize this protein and eliminate nucleotide mutations that cause disruptive changes in the protein's functions. There are still some mutations that should be highlighted, however.

We found a T5A mutation in *Arctonyx collaris*. It is reported that a T5M mutation causes nonsyndromic autosomal dominant deafness (Grifa et al. 1999). Although the
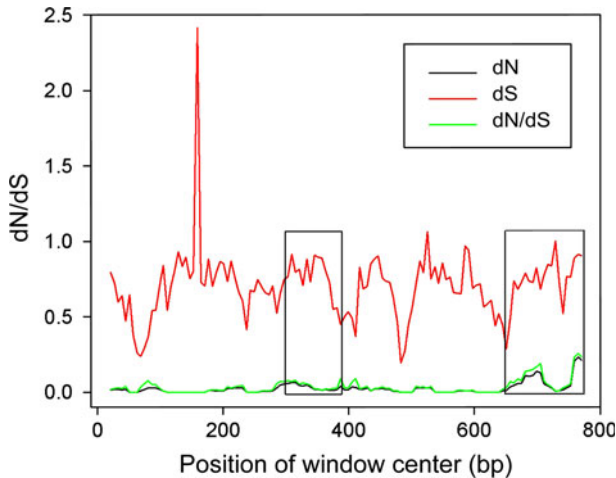
Fig. 3 Variation in the coding region of *GJB6* for 31 species of mammals. *Black (bottom) line*: Nonsynonymous (d*N*) substitution rate. *Red (top) line*: Synonymous (d*S*) substitution rate. *Green (lightest) line*: d*N*/d*S* ratio (ω). The cytoplasmic loop region is enclosed by the first (*left*) box, and the region of carboxy terminus is enclosed by the second (*right*) box. The ω rates are more variable and higher in the cytoplasmic loop and carboxy terminus than in other regions of the gene (color figure online)

ears of *A. collaris* are small, its hearing is very good (Parker 1979). This suggests that the T5A mutation is not detrimental to the bat's hearing, unlike in humans.

The results of the sliding window analysis correspond with previous reports that the parts of the cytoplasmic loop and carboxy terminus are highly variable. These regions are thought to be important for the regulation of connexin gene expression (Kumar and Gilula 1996). In addition, every domain of the protein has special amino acid sites and specific functions (Krutovskikh and Yamasaki 2000), such as a small domain of 6–7 residues in the inner loop between the second and third transform membrane domains that is critical for the permeability of gap junction channels in Connexin43 (Krutovskikh et al. 1998). We found a bat-specific amino acid substitution, amino acid Arg131, located on the inner loop between the second and third transform membrane domains, an area considered to have an important influence on pH and voltage gating sensitivity in cells (Wang et al. 1996). This substitution was caused by a nucleotide mutation (T to G) at the first codon position in the ancestral bats. Further experimental work is needed to test the relationship between this amino acid substitution and bat-specific hearing characteristics.

We also found several bat species that possess an amino acid gap in the carboxy terminus, which is the most variable region among the connexins and is also thought to be important for regulation of expression (Kumar and Gilula 1996). Although there are some phosphorylation sites in this domain, previous studies have suggested that phosphorylation might not be essential for the normal function of connexins (Krutovskikh and Yamasaki 2000). The potential effect of this amino acid deletion also needs to be studied further.

In artiodactyls, 10 artiodactyl-specific amino acid substitutions were found in different functional regions, and that is many more than were found in other orders. This may be the reason for the long-branch attraction between artiodactyls and rodents. There is no evidence to indicate that the evolutionary rate was distinctly faster in artiodactyls than in other mammals. The function of these artiodactyl-specific amino acid substitutions and the special evolution of *GJB6* in this order should be studied further.

The results of the saturation tests show that the transition and transversion of whole data and third codon site data do not reach saturation, which could explain the misplacement of the artiodactyls in the phylogenetic reconstruction. The phylogenetic construction results of *GJB6* showed high concordance between the PhyML tree and the MrBayes tree. In both of these gene trees, high statistical support (Bayesian posterior probability higher than 0.95) was obtained for the monophyly of the major clades of the bats. The estimated evolutionary ratios of all the branches in the phylogenetic tree by the free-ratio model show that purifying selection is the predominant factor in the evolution of the *GJB6* gene in mammals. The phylogenetic construction showed that the *GJB6* gene is conserved in Chiroptera, and there are no obvious changes after the origin of bats, indicating that *GJB6* plays an important role in the formation of the hearing system in bats.

In conclusion, we cloned and sequenced the *GJB6* gene from 16 bats and four other mammals and studied the evolution of the gene in six mammalian orders. The evolutionary analysis showed the high power of purifying selection in mammals. We found a novel amino acid substitution on the specific branches of bats, which has the potential to affect function. Finally, we showed relatively rapid evolution and more amino acid substitutions of *GJB6* in artiodactyls.

References

Ahmad S, Chen S, Sun J, Lin X (2003) Connexins 26 and 30 are co-assembled to form gap junctions in the cochlea of mice. Biochem Biophys Res Commun 307:362–368

Cohen-Salmon M, Ott T, Michel V, Hardelin JP, Perfettini I, Eybalin M, Wu T, Marcus DC, Wangemann P, Willecke K, Petit C (2002) Targeted ablation of connexin26 in the inner ear epithelial gap junction network causes hearing impairment and cell death. Curr Biol 12:1106–1111

Dahl E, Manthey D, Chen Y, Schwarz HJ, Chang YS, Lalley PA, Nicholson BJ, Willecke K (1996) Molecular cloning and functional expression of mouse connexin-30, a gap junction gene highly expressed in adult brain and skin. J Biol Chem 271:17903–17910

Essenfelder GM, Bruzzone R, Lamartine J, Charollais A, Blanchet-Bardon C, Barbe MT, Meda P, Waksman G (2004) Connexin30 mutations responsible for hidrotic ectodermal dysplasia cause abnormal hemichannel activity. Hum Mol Genet 13:1703–1714

Essenfelder GM, Larderet G, Waksman G, Lamartine J (2005) Gene structure and promoter analysis of the human GJB6 gene encoding connexin 30. Gene 350:33–40

Forge A, Becker D, Casalotti S, Edwards J, Marziano N, Nevill G (2003) Gap junctions in the inner ear: comparison of distribution patterns in different vertebrates and assessment of connexin composition in mammals. J Comp Neurol 467:207–229

Goodenough DA, Goliger JA, Paul DL (1996) Connexins, connexons, and intercellular communication. Annu Rev Biochem 65:475–502

Grifa A, Wagner CA, D'Ambrosio L, Melchionda S, Bernardi F, Lopez-Bigas N, Rabionet R, Arbones M, Della Monica M, Estivill X, Zelante L, Lang F, Gasparini P (1999) Mutations in *GJB6* cause nonsyndromic autosomal dominant deafness at DFNA3 locus. Nat Genet 23:16–18

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704

Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755

Jan AY, Amin S, Ratajczak P, Richard G, Sybertz VP (2004) Genetic heterogeneity of KID syndrome: identification of a C×30 gene (GJB6) mutation in a patient with KID syndrome and congenital atrichia. J Invest Dermatol 122:1108–1113

Jones G, Teeling EC (2006) The evolution of echolocation in bats. Trends Ecol Evol 21:149–156

Kelley PM, Abe S, Askew JW, Smith SD, Usami S, Kimberling WJ (1999) Human connexin 30 (*GJB6*), a candidate gene for nonsyndromic hearing loss: molecular cloning, tissue-specific expression, and assignment to chromosome 13q12. Genomics 62:172–176

Kikuchi T, Kimura RS, Paul DL, Adams JC (1995) Gap junctions in the rat cochlea: immunohisto-chemical and ultrastructural analysis. Anat Embryol 191:101–118

Krutovskikh V, Yamasaki H (2000) Connexin gene mutations in human genetic diseases. Mutat Res 462:197–207

Krutovskikh VA, Yamasaki H, Tsuda H, Asamoto M (1998) Inhibition of intrinsic gap-junction intercellular communication and enhancement of tumorigenicity of the rat bladder carcinoma cell line BC31 by a dominant-negative connexin 43 mutant. Mol Carcinog 23:254–261

Kumar NM, Gilula NB (1996) The gap junction communication channel. Cell 84:381–388

Lamartine J, Munhoz Essenfelder G, Kibar Z, Lanneluc I, Callouet E, Laoudj D, Lemaitre G, Hand C, Hayflick SJ, Zonana J, Antonarakis S, Radhakrishna U, Kelsell DP, Christianson AL, Pitaval A, Der Kaloustian V, Fraser C, Blanchet-Bardon C, Rouleau GA, Waksman G (2000) Mutations in GJB6 cause hidrotic ectodermal dysplasia. Nat Genet 26:142–144

Lautermann J, ten Cate WJ, Altenhoff P, Grummer R, Traub O, Frank H, Jahnke K, Winterhager E (1998) Expression of the gap-junction connexins 26 and 30 in the rat cochlea. Cell Tissue Res 294:415–420

Lautermann J, Frank HG, Jahnke K, Traub O, Winterhager E (1999) Developmental expression patterns of connexin26 and -30 in the rat cochlea. Dev Genet 25:306–311

Murphy WJ, Eizirik E, O'Brien SJ, Madsen O, Scally M, Douady CJ, Teeling E, Ryder OA, Stanhope MJ, de Jong WW, Springer MS (2001) Resolution of the early placental mammal radiation using Bayesian phylogenetics. Science 294:2348–2351

Nagy JI, Patel D, Ochalski PA, Stelmack GL (1999) Connexin30 in rodent, cat and human brain: selective expression in gray matter astrocytes, co-localization with connexin43 at gap junctions and late developmental appearance. Neuroscience 88:447–468

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3:418–426

Parker C (1979) Birth, care and development of Chinese hog badgers. Int Zoo Yearb 19:182–185

Posada D, Crandall KA (1998) Model test: testing the model of DNA substitution. Bioinformatics 14:817–818

Pride (2000) Swaap: a tool for analyzing substitutions and similarity in multiple alignments. Version 1.0.2. http://www.bacteriamuseum.org/SWAAP/SwaapPage.htm

Rabionet R, Lopez-Bigas N, Arbones ML, Estivill X (2002) Connexin mutations in hearing loss, dermatological and neurological disorders. Trends Mol Med 8:205–212

Smith FJ, Morley SM, McLean WH (2002) A novel connexin 30 mutation in Clouston syndrome. J Invest Dermatol 118:530–532

Tamura K, Dudley J, Nei M, Kumar S (2007) Mega4: molecular evolutionary genetics analysis software version 4.0. Mol Biol Evol 24:1596–1599

Teeling EC, Springer MS, Madsen O, Bates P, O'Brien SJ, Murphy WJ (2005) A molecular phylogeny for bats illuminates biogeography and the fossil record. Science 307:580–584

Teubner B, Michel V, Pesch J, Lautermann J, Cohen-Salmon M, Sohl G, Jahnke K, Winterhager E, Herberhold C, Hardelin JP, Petit C, Willecke K (2003) Connexin30 (Gjb6)-deficiency causes severe hearing impairment and lack of endocochlear potential. Hum Mol Genet 12:13–21

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Wang X, Li L, Peracchia LL, Peracchia C (1996) Chimeric evidence for a role of the connexin cytoplasmic loop in gap junction channel gating. Pflugers Arch 431:844–852

Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586–1591

Zhao HB, Kikuchi T, Ngezahayo A, White TW (2006) Gap junctions and cochlear homeostasis. J Membr Biol 209:177–186