Available online at www.sciencedirect.com

SCIENCE DIRECT

Genomics 87 (2006) 724–732

GENOMICS

www.elsevier.com/locate/ygeno

# Functional consequences of naturally occurring DRY motif variants in the mammalian chemoattractant receptor GPR33☆

Holger Römpler [a], Hon-Tsen Yu [b], Antje Arnold [a], Annie Orth [c], Torsten Schöneberg [a,*]

[a] *Institute of Biochemistry, Molecular Biochemistry, Medical Faculty, University of Leipzig, Johannisallee 30, 04103 Leipzig, Germany*
[b] *Institute of Zoology and Department of Life Science, National Taiwan University, Taipei, Taiwan*
[c] *Laboratoire Génome Populations Interactions Adaptation, UMR 5171 CNRS, CC063, Universite Montpellier 2, 34095 Montpellier Cedex 05, France*

Received 22 September 2005; accepted 14 February 2006
Available online 3 April 2006

**Abstract**

Most members of the large family of rhodopsin-like G-protein-coupled receptors possess an evolutionarily conserved Asp-Arg-Tyr (DRY) motif in the C-terminal region of the third transmembrane domain. Mutations of residues within this motif usually abolish receptor function and, when they occur naturally, can even cause human diseases. By analyzing over 100 mammalian orthologs of the chemoattractant receptor GPR33 we identified several polymorphic and fixed sequence variations within the DRY motif. Unexpectedly, the naturally occurring mutation of $Arg^{3.50}$ to His in mouse GPR33 showed no difference from the wild-type receptor in several functional tests. Sequence analysis of GPR33 from Asian house mice revealed the polymorphic existence of $Arg^{3.50}$ and $His^{3.50}$ alleles in wild-trapped populations, further supporting the functional equivalence of both allelic variants. In contrast, the $Arg^{3.50}$ to Gly mutation found in hamster GPR33 inactivates the receptor and may have contributed to pseudogenization of this gene in this species. Functional data with GPR33 variants indicate different receptor- and context-specific consequences of DRY mutations. Our study also reveals GPR33 as a new example illustrating missense mutations as a first step in the pseudogenization process.


*Keywords:* Evolution; DRY motif; Chemokine; G-protein-coupled receptor; Pseudogene; Constitutive activity

Despite the rather low amino acid sequence homology between G-protein-coupled receptors (GPCR) and the remarkable structural variety of their natural agonists, all GPCR share a common molecular architecture consisting of seven transmembrane domains (TMD) connected by three extra- and three intracellular loops (ICL). This global architecture is maintained, in part, by amino acid motifs and highly conserved residues, which are also used to categorize GPCR into several families. Most members of the rhodopsin-like family of GPCR possess within their TMD core a number of highly conserved motifs, such as a DRY motif ($Asp^{3.49}$-$Arg^{3.50}$-$Tyr^{3.51}$; numbering refers to the absolute-position-independent nomenclature for GPCR introduced by Ballesteros and Weinstein [1]) at the TMD3/ICL2 transition and an N/DPXXY motif in TMD7. The DRY motif plays an essential role in GPCR function (reviewed in [2]). The crystal structure of rhodopsin and numerous experimental data suggest that the residues of the DRY motif form a cage-like structure through hydrogen bonding with residues in close proximity to the DRY motif and with surrounding residues of TMD [3,4].

Previous studies examined the contribution of individual DRY motif residues to receptor function and found two critical residues. First, mutation of $Asp^{3.49}$ results in constitutive activity of many receptors [5,6]. Second, the almost fully conserved $Arg^{3.50}$ in the DRY motif is considered a key residue in GPCR signal transduction since replacement of $Arg^{3.50}$ with different amino acids virtually abolishes G-protein coupling of many GPCR [7–10]. Consequently, several naturally occurring mutations have been found at position $Arg^{3.50}$ in patients

*Abbreviations:* AVPR2, V2 vasopressin receptor; DMEM, Dulbecco's modified Eagle's medium; GPCR, G-protein-coupled receptor; HA, hemagglutinin; ICL, intracellular loop; Myr, million years; mGPR33, murine GPR33; TMD, transmembrane domain.

☆ Sequence data from this article have been deposited with the GenBank Data Library under Accession Nos. DQ156944–DQ156966 (Supplemental Table S2).

* Corresponding author. Fax: +49 341 9722 159.
*E-mail address:* schoberg@medizin.uni-leipzig.de (T. Schöneberg).

0888-7543/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2006.02.009

suffering from diseases caused by inactivating mutations in distinct GPCR [11]. Further, missense mutations in the DRY motif have been implicated in pseudogenization of GPCR [12,13]. The fact that a single missense mutation in a GPCR can functionally inactivate the receptor is of importance since most pseudogene annotations, e.g., in the odorant GPCR field, are classically based on obvious signatures like premature truncation or frame disruptions to ascertain their nonfunctionality [14,15].

We recently examined the evolution of the orphan chemoattractant receptor GPR33 [16]. This chemoattractant GPCR appeared with the first mammals but became inactivated during the past 1 million years (Myr) in humans, as well as in several great ape and rodent species. The coincidental inactivation and fixation of GPR33 in several species of distantly related mammalian orders suggest a selective pressure on this chemoattractant receptor gene [16]. GPR33 appears to be intact in most mammals; however, for a rhodopsin-like GPCR, this receptor displays an unusually high diversification at the amino acid level, including the otherwise highly conserved DRY motif.

Here, we analyzed the functional consequences of sequence variations within the GPR33 DRY motif during evolution. Despite an absolute conservation of $Asp^{3.49}$ in over 100 cloned GPR33 orthologs, this residue was substituted to Ala or Gly in all investigated species of the genus *Mus* and the subfamily Gerbillinae. The functional consequence of this change is an increase in basal receptor activity. By screening several *Mus musculus* subspecies we identified an allelic variant ($Arg^{3.50}$His) in wild-captured Asian house mouse (*M. musculus castaneus*) individuals. This amino acid change is known to inactivate many GPCR [12,17–20]; however, it does not influence GPR33 function. Our evolutionary, experimental, and in vivo findings indicate that mutations in the DRY motif do not necessarily inactivate GPCR function. Therefore, the functional relevance of residues in highly conserved motifs, like the DRY motif, cannot be generalized to all rhodopsin-like GPCR.

## Results

### *The GRP33 DRY motif displays an unusually high sequence variability in mammals*

Sequences analysis of more than 100 mammalian GPR33 orthologs revealed a high amino acid sequence variability in comparison to other mammalian GPCR ortholog sets. This variability is also observed in motifs that are highly conserved among most rhodopsin-like GPCR. In previous work we showed that after the rat–mouse evolutionary split, $Asp^{3.49}$ in the DRY motif was substituted by Ala thereby increasing basal activity of murine GPR33 [16]. The variability of the conserved DRY motif became more apparent when the number of individuals examined was increased and further mammalian species were included in the analysis (Fig. 1). First, in all gerbil GPR33 orthologs investigated, $Asp^{3.49}$ is substituted by Gly (Fig. 1B). Second, $Arg^{3.50}$His is fixed in several gerbil species and polymorphic in both the Asian house mouse (*M. m. castaneus*) and Geoffroy's marmoset (*Callithrix geoffroyi*). Further, $Arg^{3.50}$ is substituted by Gly in Syrian hamster (*Microcricetus aureus*). Because $Asp^{3.49}$ and $Arg^{3.50}$ mutations have been implicated in diverse GPCR dysfunctions, we set out to evaluate the functional relevance of these naturally occurring DRY variants. This analysis may be of importance since missense mutations are very rarely linked to pseudogenization events.

### *Functional relevance of $Asp^{3.49}$ variants in mammalian GPR33 orthologs*

Two natural DRY motif variants, $Ala^{3.49}$ and $Gly^{3.49}$ (see above), were found at position 3.49. We recently showed that basal signal transduction abilities of orphan GPCR can be studied even without agonist stimulation when the receptor and appropriate chimeric G protein are coexpressed [21]. To examine signal transduction abilities of the $G_i$-coupled GPR33 and its mutant variants we coexpressed them with $G\alpha_{\Delta 6qi4myr}$ [22] in COS-7 cells and determined basal receptor activity in inositol phosphate (IP) assays. For control purposes the ADP receptor $P2Y_{12}$, which displays a high basal activity in this system [21], was included in all assays (Fig. 2).

In all species of the subfamily Murinae investigated, $Asp^{3.49}$ is substituted by Ala and an increased (2.9-fold over basal) basal activity of the mGPR33 was observed (Fig. 2) [16]. Reversal of $Ala^{3.49}$ (ARY) to $Asp^{3.49}$ (DRY) abolishes this elevated basal activity but has no influence on cell surface expression as determined by indirect cellular ELISA. We next assessed the $Gly^{3.49}$ mutant (GRY) found in all gerbil orthologs (see Fig. 1). Compared to wild-type mGPR33 (ARY), $Ala^{3.49}$Gly showed an increased basal activity (1.7-fold elevation above basal) accompanied by a reduced cell surface expression (~76% of mGPR33). These data indicate that, in the subfamily Gerbillinae and in the genus *Mus,* the increase in basal GPR33 activity occurred by independent mutation of the DRY motif.

Previous studies with other GPCR showed that several amino acid residues (e.g., Thr, Gln) at position 3.49 lead to constitutive receptor activity [23]. To test whether constitutive activation of GPR33 is specific to Ala and Gly substitutions, we generated $Thr^{3.49}$ (TRY) and $Gln^{3.49}$ (QRY) and analyzed their activity as described above. As shown in Fig. 2, both mutants displayed no significant increase in basal IP formation compared to the control or DRY variant. Determination of cell surface expression, an essential prerequisite for proper signal transduction, revealed a loss of plasma membrane expression for both mutant GPR33. This finding indicates that GPR33 activation by mutation of the DRY motif depends on the amino acid substitution.

### *The $Arg^{3.50}$His mutation occurs naturally in GPR33 of wild M. m. castaneus*

$Arg^{3.50}$ within the DRY motif is one of the most conserved residues in rhodopsin-like GPCR. However, our sequence analysis revealed several mutations of $Arg^{3.50}$, mainly to His

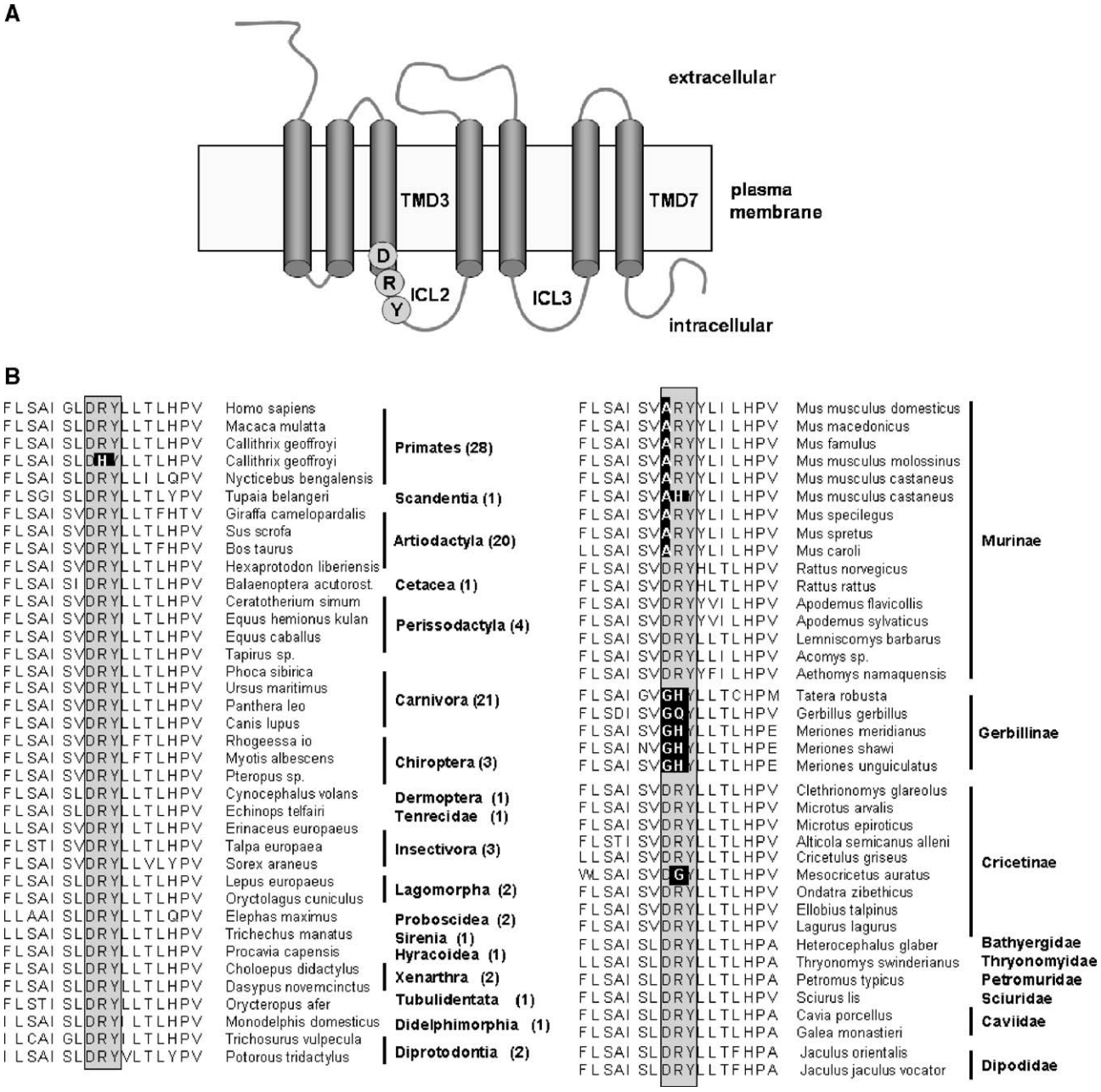


Fig. 1. Structural variability of the DRY motif in mammalian GPR33 orthologs. (A) The DRY motif is located at the TMD3/ICL2 transition of rhodopsin-like GPCR. (B) The amino acid sequence of the TMD3/ICL2 transition of various mammalian GPR33 orthologs contains a highly conserved DRY motif (gray highlight). Variations in the DRY motif are boxed. The numbers in parentheses indicate the numbers of species of that order also containing a DRY motif in GPR33 (for sequence details see Accession Nos. AY490569–AY490743, AY493989–AY494004, AY502103, AY502104, AY528865). The alignment on the right shows the DRY motif of only rodent species. ICL, intracellular loop; TMD, transmembrane domain.

(see Fig. 1), in gerbil, Syrian hamster, Asian house mouse, and Geoffroy's marmoset. Because such $Arg^{3.50}$His mutations are known to inactivate GPCR in vitro and in vivo [12,17–20], we speculated that the polymorphic occurrence of $Arg^{3.50}$His may reflect an ongoing pseudogenization of GPR33. We identified the $Arg^{3.50}$His allele in an inbred laboratory strain of *M. m. castaneus* (CIM strain according to the nomenclature of the Laboratoire Génome Populations Interactions Adaptation at the Universite Montpellier, France), which may reflect either natural polymorphisms present in the wild population or a new polymorphism that has been introduced during captivity. To distinguish between these possibilities, we sequenced the coding region from 88 wild *M. m. castaneus* trapped in Taiwan. Three allelic variants of GPR33 were found in these rural Asian house mice (Fig. 3A). Allele 1 and allele 2 are almost identical to those of *M. m. domesticus* (10 individuals tested), but allele 3 ($Arg^{3.50}$His allele) contains five substitutions, three of which are nonsynonymous. Analysis of the geographic distribution of allele 3 showed its abundant presence in all regions of Taiwan (Fig. 3B).

The $Arg^{3.50}$His mutation in the Asian house mouse took place after divergence from other *M. musculus* subspecies. To

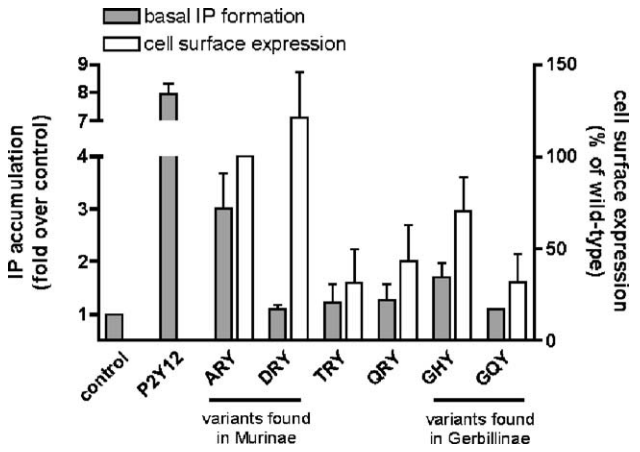


Fig. 2. Functional characterization of mutations at relative position 3.49 in mouse GPR33. In most mammalian GPR33 orthologs the relative position 3.49 is an Asp. However, in *Mus* and Gerbillinae Asp[3.49] is substituted by Ala and Gly, respectively. Replacement of the four to five C-terminal amino acids of $G\alpha_q$ with the corresponding $G\alpha_i$ residues (referred to as $G\alpha_{\Delta 6qi4myr}$) confers the ability to stimulate the PLC-β pathway onto $G_i$-coupled receptors [22]. To test for basal receptor activity, GPR33 receptor constructs were coexpressed with the chimeric $G\alpha_{\Delta 6qi4myr}$ protein in COS-7 cells and inositol phosphate (IP) assays were performed as described (Materials and methods). As controls, plasmids encoding GFP and the human $P2Y_{12}$ were cotransfected with $G\alpha_{\Delta 6qi4myr}$. Basal IP formation is expressed as fold over basal levels of GFP-transfected cells (544 ± 15 cpm/well). Data are presented as means ± SD of two independent experiments, each carried out in triplicate. Cell surface expression levels of the GPR33 mutant receptors were measured by indirect cellular ELISA (Materials and methods). Specific optical density (OD) readings (OD value of HA-tagged GPR33 construct minus OD value of GFP-transfected cells) are given as a percentage of wild-type HA-tagged GPR33. The nonspecific OD value (GFP) was 0.089 ± 0.031, and the OD value of the wild-type HA-tagged GPR33 was 0.485 ± 0.116. ELISA data are given as means ± SD of eight independent experiments, each carried out in quadruplicate.

delimit further the time point of this event, we sequenced the 5′ and 3′ noncoding genomic regions flanking the GPR33 coding region of *M. m. castaneus* (allele 1 and allele 3), *M. m. domesticus* (one allele), *Rattus norvegicus* (two alleles) [16], and *Rattus rattus.* Comparison of ~2.6 kb of homologous genomic sequence revealed an average of 140.4 ± 0.5 mutations/1000 bp between *Mus* alleles and the corresponding *Rattus* sequences (see Supplemental Table S1). The Asian house mouse allele 1 and allele 3 differ from each other by 6.6 substitutions/1000 bp. Thus, by comparison with the divergence of *Mus* and *Rattus* (about 12 to16–23 Myr ago) [24,25] and that of *R. rattus* and *R. norvegicus* (about 5.5 Myr ago) [26], we estimate that the divergence of the *M. m. castaneus* His[3.50]-containing allele 3 from both the Arg[3.50]-containing alleles of *M. m. domesticus* and *M. m. castaneus* (allele 1) occurred about 1.07 ± 0.37 Myr ago (range min 0.64 to max 1.56 Myr, depending on the time point for *Mus*/*Rattus* divergence [24,25]).

### *Arg to His mutation does not influence receptor function in vitro*

Previous studies with the V2 vasopressin receptor (AVPR2), the angiotensin type 1 receptor, and β2 adrenergic receptor showed that the Arg[3.50]His mutation leads to a loss of cell surface expression due to constitutive receptor internalization [20,27]. To analyze the functional consequence of Arg[3.50]His in mGPR33 we expressed wild-type (ARY) and mutant (AHY) receptors in COS-7 cells, performed second messenger assays, and measured cell surface expression by ELISA. Interestingly, both wild-type and Arg[3.50]His receptors showed similar cell surface expression levels and basal activities (Table 1). Similarly, cell surface expression measured in transiently transfected CHO cells (transfectable cell system with lower receptor expression levels when pcDps vector is used) was indistinguishable between the mutant (Arg[3.50]His) and the wild-type mGPR33 (data not shown).

In the *M. m. castaneus* GPR33 ortholog, the Arg[3.50]His mutation cosegregates with other synonymous and nonsynonymous substitutions (Fig. 3A). To exclude the possibility that Arg[3.50]His inactivates GPR33 only in the context of the *M. m. castaneus* allele 3 (see Fig. 3A) we cloned and expressed an epitope-tagged version of allele 3. Again, no differences in signaling abilities or cell surface expression were found compared with the wild-type mGPR33 (allele 1 in Fig. 3A, see Table 1).

As mentioned above, previous studies showed that in several GPCR the Arg[3.50]His mutation triggers constitutive β-arrestin-mediated desensitization. Even in the absence of agonist the sequestered receptors were found in β-arrestin-associated intracellular vesicles [27]. We examined GPR33-Arg[3.50]His localization with respect to β-arrestin-associated vesicles by confocal microscopy of transfected human embryonic kidney (HEK) cells. For control purposes AVPR2 was cotransfected with β-arrestin–GFP and stimulated with 10 nM vasopressin (AVP). As shown in Fig. 4, AVP stimulation leads to β-arrestin-mediated desensitization of AVPR2 by formation of internalized vesicles with colocalized AVPR2 and β-arrestin. The wild-type mGPR33 (ARY variant in Fig. 4) showed membrane localization while β-arrestin–GFP remained cytosolic. The same pattern of localization was also seen with the Arg[3.50]His mutant (AHY in Fig. 4), indicating the lack of constitutive internalization of the mutant receptor.

The Arg[3.50]His mutation-mediated constitutive receptor internalization was originally found in GPCR containing a classical DRY motif [20,27]. In the wild-type murine GPR33, Asp[3.49] of the DRY motif is substituted by Ala (ARY motif) [16]. To test whether the unchanged cell surface expression of GPR33-Arg[3.50]His is caused by this difference in the DRY motif we analyzed Arg[3.50]His in the DRY background (GPR33-DHY mutant). The DHY variant naturally occurs in marmoset (*C. geoffroyi,* see Fig. 1B). As shown in Table 1, cell surface expression of GPR33-DHY is slightly reduced compared with ARY and AHY. The decrease in cell surface expression was not due to a reduction in total cellular expression of the mutant receptor as determined by a sandwich ELISA (Table 1). In the Syrian hamster GPR33 ortholog Arg[3.50] is mutated to Gly in the DRY context (DGY, see Fig. 1B). In contrast to mGPR33-DHY, Arg[3.50] substitution by Gly results in an almost complete loss of cell surface expression (Table 1).

**A**

```
codon          109    120    126    162    227    248
               Ser    Ser    Arg    Pro    Arg    Val
               TCG....TCC....CGT....CCC....CGA....GTG
mouse ..........G......C.....G.......C.....G.....G
allele-1........G......C.....G.......C.....G.....G
allele-2........A......C.....G.......C.....G.....G

allele-3........G......T.....A.......G.....A.....C
               Ser    Ser    His    Pro    Gln    Leu
```

**B**

| location | numbers of individuals | frequency (%) of allele-1 (absolute number) | frequency (%) of allele-2 (absolute number) | frequency (%) of allele-3 (absolute number) |
|---|---|---|---|---|
| Northwest | 17 | 38.2 (13) | 17.6 (6) | 44.1 (15) |
| Northeast | 13 | 38.4 (10) | 7.7 (2) | 53.8 (14) |
| Central | 15 | 70.0 (21) | 13.3 (4) | 16.6 (5) |
| Southwest | 17 | 88.2 (30) | 0 (0) | 11.8 (4) |
| Southeast | 16 | 81.2 (26) | 0 (0) | 18.8 (6) |
| Outer islands | 10 | 60.0 (12) | 20.0 (4) | 20.0 (4) |
| **total** | **88** | **63.6 (112)** | **9.1 (16)** | **27.3 (48)** |

Fig. 3. GPR33 alleles in Asian house mouse and their frequency and geographic distribution in Taiwan. (A) Sequencing of the coding region of GPR33 from Asian house mouse (*M. m. castaneus*) revealed three alleles. Allele 1 is identical to that of *M. m. domesticus,* whereas allele 2 and allele 3 contain differences at the indicated codon positions. (B) To determine allele frequencies (given in %) 88 wild-captured Asian house mice from six different regions in Taiwan [31] were analyzed by sequencing.

To evaluate whether the small reduction in cell surface expression of mGPR33-DHY is due to β-arrestin-mediated desensitization, immunofluorescence studies (see above) were performed. As expected from ELISA data, both GPR33 variants, GPR33-DRY and GPR33-DHY, were expressed at the plasma membrane. Additionally, an increase in intracellular reticular fluorescence was found for GPR33-DHY (Supplemental Fig. S1). However, no differences in β-arrestin distribution and no β-arrestin-containing vacuoles were found in GPR33-DHY, indicating that $Arg^{3.50}$His mutation in the DRY context slightly interferes with proper cell surface expression.

To test whether the lack of an arrestin-mediated internalization is restricted to $Arg^{3.50}$ mutants of GPR33, the ADP receptor $P2Y_{12}$ was mutated in its DRY motif to DHY. Similar to the finding with GPR33-DHY, $Arg^{3.50}$His mutation of $P2Y_{12}$ partially reduces receptor cell surface expression but no β-arrestin-mediated internalization was found (Table 1, Supplementary Fig. S1).

To evaluate further the individual influence of amino acids at position 3.50 on basal receptor activity and cell surface expression, $Arg^{3.50}$ was substituted by other amino acids in mGPR33. As shown in Table 1, basic residues (Arg, His, Lys) are well tolerated with respect to plasma membrane expression and basal activity of mouse GPR33. However, all other mutations lead to loss of basal activity (Gln, Trp, Thr) and/or cell surface expression (Ser, Leu, Glu, Gly). The latter phenotype was probably due to intracellular retention because total cellular expression remained unchanged in most mutant GPR33 (Table 1).

## Discussion

Although mammalian genomes have maintained a relatively constant number of GPCR throughout evolution, several of these receptor genes have become inactivated in some species. One such example is the chemoattractant receptor GPR33, which became inactivated in humans, as well as in several great ape and rodent species in fairly recent time [16]. Pseudogenization of a gene can often be detected when hallmark disruptions to the reading frame, such as deletions, insertions, or stop codons, are present. However, about 65% of all inactivating mutations found naturally in GPCR are missense mutations, which can be identified only by functional testing [11]. More than 80% of such missense mutations hit highly conserved residues. Therefore, it is likely that GPCR pseudogenization, especially if very recent, can escape detection by undirected sequence analysis.

By examining the DRY motif of over 100 mammalian GPR33 orthologs we identified different sequence variations which could alter receptor function. $Asp^{3.49}$ in the DRY motif was found to be substituted by Ala (*Mus*) and Gly (Gerbillinae), both of which mutants were compatible with functional receptor trafficking and signal transduction. Moreover, both mutations increased basal receptor activity. In the crystal structure of rhodopsin [3], the acidic residue at position 3.49 forms a salt bridge with $Arg^{3.50}$ of the DRY motif. Numerous reports show that mutation of the acidic residue in the DRY motif disrupts this ionic interaction and results in increased basal activity of many receptors [28]. But there are several functional rhodopsin-

Table 1
Functional characterization of $Arg^{3.50}$ mutations introduced into mouse GPR33

| Mutation | Basal IP formation (fold over basal) | Cell surface expression (% of ARY) | Total cellular expression (% of ARY) |
|---|---|---|---|
| GFP | 1 | 0[a] | 0[b] |
| ARY (wild type) | 3.2 ± 0.2 | 100[c] | 100[d] |
| AHY | 3.2 ± 0.1 | 138 ± 21 | 104 ± 6 |
| AHY (*M. m. castaneus.*) | 3.2 ± 0.2 | 87 ± 18 | 116 ± 13 |
| DHY | 1.0 ± 0.1 | 69 ± 10 * | 106 ± 7 |
| DRY | 1.1 ± 0.1 | 150 ± 32 | 93 ± 5 |
| DGY | 1.1 ± 0.1 | 38 ± 16 * | 103 ± 30 |
| $P2Y_{12}$ (DRY) | 5.7 ± 0.1 | 291 ± 46 | 97 ± 13 |
| $P2Y_{12}$-DHY | 2.2 ± 0.1 | 219 ± 38 | 110 ± 4 |
| AKY | 3.2 ± 0.2 | 102 ± 22 | 110 ± 8 |
| AQY | 1.4 ± 0.3 | 165 ± 38 | 112 ± 7 |
| AWY | 1.1 ± 0.1 | 120 ± 35 | 129 ± 19 |
| ATY | 1.0 ± 0.1 | 99 ± 38 | 120 ± 4 |
| ASY | 1.0 ± 0.1 | 33 ± 20 * | 125 ± 16 |
| AAY | 1.0 ± 0.2 | 29 ± 13 * | 114 ± 17 |
| ALY | 1.1 ± 0.1 | 14 ± 10 * | 123 ± 12 |
| AEY | 1.0 ± 0.2 | 22 ± 10 * | 104 ± 12 |
| AGY | 1.1 ± 0.1 | 25 ± 15 * | 49 ± 24 |

Wild-type and the mutant mouse GPR33 and human $P2Y_{12}$ constructs were coexpressed with the chimeric $G\alpha_{\Delta 6qi4myr}$ protein in COS-7 cells and IP assays were performed as described (Materials and methods). As a negative control a plasmid encoding GFP was cotransfected with $G\alpha_{\Delta 6qi4myr}$. Basal IP formation is presented as fold over GFP-transfected cells (544 ± 15 cpm/well). Data are presented as means ± SD of two to four independent experiments, each carried out in triplicate. Cell surface and total cellular level expression of the wild-type and mutant receptors were measured by a cell surface ELISA and sandwich ELISA. Specific optical density (OD) readings (OD value of HA-tagged construct minus OD value of GFP-transfected cells) are given as a percentage of wild-type HA-tagged GPR33. The nonspecific OD value (GFP) was 0.089 ± 0.031 ([a]set as 0%) and 0.175 ± 0.033 ([b]set as 0%), and the OD value of the wild-type HA-tagged GPR33 was 0.485 ± 0.116 ([c]set as 100%) and 1.144 ± 0.116 ([d]set as 100%). ELISA data are given as means ± SD of four to seven (cell surface ELISA) and two or three (sandwich ELISA) independent experiments, each carried out in quadruplicate.

* Significant reduction in cell surface expression ($P < 0.005$).

like GPCR in which the acidic residue within this motif is naturally substituted by Ala (relaxin receptors) or Gly (free fatty acid receptor 1, GPR40). Therefore, our functional analysis of GPR33 orthologs suggests that the $Asp^{3.49}$ substitutions identified in mouse and gerbil species are compatible with proper receptor function. It should be noted that GPR33 is pseudogenized in all gerbil species because of truncating or frame-disrupting mutations that obviously occurred after speciation. Since $Asp^{3.49}$Gly is found in all gerbil species, this substitution must have occurred before receptor inactivation.

In contrast to position 3.49, naturally occurring substitutions of $Arg^{3.50}$ in GPCR without obvious functional changes are very rare but include LGR8, CCBP2, and GPR1. Most mutations of $Arg^{3.50}$ in rhodopsin-like GPCR, including mutations to His, lead to receptor inactivation (see references above). Although there is only one experimentally proven example of GPCR pseudogenization by a missense mutation [12], we speculated that the $Arg^{3.50}$His mutation found to be polymorphic in the Asian house mouse may inactivate GPR33. However, all tests (cell surface expression, IP assay, β-arrestin recruitment) revealed equivalent function of $Arg^{3.50}$His compared with wild-type GPR33. In respect to cell surface expression, minor context-specific differences were observed (AHY versus DHY, see Table 1). Systematic substitution of $Arg^{3.50}$ in mGPR33 showed that not only His but all basic amino acid residues are functionally tolerated in this position. We cannot exclude more distinct functional differences between the $Arg^{3.50}$ and the $His^{3.50}$ variants, as shown for the agonist-induced internalization kinetics of the V1a vasopressin receptor [29]. However, the allele frequency, the presence of several allelic variants at other sites in complete linkage disequilibrium with $His^{3.50}$, and the allele age estimates all suggest a long natural coexistence of both alleles. Further, initial characterization of mouse strains homozygous for either $Arg^{3.50}$ or $His^{3.50}$ revealed no obvious differences in mating, behavior, blood cell counts, and morphological characteristics (Römpler and Schöneberg, unpublished data).

Until the agonist for GPR33 is identified the functional relevance of Gln, Trp, and Thr substitution at position 3.50 will remain an open question because receptors bearing these mutations are still expressed at the cell surface and only a loss of basal receptor activity is observed. The loss of basal activity does not automatically implicate a loss of agonist-induced receptor function. However, significant alteration of receptor function (loss of cell surface expression and basal activity) was observed by mutating $Arg^{3.50}$ to Ser, Leu, Glu, and Gly. The majority (more than 80%) of disease-causing inactivating missense mutations found in GPCR lead to improperly folded proteins that are retained intracellularly by the endoplasmic reticulum quality control system [30] and, therefore, escape agonist stimulation and signal transduction through the plasma membrane. It is therefore very likely that the $Arg^{3.50}$Gly mutation found in the Syrian hamster GPR33 ortholog inactivates the receptor. However, the hamster ortholog displays a frameshifting mutation; thus it remains unclear whether this deletion or the $Arg^{3.50}$Gly mutation was responsible for pseudogenization.

## Conclusion

Our analysis of more than 100 mammalian orthologs of the chemoattractant receptor GPR33 provided an excellent evolutionary history of a GPCR that underwent a variety of functional fates. First, focusing on the highly conserved DRY motif, we found that mutations increased the basal activity of this receptor in *Mus* and Gerbillinae species before the receptor became pseudogenized in the Gerbillinae subfamily. Second, because of the DRY motif diversity observed in vivo, the functional relevance of residues within this motif appears to be receptor and context-specific. In contrast to many other GPCR, the conserved $Arg^{3.50}$ of mGPR33 can be substituted by other basic amino acids without obvious loss of receptor function as shown in in vitro tests and as implicated by the long natural coexistence of $Arg^{3.50}$ and $His^{3.50}$ variants. On the other hand there is evidence that other $Arg^{3.50}$ mutations in GPR33 can be inactivating, as illustrated by $Arg^{3.50}$Gly in hamster. The latter

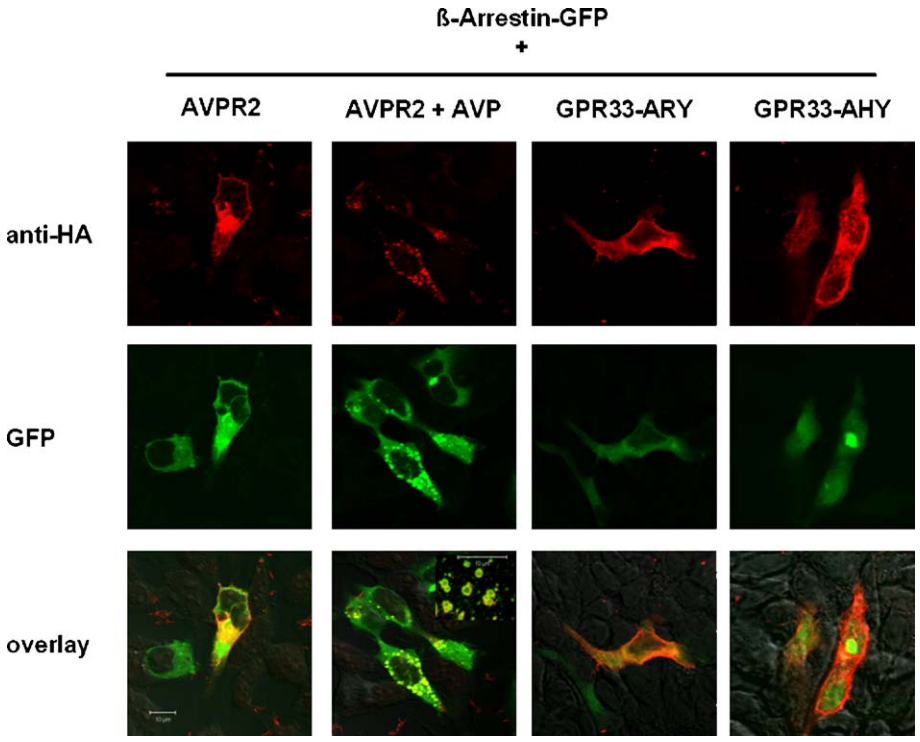


Fig. 4. Lack of β-arrestin-mediated desensitization of the AHY mGPR33 variant. HEK cells were cotransfected with β-arrestin–GFP and HA-tagged versions of the V2 vasopressin receptor (AVPR2), wild-type mGPR33 (GPR33-ARY), or the mGPR33-Arg$^{3.50}$His (GPR33-AHY) variant. Cells were fixed and permeabilized and HA-tagged GPCR were detected with an anti-HA monoclonal antibody and a TRITC-labeled anti-mouse antibody. For control purposes, AVPR2-transfected cells were incubated 10 nM AVP for 20 min at 37°C prior to fixation. Specific fluorescence of HA-tagged GPCR (red) and GFP (green) and the overlay of both fluorescences plus the differential interference contrast image are shown. The inset (overlay AVPR2 + AVP) depicts internalization vesicles at a higher magnification. Scale bars, 10 μm. The pictures shown are representative of at least four additional independent experiments. The cotransfection efficiency (obvious receptor and arrestin fluorescence related to transfected cells) was 51 ± 8% (20 randomly chosen transfected cells were counted per independent experiment). In the case of AVPR2 90 ± 6% of all cotransfected cells displayed double-fluorescent internalization vacuoles upon AVP stimulation.

finding further supports the idea that recent pseudogenization events due to inactivating missense mutations in GPCR, and likely in other proteins as well, can escape simple bioinformatics inspection of genes.

## Materials and methods

### *GPR33 ortholog identification and site-directed mutagenesis*

To analyze the sequence of GPR33 orthologs, genomic DNA samples were prepared from tissue or peripheral mononuclear blood cells of various mammalian species (sources are given in Supplemental Table S2). Tissue samples were digested in lysis buffer (50 mM Tris/HCl, pH 7.5, 100 mM EDTA, 100 mM NaCl, 1% SDS, 0.5 mg/ml proteinase K) and incubated at 55°C for 18 h. DNA was purified by phenol/chloroform extraction and ethanol precipitation. Degenerate primer pairs (Supplemental Table S3) were applied to amplify GPR33-specific sequences. PCRs were performed with *Taq* polymerase under variable annealing and elongation conditions. Conditions of a standard PCR were as follows: Genomic DNA (100 ng) was used in PCRs (50 μl) with primers (10 pmol each), standard buffer (Perkin–Elmer), dNTP (200 μM), and *Taq* polymerase (1 U; Perkin–Elmer). The reactions were initiated with a denaturation at 94°C for 3 min, followed by 35 cycles of denaturation at 94°C for 45 s, annealing at 60°C for 45 s, and elongation at 72°C for 2 min. A final extension step was performed at 72°C for 10 min. Specific PCR products were sequenced directly and/or subcloned into the pCR2.1-TOPO vector (Invitrogen, La Jolla, CA, USA) for sequencing. In the case of heterozygosity allelic separation was performed by subcloning and subsequent sequencing. Sequencing reactions were performed on PCR products with a dye-terminator cycle sequencing kit (Applied Biosystems) on an ABI 3700 automated sequencer (Applied Biosystems). Based on considerable sequence similarities of the 5′ and 3′ untranslated regions of GPR33 genes, primers were designed (Supplemental Table S3) that allowed for the identification of sequences encoding the N and C termini of mammalian GPR33 orthologs.

To determine the frequency of allelic variants in *M. m. castaneus* we sequenced the GPR33 coding region of 88 wild-caught animals from Taiwan [31]. Here, PCR fragments were amplified with genomic DNA samples and primers (Supplemental Table S3). PCR products were separated in a 1% agarose gel, purified by using a gel extraction kit (Qiagen), and sequenced.

The full-length mGPR33 and the human $P2Y_{12}$ were inserted into the mammalian expression vector pcDps and epitope-tagged with an N-terminal HA epitope and a C-terminal FLAG epitope by a PCR-based overlapping fragment mutagenesis approach [16,21]. Mutations were introduced into the double-tagged versions of the murine GPR33 and human $P2Y_{12}$ using a PCR-based site-directed mutagenesis and restriction fragment replacement strategy. The identity of the various constructs and the correctness of all PCR-derived sequences were confirmed by restriction analysis and sequencing.

### *Cell culture and functional assays*

COS-7 cells and HEK 293 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum, 100 U/ml penicillin, and 100 μg/ml streptomycin at 37°C in a humidified 7% $CO_2$ incubator. LipofectAMINE (Invitrogen) was used for transient transfection of COS-7 cells. Subsequently cells were split into 12-well plates ($1.5 \times 10^5$ cells/well) and transfected with a total amount of 1 μg of plasmid DNA/well. To measure IP formation, transfected COS-7 cells were incubated with 2 μCi/ml myo-[3H]inositol (18.6 Ci/mmol; Perkin-Elmer) for 18 h. Thereafter, cells were washed once with serum-free DMEM containing 10 mM LiCl followed by

incubation for 1 h at 37°C. Intracellular IP levels were determined by anion-exchange chromatography as described [32].

### *ELISA and immunofluorescence studies*

To estimate cell surface expression of receptors carrying an N-terminal HA tag, we used an indirect cellular ELISA [33]. To assess further the amounts of full-length HA/FLAG-double tagged GPR33 and $P2Y_{12}$ constructs and to demonstrate that the reduction of cell surface expression levels is not due to a decrease in receptor expression in general, a previously developed "sandwich ELISA" was used [17]. In brief, transfected cells were harvested from 6-cm dishes, and membrane preparations were solubilized in lysis buffer (10 mM Tris–HCl, pH 7.4, 150 mM NaCl, 1 mM DTT, 1 mM EDTA, 1% desoxycholate, 1% Nonidet NP-40, 0.2 mM PMSF, 10 μg/ml aprotinin) overnight. Microtiter plates (Maxi Sorp, Nunc Immuno plates; Nunc) were coated with a monoclonal antibody directed against the C-terminal FLAG tag (10 μg/ml in 0.05 M borate buffer, M2 antibody; Sigma). After incubation with the solubilized membranes, bound full-length receptor proteins were detected with the combination of a biotin-labeled anti-HA monoclonal antibody (12CA5; Roche Molecular Biochemicals) and a peroxidase-labeled streptavidin conjugate (Sigma) and a color reaction [17].

Immunofluorescence studies were carried out to examine the subcellular distribution of the various receptor constructs and bovine β-arrestin-2-GFP (generous gift of Professor M. Lohse, Würzburg). HEK cells were transferred into six-well plates containing sterilized glass coverslips and transfected. For immunofluorescence staining cells were fixed 48 h after transfection, permeabilized with 0.1% Triton X-100 in PBS (PBS-T), and probed with a monoclonal anti-HA antibody (Roche; 10 μg 12CA5 in PBS-T). The primary mouse antibody was detected using an anti-mouse-IgG TRITC-labeled secondary antibody (Sigma). Fluorescence images were obtained with a confocal laser-scanning microscope (LSM 510; Carl Zeiss Jena, Jena, Germany).

## Acknowledgments

We thank the numerous contributors of species samples (Supplemental Table S2) and Thomas Hermsdorf for providing the double-tagged human $P2Y_{12}$. We are grateful to Hilliary Creely, Manyuan Long, and the anonymous reviewers for many suggestions and critical reading of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft, IZKF, Leipzig and Bundesministerium für Bildung und Forschung.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ygeno.2006.02.009.

## References

[1] J.A. Ballesteros, H. Weinstein, Integrated methods for the construction of three dimensional models and computational probing of structure–function relations in G-protein coupled receptors, Methods Neurosci. 25 (1995) 366–428.
[2] J. Wess, G-protein-coupled receptors: molecular mechanisms involved in receptor activation and selectivity of G-protein recognition, FASEB J. 11 (1997) 346–354.
[3] K. Palczewski, T. Kumasaka, T. Hori, et al., Crystal structure of rhodopsin: a G protein-coupled receptor, Science 289 (2000) 739–745.
[4] J. Li, P.C. Edwards, M. Burghammer, C. Villa, G.F. Schertler, Structure of bovine rhodopsin in a trigonal crystal form, J. Mol. Biol. 343 (2004) 1409–1438.
[5] A.E. Alewijnse, H. Timmerman, E.H. Jacobs, M.J. Smit, E. Roovers, S. Cotecchia, R. Leurs, The effect of mutations in the DRY motif on the constitutive activity and structural instability of the histamine H2 receptor, Mol. Pharmacol. 57 (2000) 890–898.
[6] S.G. Rasmussen, A.D. Jensen, G. Liapakis, P. Ghanouni, J.A. Javitch, U. Gether, Mutation of a highly conserved aspartic acid in the β2 adrenergic receptor: constitutive activation, structural instability, and conformational rearrangement of transmembrane segment 6, Mol. Pharmacol. 56 (1999) 175–184.
[7] R.R. Franke, T.P. Sakmar, R.M. Graham, H.G. Khorana, Structure and function in rhodopsin: studies of the interaction between the rhodopsin cytoplasmic domain and transducin, J. Biol. Chem. 267 (1992) 14767–14774.
[8] P.G. Jones, C.A.M. Curtis, E.C. Hulme, The function of a highly-conserved arginine residue in activation of the muscarinic M1 receptor, Eur. J. Pharmacol. 288 (1995) 251–257.
[9] A. Scheer, F. Fanelli, T. Costa, P.G. De Benedetti, S. Cotecchia, Constitutively active mutants of the a1B-adrenergic receptor: role of highly conserved polar amino acids in receptor activation, EMBO J. 15 (1996) 3566–3578.
[10] S.Z. Zhu, S.Z. Wang, J. Hu, E.E. el-Fakahany, An arginine residue conserved in most G protein-coupled receptors is essential for the function of the m1 muscarinic receptor, Mol. Pharmacol. 45 (1994) 517–523.
[11] T. Schöneberg, A. Schulz, H. Biebermann, T. Hermsdorf, H. Römpler, K. Sangkuhl, Mutant G-protein-coupled receptors as a cause of human diseases, Pharmacol. Ther. 104 (2004) 173–206.
[12] I. Gaillard, S. Rouquier, A. Chavanieu, P. Mollard, D. Giorgi, Amino-acid changes acquired during evolution by olfactory receptor 912-93 modify the specificity of odorant recognition, Hum. Mol. Genet. 13 (2004) 771–780.
[13] I. Menashe, O. Man, D. Lancet, Y. Gilad, Population differences in haplotype structure within a human olfactory receptor gene cluster, Hum. Mol. Genet. 11 (2002) 1381–1390.
[14] Z. Zhang, N. Carriero, M. Gerstein, Comparative analysis of processed pseudogenes in the mouse and human genomes, Trends Genet. 20 (2004) 62–67.
[15] Z. Zhang, M. Gerstein, Large-scale analysis of pseudogenes in the human genome, Curr. Opin. Genet. Dev. 14 (2004) 328–335.
[16] H. Römpler, A. Schulz, C. Pitra, G. Coop, M. Przeworski, S. Pääbo, T. Schöneberg, The rise and fall of the chemoattractant receptor GPR33, J. Biol. Chem. 280 (2005) 31068–31075.
[17] T. Schöneberg, A. Schulz, H. Biebermann, A. Grüters, T. Grimm, K. Hübschmann, G. Filler, T. Gudermann, G. Schultz, V2 vasopressin receptor dysfunction in nephrogenic diabetes insipidus caused by different molecular mechanisms, Hum. Mutat. 12 (1998) 196–205.
[18] H.B. Schiöth, S.R. Phillips, R. Rudzish, M.A. Birch-Machin, J.E. Wikberg, J.L. Rees, Loss of function mutations of the human melanocortin 1 receptor are common and are associated with red hair, Biochem. Biophys. Res. Commun. 260 (1999) 488–491.
[19] A. Chen, Z.G. Gao, D. Barak, B.T. Liang, K.A. Jacobson, Constitutive activation of A(3) adenosine receptors by site-directed mutagenesis, Biochem. Biophys. Res. Commun. 284 (2001) 596–601.
[20] A.M. Wilbanks, S.A. Laporte, L.M. Bohn, L.S. Barak, M.G. Caron, Apparent loss-of-function mutant GPCR revealed as constitutively desensitized receptors, Biochemistry 41 (2002) 11981–11989.
[21] A. Schulz, T. Schöneberg, The structural evolution of a P2Y-like G-protein-coupled receptor, J. Biol. Chem. 278 (2003) 35531–35541.
[22] E. Kostenis, Is Galpha16 the optimal tool for fishing ligands of orphan G-protein-coupled receptors? Trends Pharmacol. Sci. 22 (2001) 560–564.
[23] A. Scheer, F. Fanelli, T. Costa, P.G. De Benedetti, S. Cotecchia, The activation process of the alpha1B-adrenergic receptor: potential role of protonation and hydrophobicity of a highly conserved aspartate, Proc. Natl. Acad. Sci. USA 94 (1997) 808–813.
[24] L.L. Jacobs, D. Pilbeam, Of mice and men: fossil-based divergence dates and molecular 'Clocks', J. Hum. Evol. 9 (1980) 551–555.
[25] M.S. Springer, W.J. Murphy, E. Eizirik, S.J. O'Brien, Placental mammal diversification and the Cretaceous–Tertiary boundary, Proc. Natl. Acad. Sci. USA 100 (2003) 1056–1061.
[26] J.Y. Dubois, P.A. Jekel, P.P. Mulder, A.P. Bussink, F.M. Catzeflis, A.

Carsana, J.J. Beintema, Pancreatic-type ribonuclease 1 gene duplications in rat species, J. Mol. Evol. 55 (2002) 522–533.

[27] L.S. Barak, R.H. Oakley, S.A. Laporte, M.G. Caron, Constitutive arrestin-mediated desensitization of a human vasopressin receptor mutant associated with nephrogenic diabetes insipidus, Proc. Natl. Acad. Sci. USA 98 (2001) 93–98.

[28] T. Schöneberg, A. Schulz, T. Gudermann, The structural basis of G-protein-coupled receptor function and dysfunction in human diseases, Rev. Physiol. Biochem. Pharmacol. 144 (2002) 143–227.

[29] S.R. Hawtin, Charged residues of the conserved DRY triplet of the vasopressin V1a receptor provide molecular determinants for cell surface delivery and internalization, Mol. Pharmacol. 68 (2005) 1172–1182.

[30] R. Schülein, The early stages of the intracellular transport of membrane proteins: clinical and pharmacological implications, Rev. Physiol. Biochem. Pharmacol. 151 (2004) 45–91.

[31] S.W. Huang, K.G. Ardlie, H.T. Yu, Frequency and distribution of t-haplotypes in the Southeast Asian house mouse (*Mus musculus castaneus*) in Taiwan, Mol. Ecol. 10 (2001) 2349–2354.

[32] M.J. Berridge, Rapid accumulation of inositol trisphosphate reveals that agonists hydrolyse polyphosphoinositides instead of phosphatidylinositol, Biochem. J. 212 (1983) 849–858.

[33] T. Schöneberg, V. Sandig, J. Wess, T. Gudermann, G. Schultz, Reconstitution of mutant V2 vasopressin receptors by adenovirus-mediated gene transfer: molecular basis and clinical implication, J. Clin. Invest. 100 (1997) 1547–1556.