

Chapter 2

Examining the Dynamic Evolution of G Protein-Coupled Receptors

Claudia Stäubert, Diana Le Duc, and Torsten Schöneberg

Abstract

The valuable source of large-scale genomic information initiated attempts to identify the origin(s) of G protein-coupled receptors (GPCR), count and categorize those genes, and follow their evolutionary history. Being present in fungi, plants, and unicellular eukaryotes, GPCR must have evolved before the plant-fungi-animal split about 1.5 billion years ago. Phylogenetic analyses revealed several kinds of evolutionary patterns that occurred during GPCR evolution including one-to-one orthologous relationships, species-specific gene expansion, and episodic duplication of the entire GPCR repertoire in certain species lineages. These data document the highly dynamic process of birth and death of GPCR genes since hundreds of millions of years. Genetic drift and selective forces have shaped the individual structure of a given receptor gene but also of the species-specific receptor repertoire – a process that is still ongoing. These processes have left footprints in the genomic sequence that can be detected by bioinformatic methods and may help to interpret receptor function in the light of a given species in its environment. Reasonable intraspecies sequence variability in GPCR is either physiologically tolerated or promotes individual phenotypes and adaptation, but also susceptibilities for diseases. Therefore, the impact of GPCR variants in epistatic networks will be an important task of future GPCR research. The chapter summarizes evolutionary processes working on GPCR genes and sheds light on their consequences at the levels of receptor structure and function.

Key words G protein-coupled receptors, GPCR, Selection, Genetic variability

1 Introduction to the Evolutionary Dynamics of G Protein-Coupled Receptors

Once introduced into the early eukaryotic genome, G protein-coupled receptors (GPCRs) evolved in various structural classes, eventually resulting in hundreds of members in invertebrate and vertebrate genomes. The conserved molecular architecture formed of seven-transmembrane helices and the highly preserved set of intracellular signaling mechanisms were kept over 1 billion years of eukaryotic evolution. However, manifold combinations of amino acids within the transmembrane core and the loop regions produced a versatile binding pocket for almost every natural compound that may serve as signal. This balanced composition of conserved and variable structures is the key to the evolutionary

success of GPCRs. The repertoires of GPCRs retrieved from extant genomes are the recent endpoints of parallel evolutionary processes. Since genomes have been shaped by genetic drift and selective forces over millions of years, understanding structure-function relationships and the physiological relevance of individual GPCRs only makes sense in the light of evolution. With the availability of numerous invertebrate and vertebrate genome sequences in public databases, there has been much effort to categorize GPCR genes and to follow the evolutionary history of these receptor genes.

The dynamic evolution of GPCR can be analyzed from different perspectives: either by sequence comparisons between species or between populations within one species. The interspecies perspective provides phylogenetic information about gain and loss and thereby the relation of individual receptor genes, receptor families, and classes. Further, sequence comparison between orthologs of different species reveals conserved and variable regions within a receptor molecule. Interspecies as well as population genetic data are used to scan for signatures of selection in GPCR gene-containing loci. Since GPCRs are very central in the regulation of almost every physiological process, it is not surprising that several receptor variants appear to be involved in adaptation to environmental changes and niches. This chapter sheds light on the origin(s), birth and death of GPCR genes, and their functions and focuses on elucidating selective mechanisms (still) driving these processes.

1.1 The Origin(s) of Recent G Protein- Coupled Receptors

Most mammalian GPCR can be classified into the Glutamate, Rhodopsin, Adhesion, Frizzled, and Secretin classes (GRAFS system) [1]. Because those analyses are only based on sequence data from recent species, the relations of these classes to each other and their origin(s) are still under investigation. The central requirements for constructing phylogenetic trees and timing of origin/branch points are (1) a correct phylogenetic tree of species, (2) sequence differences between receptors, and (3) fossil or other datable records. Fossil records older than 1 million years usually do not contain authentic DNA of sufficient quality [2]. For those species that are assumed to be very basic from evolutionary perspective (e.g., unicellular eukaryotic organisms), fossils are generally absent. Further, many other factors can influence the reconstruction of the phylogeny of GPCR (and other) genes and the timing of branch points. Gene loss without sequence trace in the genome, population bottlenecks, the effective population size, and the generation time of extinct species as well as horizontal (lateral) gene transfer are mostly unknown variables in sequence-based phylogenetic analyses. Since most phylogenetic reconstructions are based on sequence alignments, methodical issues (e.g., alignment problems of distantly related genes, sequence gaps) can significantly influence the outcome. Nevertheless, all studies start with

mining sequence data which at least provides more or less complete GPCR gene repertoires in the genomes of given species.

Signaling through seven-transmembrane (7TM) receptors is a feature of eukaryotic organisms, but proteins with a 7TM topology are already encoded in prokaryote genomes. This includes light-sensitive proteorhodopsins, bacteriorhodopsins, and halorhodopsins that are involved in non-photosynthetic energy harvesting in archaea and bacteria [3, 4]. Although structurally similar sensory rhodopsins are also found in eukaryotes [5, 6], their phylogenetic relation to GPCRs remains unsolved [7]. As the divergence between prokaryotes and eukaryotes has taken place roughly 1.6–2.1 billion years ago, the sequence similarity between the 7TM domains is low and a common origin cannot be derived from sequence data. The answer to how the 7TM core of eukaryotic GPCRs has evolved still needs to be found. Structural and functional data clearly show that G protein signaling via GPCR is present in yeast/fungi [8], plants [9], and unicellular eukaryotes [10]. Therefore, this receptor-signaling complex must have evolved before the plant-fungi-animal split about 1.5 billion years ago. The most prominent and eponymous feature of GPCRs is their signal transduction through G proteins. However, one has to consider that GPCRs signal not only via G proteins but also via alternative, non-G-protein-linked signaling pathways [11]. Therefore, it remains open whether G proteins were involved in GPCR signaling from the very evolutionary beginning or if the prototypes of what we now call GPCR initially fulfilled other functions [12].

Phylogenetic studies suggest that cAMP receptors are the most ancient GPCR, being present already in *Alveolate* genomes such as ciliated protozoa and dinoflagellates [13]. A recent study provides evidence that, besides fungi-specific GPCRs, some fungi genomes contain metazoan-like GPCRs revealing early evolutionary history of several GPCR classes [14]. Thus, based on the existence in recent eukaryotic genomes, the origin of the rhodopsin-like class can be traced back to the ancestor of Opisthokonts (~1.1 billion years), the Adhesion and the Frizzled classes to Unikonts (~1.3 billion years), and the Glutamate and the cAMP receptor classes to the common ancestor of Alveolates and Unikonts, early in eukaryotic evolution more than 1.4 billion years ago [14]. Although sequence divergence between fungal and metazoan rhodopsin-like GPCRs suggests long separate evolution, horizontal gene transfer in early metazoan evolution cannot be completely ruled out. Indeed, the above study provides evidence that adhesion GPCRs in *Alveolata* may have arisen from horizontal gene transfer since the *Alveolata* and *Fungi* are distantly related, but adhesion GPCRs from both superphyla show pairwise identity of about 50 % in the 7TM region. If we consider that horizontal gene transfer had only minor impact in GPCR history, most recent GPCR classes (rhodopsin, adhesion, frizzled, glutamate) were most likely evolved

from ancestral cAMP receptors [14]. It is also likely that the Secretin class of GPCRs emerged secondarily from the Adhesion class of GPCRs [15].

1.2 Expansion of the G Protein-Coupled Receptors Repertoire

Phylogenetic analyses revealed that there are several kinds of evolutionary patterns that occurred during GPCR evolution: (1) one-to-one orthologous relationships, (2) species-specific expansion, and (3) episodic duplication of the entire GPCR repertoire in certain species lineages. The one-to-one relationship of receptor orthologs is mostly seen in closely related species though some GPCRs are highly stable in their number and presence throughout animal genomes. For example, 2–3 subtypes of latrophilins, which are neuronal adhesion G protein-coupled receptors and the major receptors for α -latrotoxin (black widow spider toxin), are found in almost all invertebrate and vertebrate genomes sequenced so far. The qualitative and quantitative constancy of latrophilins appears to be essential for the animal blueprint since they play a fundamental role in anterior-posterior tissue polarity and development [16]. Similarly, the number of glutamate receptors remained relatively constant during animal evolution [14].

Comparisons of vertebrate and invertebrate genomes indicate that invertebrates contain the basic ancestral complement of vertebrate GPCR genes [17–19]. However, the number of most GPCRs in non-vertebrate genomes (exceptions are the chemokine receptors in worms) is substantially lower than in vertebrate genomes [1]. This is especially true for rhodopsin-like GPCRs that evolved and expanded in the very early vertebrate evolution about 500 million years (Myr) ago during a time called the “Cambrian Explosion.” For example, purinergic P2Y-like GPCRs constitute a rhodopsin-like family with several dozens of members in vertebrates that is virtually absent in non-vertebrates [20].

Species-specific expansion of GPCR genes is attributed to adaptive processes of species lineages. GPCR genes of the sensory system are the prototypical example for this very dynamic process. Expansion of GPCR odorant receptors (OR) due to gene or even locus duplications may have expanded the repertoire of odorant signals that species detect, allowing them to occupy new ecological niches. The number of rhodopsin-like OR varies greatly among vertebrates, ranging from approximately 22 genes in the pufferfish, *Tetraodon nigroviridis* [18], to 1,234 OR genes in rats [21]. In general, terrestrial vertebrates have a higher number of intact OR genes compared to aquatic vertebrates. Specifically, mammals are equipped with thousands of OR genes. This OR gene expansion occurred mainly during early mammalian evolution between the bird-mammal split and marsupial-placental split, coinciding with the reduction of opsin genes in primitive mammals. It was suggested that the nocturnal adaptation of early mammals might have triggered this OR gene expansion [22]. However, genomes of human

and great apes contain almost 1,000 OR genes, but only one-third appear to be functional. The “vision-priority” hypothesis implies that the independent acquisition of trichromatic color vision in primates caused the parallel pseudogenization of OR genes through relaxed purifying selection [23]. Although the loss of function of OR genes within primates has been mainly attributed to a “trade-off” between vision and olfaction, functional OR gene repertoires were reduced independently in aquatic mammals with their multiple origins of vision, some of those being L cone visual monochromats [24]. Therefore, vision and olfaction do apparently not always show an evolutionary link, and thus other ecological factors may account for OR expansion.

An obvious expansion of OR occurred also during bird evolution. Compared to the repertoire of OR genes of the green anole lizard (approximately 156, including 42 pseudogenes), the OR gene repertoire of two bird species is substantially larger with 479 and 553 OR gene homologs in chicken and the zebra finch, respectively (including 111 and 221 pseudogenes). Analysis of selective pressures on the paralogous genes revealed that they have been subjected to adaptive evolution [25]. Note, the green anole has a higher fraction of intact OR genes compared to the zebra finch. These data document the highly dynamic process of birth and death of OR since the bird-reptile split occurred over 200 Myr ago [26].

The number of OR genes in insects is relatively low compared to vertebrates. There are approximately 10 ORs in human body louse, 60 in fruit fly, and 165 in honey bees. However, more than 400 putative olfactory receptor genes were identified in fire ants. This represents the largest repertoire of ORs reported so far in any insect [27]. Therefore, the great or varied dynamics of OR evolution found in vertebrate genes also independently occurred in invertebrates.

Expansion of the chemosensory receptor repertoire is not only found in rhodopsin-like OR. In the genome of the frog *Xenopus tropicalis*, more than 330 vomeronasal 2 receptors were identified [28]. Vomeronasal 2 receptors are pheromone receptors that structurally belong to the class of metabotropic glutamate receptors. Together with the 665 rhodopsin-like OR, the olfactory/pheromone system of *Xenopus* may reflect lineage-specific adaptation to both aquatic and terrestrial environments.

Recent studies have suggested that the common ancestor of the extant vertebrates underwent two rounds of whole-genome duplication. Phylogenetic studies of individual gene families indicate that the gene repertoire, which was expanded in the common ancestor of jawed vertebrates, was later reshaped. This already implicates dynamic processes in gain and loss of individual receptor genes in early vertebrate evolution and before. Indeed, genome comparison revealed that secretin-like receptors emerged through local tandem gene duplications before the two rounds of

whole-genome duplication. These genes have then been amplified by whole-genome duplication, followed by additional local duplications and gene losses prior to the divergence of tetrapods and teleost fishes [29].

There is evidence that within the GPCR superfamily the expansion of different classes and families was asymmetric. GPCRs with non-peptide ligands were reported to have a significantly higher retention rate than GPCRs with peptide ligands after lineage-specific whole-genome duplications in the pufferfish *T. nigroviridis* [30]. A more recent study examined if the same is true for GPCRs in general [31]. In contrast to the earlier report [30], it was obvious that more duplication events were found among GPCRs that bind peptide ligands than those binding non-peptide ligands in more recent evolutionary stages (after 400 Myr). This may reflect that the selective pressures driving the fixation of different types of duplicated GPCRs were different in different species and at different evolutionary stages. Indeed, before 600 Myr, more duplication events were observed in GPCRs with non-peptide ligands than GPCR with peptide ligands [31].

1.3 G Protein-Coupled Receptor Pseudogenes

Pseudogenes are inheritable and characterized by homology to a known gene but lack of functionality [32]. Depending on the mechanism by which they evolved, the majority of mammalian pseudogenes can be classified as duplicated pseudogenes or retrotransposed pseudogenes (also called processed pseudogenes). The latter are generated by reverse transcription of mRNAs, followed by genomic integration. The human GnRH type II receptor homolog is one well-characterized example of GPCR pseudogenization caused by retrotransposition [33]. Duplicated pseudogenes arise from local duplication or unequal crossing-over. This is frequently seen in genomic clusters of OR and trace-amine-associated receptors, called TAARs [34]. Thus, GPCR pseudogenes often retain the original exon-intron structures of their parental genes.

When signals from GPCRs do not provide a selective advantage, the repertoire of receptors is reduced by gene inactivation as shown for OR (see Sect. 1.2 above) and taste receptors [35]. This process is called pseudogenization and may occur from duplication or retrotransposition (see above). As a result of their nonfunctionality, most pseudogenes are released from selective pressure. Therefore, compared to functional genes, pseudogenes, if old enough, display a ratio of non-synonymous to synonymous substitution rates (K_a/K_s) of 1 [36] and accumulate missense, frame shifting, and nonsense mutations. Once released from purifying selection, it will take several Myr for obvious signatures of inactivation (premature stop codon, frame-shifting mutations) to become fixated in the coding sequence. In primate TAAR, for example, rough estimates suggest that 7.3 ± 2.7 Myr are required to obtain and fixate at least one such obvious signatures of receptor

inactivation [37]. On the other hand, signatures of the original sequence will gradually disappear over time. Consequently, a pseudogene may escape present-day detection depending on the date and mechanism of its pseudogenization. Rough estimates suggest that signatures of genes can be detected from more than 80 Myr of neutral evolution. For example, the neuropeptide Y receptor type 6 (Y6R) is a pseudogene in all primate genomes investigated so far, originated from an inactivating deletion [38] which occurred in the common ancestor of primates approximately 80 Myr ago [39]. However, large deletions can remove informative sequences in a shorter time period as found in rodent Y6R evolution where it disappeared in rats after the mouse/rat split 14–16 Myr ago.

One should keep in mind that missense mutations impairing protein function may render a gene functionally inactive although there is none of the critical pseudogene features present. This has been described for a number of GPCRs including chemoattractant and olfactory receptors [40–42]. Unless tested in functional assays, these pseudogenes thus escape detection. Because of this and the many other difficulties in identifying pseudogenes [43, 44], the number of pseudogenes is just an estimate.

Because pseudogenes are considered nonfunctional, they have long been neglected and are considered only as genomic fossils. Recent studies have established that the DNA segment of a pseudogene, the RNA transcribed from a pseudogene, or the protein translated from a pseudogene can have multiple functions [44, 45]. Indeed, transcripts of chemokine, serotonin, and OR pseudogenes have been described [40, 46–48], but the functional relevance of those transcripts is not yet established.

2 Techniques to Detect and Evaluate the Variability of G Protein-Coupled Receptor Genes

Natural selection of favorable gene variants has left footprints in the genomic sequence that can be detected by suitable bioinformatic methods. The availability of large data sets of single nucleotide polymorphisms (SNP) in humans [49] and many other species [50, 51] opened a new dimension for the search of signatures of recent selection [52]. It was now possible to investigate mutations causing variation in complex traits by correlation of genotype and phenotype in genome-wide association studies, called GWAS [53]. Since 3–4 % of all protein-coding genes encode GPCRs, it is not surprising that GWAS and studies on selected genomic loci in humans and animals revealed a number of GPCR genes to be responsible for previously disregarded phenotype variations and signatures of selection, respectively, at least on a statistical level. The following sections will roughly describe and explain the methods and techniques used to study genetic variability and selection in GPCR genes.

2.1 High-Throughput Sequencing Projects

The availability of high-throughput (HTP) sequencing technologies and continuously decreasing costs of sequencing produced a wealth of genome sequences. Whole-genome shotgun sequencing with Illumina or 454 technologies is currently used, but, after assembly of sequence fragments, proper gene annotation is required [54]. Currently, three major lines of HTP sequence data sets are produced: (1) genome data from different species, (2) genome data from different individuals of one species (population genetic data), and (3) transcriptome data from different species, individuals, and cell/tissue types. To properly assemble *de novo* and annotate whole genomes of yet unsequenced species, all three types of data sets are highly recommended to use. Sequencing several individuals of one species unveils individual genomic variants, providing information about the intraspecies variability, and allows estimates of the effective population size. Also, high coverage of the genome is required for a proper assembly with low number of gaps. Transcriptome data is a great help for genome assembly by joining genomic fragments and contigs and also for gene annotation. Genome-wide annotation usually undergoes two phases: (1) genes are predicted *ab initio* using transcriptome and protein structure data and (2) the actual annotation step, which incorporates all the available evidence in the newly sequenced species and profits from other annotated genomes of closely related species [54]. Once assembled, specific methods can be used to mine genomic sequences for GPCR genes [14]. Then, the extracted GPCR repertoire can be compared with those of other species or searched for intraspecies variability in specific GPCR genes. Several public sources can be mined for GPCR ortholog data of different species, such as the 10 K vertebrate genome project [55, 56]. The 10 K vertebrate genome project will provide a collection of DNA sequences representing the genomes of 10,000 vertebrate species, approximately one for every vertebrate genus (<https://genome10k.soe.ucsc.edu/>). Once completed and released, this data set will be a valuable source to analyze the dynamic evolution of the GPCR repertoires in vertebrates and will help to identify variable and invariable positions in the sequence of distinct GPCR orthologs, families, and classes (see below). Genetic variability within the *Homo sapiens* species is available by the data set cataloged in the 1,000 human genomes project at <http://www.1000genomes.org> [49]. In this project, the genome sequences of a large number of people provide a comprehensive resource on human genetic variation. The platform gives access to variations and their frequency identified in more than 1,000 human genomes of different ethnic background and populations.

It should be mentioned that genomic sequence information is not only available from extant species but increasingly also from extinct species. Nuclear genomes from ancient hominids, such as Neanderthals [57] and Denisovans [58], and from several Pleistocene

species, such as mammoths [59], mastodons [60], and cave bears [61], have been sequenced, annotated, and are available online.

In the future, HTP sequencing technologies will offer even more valuable information relevant for GPCR research. High coverage RNASeq data will provide information about the cellular repertoire of expressed GPCRs, their splice variants, and allele-specific transcription. Going forward, it is thus very likely that bioinformatics will continue to contribute significantly to the understanding of GPCR variation, physiology, and molecular function.

2.2 Genetic Variability of G Protein-Coupled Receptors

Genetic variability (polymorphisms) in GPCR genes was recognized at the time when identification of GPCR genes first began [62, 63]. To date an almost uncountable number of GPCR variants have been identified, and many of the variants have been associated with clinical phenotypes and their therapeutic correlates [64]. GPCR variants were detected even in extinct species [65, 66]. GPCR polymorphisms and rare variants were first collected in databases for individual GPCRs [67, 68] and to a greater degree in the GPCR Natural Variants database [69]. Now, comprehensive data on GPCR variability and SNP frequency within over 1,000 sequenced human individuals is provided by the Blast function of the 1,000 human genomes project. Blast analysis of 1,000 human genome sequences revealed that the coding regions of randomly picked rhodopsin-like GPCR genes contain 15 non-synonymous, 10 synonymous, 0.5 premature stop, and 0.3 frame-shifting variants per 100 codons on average. However, there is a reasonable error rate in next-generation sequencing technologies [70], and a careful reevaluation of occasionally identified variants is required [71]. Nevertheless, so far the amount of intraspecies sequence variability in GPCRs that are physiologically tolerated may be underestimated since genomic samples were taken from juvenile and adult humans. Such genomic data will also be relevant for therapeutic purposes since variants of GPCR can contribute to interindividual differences in disease susceptibility and progression and to the variability in drug responses [72].

2.3 Determining Genomic Signatures of Selection

Because GPCRs are involved in the regulation of numerous physiological functions, GPCR variants may have conferred important selective advantages during periods of human evolution. Indeed, several genomic loci with signatures of recent selection in humans contain GPCR genes [12]. Most methods aiming at detecting signatures of recent selection in the human genome have been designed with the paradigm that adaptations were mainly driven by classical positive selection. Thus, beneficial alleles should go to fixation, strongly reducing diversity and increasing levels of linkage disequilibrium in the surrounding genomic regions (selective sweeps). Large-scale SNP data sets from selected human population samples [73] and other species [50] are used to scan genomes

for signatures of recent positive selection [74, 75]. Commonly analyzed parameters include the reduction in genetic diversity, frequency of derived alleles, allele frequency differences, and fixed alleles between populations and haplotype length, as discussed at length earlier [76]. However, positive selection but also bottleneck, founder, and other demographic events can leave very similar signatures in the genome [40]. The reduction in neutral variation can also be caused due to background selection [77]. One should also note that methods designed to detect signatures of positive selection may be sensitive only for a defined historic period [76]. Further, the paradigm that adaptations were mainly driven by positive selection has recently eroded with the realization that the human genome does not show many sites that are fixed between human populations and that fixed differences are always between populations from different continents. This suggests that strong selective events rarely occurred in response to local adaptation [78]. Therefore, results from studies detecting signatures of selection should always be reevaluated by different methods.

Besides approaches searching for positive selection (see above), there are methods detecting purifying selection. Purifying selection (also called negative selection) eliminates strongly deleterious mutations in functional regions. Signatures of purifying selection can be detected by comparing sequences between populations or between species. For example, the biological relevance of a given GPCR can be estimated from its presence or absence in different species [40, 79]. Thus, for this type of analysis, large ortholog data sets from many species are required. Further, amino acid positions that are essential for receptor function can be determined by comparing ortholog sequences [80, 81]. Hence, the conservation of an amino acid at a specific position within a receptor molecule is an estimate of its functional relevance.

3 Involvement of G Protein-Coupled Receptor Variants in Adaptive Evolution

Numerous studies have shown that GPCRs do contribute not only to regulation of very basic processes in organismal development but also to more distinct functions such as immunity, sensory system, growth, energy and water homeostasis, and stress response. It is therefore reasonable to assume that selected variants of GPCRs contribute to adaptation of species to ecological niches. Several hypothesis-free but also hypothesis-driven studies revealed GPCR variants possibly involved in adaptation processes during evolution.

3.1 Signatures of Selection in G Protein-Coupled Receptors

As mentioned above, positive selection of complete GPCR genes occurred after whole-genome duplications in early vertebrate evolution or gene family expansion by gene or loci duplications. After GPCR duplications, one copy may mutate and acquire unique

functionality without risking the fitness of the organism, which is ensured by the duplicate homolog. Further, gene duplicates often retained overlapping expression patterns and preserve some redundancy. This is consistent with a mechanism that boosts gene robustness or increases gene dosages. On the other hand, if not advantageous, continuous accumulation of mutations (neutral evolution) will eliminate one of the duplicated genes. Therefore, GPCR family expansion and fixation of duplicated genes, as found, for example, in the OR family and opsins [82], are a signature of adaptive selection. In the metabotropic glutamate receptor class, for example, continuous duplication events occurred and structurally shaped the extracellular domain (the Venus flytrap module). Phylogenetic analyses show that functional divergence within this GPCR class also involved positive selection [83].

Signatures of recent positive selection at loci containing GPCR genes have been found for several adhesion GPCRs (BAI3, GPR11, CELSR1), P2Y12-like receptors (GPR34, GPR82), chemokine receptors (CXCR4), and ORs [12]. In all cases it remains unclear whether the signals are caused by some selective pressure or by processes which produce very similar footprints in the genome. However, elucidation of the physiological function of such loci may suggest some involvement in adaptive changes. For example, the two orphan receptors GPR34 and GPR82 are involved in modulation of the immune system [84] and energy metabolism [85], respectively, and thus suitable candidates for adaptive processes. There is also evidence that purifying selection and changes in the sequence of calcium-sensing receptors in tetrapod vertebrates may have contributed to adaptive evolution among some major vertebrate clades, reflecting clade-specific differences in natural history and organismal biology, including skeletal involvement in calcium homeostasis [86].

Reports of more distinct positive selection of GPCR variants are rather rare. Most examples are related to coat color changes. Variation in pigmentation is one of the most conspicuous phenotypic traits in vertebrates, including humans. More than 170 genes have been identified that influence pigmentation – among them GPCRs (<http://www.espcr.org/micemut/>). The cause for pigmentation differences between color morphs has been identified down to the underlying nucleotide changes and many cases are related to adaptation. The melanocortin 1 receptor (MC1R), a rhodopsin-like GPCR, is one of the key regulators in melanogenesis [87]. There are numerous examples suggesting that intraspecies variations in color are associated with MC1R variants. For example, melanic plumage in swans is related to amino acid changes at important functional sites in MC1R that are consistent with increased MC1R activity and melanism. Since the putative melanizing mutations were independently derived in the two melanic swan lineages, this is an example of convergent evolution of MC1R [88].

At least in some galliform birds, there is evidence for sexual selection at the MC1R locus [89]. The melanism and MC1R variants in pocket mice are strongly associated, and the melanism is a derived trait from their habitat on dark-colored lava [90]. Adaptation processes can also involve the perception of colors, and variants of visual opsin sometimes form the molecular basis for these differences. There is strong evidence of positive selection of a duplicated UV-sensitive visual opsin (UVRh2) which coincides with wing pigment evolution in *Heliconius* butterflies [91]. The sense of smell and ORs are also subjected to positive selection in vertebrates including humans [92, 93]. However, this is controversial as some reports suggest that the OR diversity is maintained to some degree by balancing selection [94]. Balancing selection describes the case where the heterozygotes for the alleles under consideration have a higher adaptive value than the homozygote. For GPCRs, balancing selection may lead to enhanced ligand recognition success at the population level.

A variant in the human receptor for bitter taste (TAS2R16) is associated with an increased sensitivity to salicin, arbutin, and different cyanogenic glycosides. It was speculated that the increased sensitivity towards harmful cyanogenic glycosides may have driven the signal of selection at an early stage of human evolution [95]. Signatures of positive selection in humans have also been found for the phenylthiocarbamide-sensitive bitter taste receptor TAS2R38 [96, 97].

GPCR variants can modulate receptor activity in both directions – activation and inactivation. In humans activating and inactivating mutations in GPCR can cause diseases with contrary phenotypes. V2 vasopressin receptor (V2R) is the key component involved in renal water reabsorption. Inactivating mutations cause reduced water reabsorption (congenital nephrogenic diabetes insipidus), whereas activating mutations cause increased water reabsorption with serum hypoosmolality and high urinary sodium levels (nephrogenic syndrome of inappropriate antidiuresis or NSIAD). Several marsupial V2R orthologs show a significant increase in basal receptor activity which would cause NSIAD in humans [80]. Only a few amino acid changes compared to other mammalian V2R orthologs must be responsible for this functional change. The gain of basal V2R function in several marsupials may contribute to the increased urine concentration abilities and, therefore, provide an advantage to maintain water and electrolyte homeostasis under limited water supply conditions.

3.2 Purifying Selection on G Protein-Coupled Receptors

The functionality as well as the sequence of many GPCRs remained very constant during hundreds of millions of years. As for other genes, disadvantageous mutations in GPCRs are removed from a population through a process called purifying or negative selection. For many GPCRs, the transmembrane domains (TMD) are

the most conserved compared to the loop domains and the N and C termini [79, 81, 82, 98]. Although the majority of residues across GPCRs experience purifying selection, many GPCRs show positive selection at both TMD and extracellular residues, albeit with a slight bias towards the extracellular loops [98]. It was suggested that the low evolutionary rate of TMD sequences is caused by their biophysical constraints to form alpha-helices and their high percentage of buried residues. However, it was found that the attenuated rate of TMD evolution may result from the strong biophysical constraints of the plasma membrane itself, or by other functional requirements.

Several key residues in TMD are fully conserved in their respective GPCR classes. Most are conserved proline residues which play a major role in the overall fold of these receptors due to proline-induced helical distortions. Therefore, the proline pattern can be used as an evolutionary marker and helps to trace the molecular evolution of GPCRs [99]. But there are also other key residues, such as the fully conserved tryptophan in TMD4 of rhodopsin-like GPCRs, which are under strong purifying selection. Interestingly, key positions of receptors for small molecule natural ligands were under strong negative selection, whereas GPCRs naturally activated by lipids had weaker purifying selection in general [100]. In spite of the strong evolutionary constraints acting on the TMD of GPCRs, positive selection at residues in TMD can occur. For example, signatures of strong positive selection were found at key position 3.29 (Ballesteros-Weinstein numbering scheme) in the large family of G protein-coupled receptors from Mas-related genes (MRGX receptors) which are selectively expressed in small-diameter sensory neurons of dorsal root ganglia [100–102].

3.3 Adaptive Evolution by G Protein-Coupled Receptor Gene Loss

A gain of gene function may not always provide a selective advantage because the maintenance of every functional gene costs energy. Maintaining required genes and the elimination of unconstrained genes by random inactivation are well balanced at an energetic optimum for a given species in its environment. Therefore, gene inactivation is an integral part and may even provide selective advantage during adaptation. Unfortunately, there are no methods to detect selection on pseudogenes except of those identifying signature of recent selection on genomic loci containing pseudogenes.

Pseudogenization is particularly frequent in OR [103], while signatures from only 30 nonolfactory rhodopsin-like GPCR pseudogenes have been detected so far in the human genome [104]. The human OR repertoire has 851 OR loci, whereby more than 50 % of the loci are annotated as nonfunctional due to frame-disrupting mutations. Furthermore, some apparently intact human OR genes lack motifs that are very highly conserved in their mouse orthologs, suggesting that not all human OR genes with complete open reading frames encode functional OR proteins [94].

A similar accumulation of pseudogenes specific to a GPCR family has been observed in human bitter taste receptors [105], the vomeronasal (pheromone) receptors [106] and trace-amine-associated receptors, known as TAARs [37].

There is evidence that loss of GPCR function may have had a selective advantage and the fixation of the pseudogenes was promoted by some selective pressure. For example, analysis of 120 mammalian GPR33 orthologs from almost all eutherian orders of animals revealed an inactivation of this chemoattractant GPCR only in humans, several great apes, and rats. Estimates of the age of the human and rat pseudogenes suggested inactivation in the past 1 million years. The coincidental inactivation and its fixation in several species of distantly related mammalian orders suggest a selective pressure on this chemoattractant receptor gene [40]. A similar scenario of convergent evolution is found for the trace-amine-associated receptors TAAR3 and TAAR4 [37]. Here, species carrying TAAR3 and TAAR4 pseudogenes significantly overlap. Pseudogenization events occurred on the same ape lineages in both receptors, as in the common ancestor of humans, chimpanzees, and gorillas, and on the lineage leading to the white-handed gibbon. In two marmoset species, the loss of constraint occurred in their common ancestor for TAAR3 and TAAR4, but inactivating mutations occurred and became fixed before or after the lineage split. One may speculate that TAAR3 and TAAR4 have probably similar functions and common constraint-determining factors. For unknown reasons the function and/or constraint-determining factors did not provide an advantage anymore and led to parallel but independent fixation of the inactive TAAR3 and TAAR4 variants in different primate species.

Reduction or loss of GPCR function can also be restricted to distinct populations of a species. Differences in environmental conditions may relax the constraint and promote neutral evolution of a gene formerly under purifying selection. Again the MC1R gene is an illustrating example. In African populations, there is strong purifying selection on the MC1R gene, whereas the great MC1R sequence diversity in European populations suggests a relaxation of constraint [107, 108]. Many of the European MC1R variants show reduction or loss of function causing pale skin color and red hair [109]. It is still a matter of debate whether the loss of MC1R functionality in vertebrates is always due to a loss of constraint as a result of adaptation to habitats in which protection from sunlight is less relevant as in areas outside of the equatorial region or for cave-adapted species [110]. Reduced pigmentation may increase reproductive fitness. The reduced MC1R activity in some Pleistocene species, for example, Neanderthals and mammoths, may have promoted vitamin D synthesis in skin under the extreme climate conditions during Pleistocene ice ages [65, 66]. Further,

reduced pigmentation due to loss of MC1R function can provide an advantage against predation. Illustrating examples are the beach mice in Florida [111, 112] and a lizard species in the Chihuahuan Desert [113]. Both animals have lighter-colored coats than their mainland counterparts, driven by natural selection for camouflage against the pale sand dunes. Using a quantitative trait locus (QTL) approach combined with sequence and functional analyses, two distinct genetic alterations in the coding sequence of MC1R cause reduced pigmentation in Mexican cave fishes [110]. Interestingly, the depigmented phenotype has arisen independently in geographically separate caves, mediated through different mutations in the same gene and probably due to loss of constraint – a perfect example for parallel evolution targeting one gene.

These examples demonstrate that not only the gain but also the elimination of a GPCR function (pseudogenization) may have an evolutionary advantage and may also trigger adaptation.

4 Future Directions

A central question in GPCR research addresses understanding of the molecular relevance of every single amino acid position in a given GPCR. Solving the crystal structure of rhodopsin and many other GPCRs was a milestone towards this goal [114]. However, the predictive value of GPCRs regarding the individual relevance of residues and structural elements from crystal structures is still very limited [115] because crystal structures provide only snapshots of certain activity state-related conformations. But can an evolutionary view help and contribute to interpret receptor positions functionally?

In the light of evolution, the conservation and the variability of GPCR function at the molecular level are balanced between two processes – purifying and positive selection. Sequence comparison of large ortholog data sets from extant species may represent the status quo of GPCR evolution and enables us to extract variable and conserved positions assuring receptor function. However, is such ortholog sequence data indeed suitable to extrapolate the functional relevance of every single position in a given GPCR? One should keep in mind that discrepancies between presence/absence of *in vivo* (ortholog data) and *in vitro* function (experimental data) may be caused by the structural environment because orthologs often differ in more than one position. For example, inactive mutants regain their functionality by second site mutations [116]. Also, some artificially generated variants with wild-type function will never occur in nature because its generation would require more than one mutation within the codon. In nature, such variants arise by sequential mutations (step-by-step). However, this process requires an

intermediate variant which may be functionally inactive and is therefore eliminated from the gene pool (a “dead-end” intermediate). Further, in some GPCR variants, even synonymous substitution may have a disastrous effect at mRNA or translational levels (mutational induction of splice sites, change of RNA folding and stability, species-specific codon usage). Nevertheless, recent advances in gene synthesis technology, in high-throughput sequencing, as well as in high-throughput assays have made it possible to generate and functionally test comprehensive mutant libraries. Thus, the functional effect of every possible substitution at each amino acid position within GPCR can be now individually determined [81, 117]. For example, sequence analysis of vertebrate orthologs of the ADP receptor P2Y₁₂ revealed that this amino acid variability ensuring proper receptor function *in vivo* highly correlates (>90 %) with *in vitro* experimental data [81]. Such *in vitro* evolutionary approaches can provide information about biophysical receptor stability for structural analysis [117, 118] and can even generate GPCRs with artificial pharmacology [119–121].

In a few years, the consequence of gene loss for all GPCRs will be determined by gene-deficient mouse models. Complete gene loss is a rather rare event in humans, and only 30–40 monogenetic diseases caused by mutations in GPCR are described [64]. More than 50 % of all GPCR-deficient mouse models show no phenotype or must be challenged to expose phenotypic differences compared to the wild-type strain. This indicates that the many receptor variants found in human populations and cohorts may not drastically influence receptor function and, therefore, contribute to more distinct human phenotypes. For instance, variations in the thyrotropin-releasing hormone receptor gene, melanocortin receptor type 4, chemokine receptor CXCL2, and GPR133 gene are associated with lean body mass [122], body weight [123], white blood cell count [124], and body height [125], respectively. GWAS data correlates with GPCR variation in pathologies such as attention-deficit/hyperactivity disorder (ADHD) and the glutamate-like receptor, GPRC5B [126]; early-onset venous thromboembolism and the brain-specific angiogenesis inhibitor secretin family GPCR, BAI3 [127]; and primary biliary cirrhosis and the chemokine receptor, CXCR5 [128]. Further, phenotypes in response to drug application can be associated with variants in GPCR genes, such as platelet aggregation and the adrenergic α_{2A} receptor [129]. It should be noted that in most cases signals in GWAS come from linked SNPs and the underlying variant is usually unknown. Additionally, phenotypes are probably not directly linked to a GPCR variant but depend on the presence of one or more gene variants. Finally, the impact of GPCR variants in the expression of other genes (epistatic networks) will be an important task of future GPCR research.

Acknowledgments

My special thanks go to all current and past members of our group who worked with us on GPCR structure-function relationships and GPCR evolution over the years. Most of our work was supported by the German Research Foundation (DFG) and the Germany Federal Ministry of Education and Research (BMBF).

References

1. Schiöth HB, Fredriksson R (2005) The GRAFS classification system of G-protein coupled receptors in comparative perspective. *Gen Comp Endocrinol* 142:94–101
2. Allentoft ME, Collins M, Harker D et al (2012) The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proc Biol Sci* 279:4724–4733
3. Oesterhelt D (1998) The structure and mechanism of the family of retinal proteins from halophilic archaea. *Curr Opin Struct Biol* 8:489–500
4. Fuhrman JA, Schwalbach MS, Stingl U (2008) Proteorhodopsins: an array of physiological roles? *Nat Rev Microbiol* 6:488–494
5. Sineshchekov OA, Jung KH, Spudich JL (2002) Two rhodopsins mediate phototaxis to low- and high-intensity light in *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci U S A* 99:8689–8694
6. Waschuk SA, Bezerra AG, Shi L et al (2005) *Leptosphaeria rhodopsin*: bacteriorhodopsin-like proton pump from a eukaryote. *Proc Natl Acad Sci U S A* 102:6879–6883
7. Soppa J (1994) Two hypotheses—one answer. Sequence comparison does not support an evolutionary link between halobacterial retinal proteins including bacteriorhodopsin and eukaryotic G-protein-coupled receptors. *FEBS Lett* 342:7–11
8. Ea E (2000) Pheromone response, mating and cell biology. *Curr Opin Microbiol* 3: 573–581
9. Fujisawa Y, Kato H, Iwasaki Y (2001) Structure and function of heterotrimeric G proteins in plants. *Plant Cell Physiol* 42:789–794
10. Kim JY, Haastert PV, Devreotes PN (1996) Social senses: G-protein-coupled receptor signaling pathways in *Dictyostelium discoideum*. *Chem Biol* 3:239–243
11. Violin JD, Lefkowitz RJ (2007) Beta-arrestin-biased ligands at seven-transmembrane receptors. *Trends Pharmacol Sci* 28:416–422
12. Strotmann R, Schrock K, Boselt I et al (2011) Evolution of GPCR: change and continuity. *Mol Cell Endocrinol* 331:170–178
13. Nordstrom KJ, Sallman Almen M, Edstam M et al (2011) Independent HHsearch, Needleman–Wunsch-based, and motif analyses reveal the overall hierarchy for most of the G protein-coupled receptor families. *Mol Biol Evol* 28:2471–2480
14. Krishnan A, Almen MS, Fredriksson R et al (2012) The origin of GPCRs: identification of mammalian like Rhodopsin, Adhesion, Glutamate and Frizzled GPCRs in fungi. *PLoS One* 7:e29817
15. Nordstrom KJ, Lagerstrom MC, Waller LM et al (2009) The Secretin GPCRs descended from the family of Adhesion GPCRs. *Mol Biol Evol* 26:71–84
16. Langenhan T, Promel S, Mestek L et al (2009) Latrophilin signaling links anterior-posterior tissue polarity and oriented cell divisions in the *C. elegans* embryo. *Dev Cell* 17:494–504
17. Metpally RP, Sowdhamini R (2005) Cross genome phylogenetic analysis of human and *Drosophila* G protein-coupled receptors: application to functional annotation of orphan receptors. *BMC Genomics* 6:106
18. Metpally RP, Sowdhamini R (2005) Genome wide survey of G protein-coupled receptors in *Tetraodon nigroviridis*. *BMC Evol Biol* 5:41
19. Kamesh N, Aradhyam GK, Manoj N (2008) The repertoire of G protein-coupled receptors in the sea squirt *Ciona intestinalis*. *BMC Evol Biol* 8:129
20. Schöneberg T, Hermsdorf T, Engemaier E et al (2007) Structural and functional evolution of the P2Y(12)-like receptor group. *Purinergic Signal* 3:255–268
21. Gloriam DE, Fredriksson R, Schiöth HB (2007) The G protein-coupled receptor subset of the rat genome. *BMC Genomics* 8:338
22. Kishida T (2008) Pattern of the divergence of olfactory receptor genes during tetrapod evolution. *PLoS One* 3:e2385
23. Gilad Y, Przeworski M, Lancet D (2004) Loss of olfactory receptor genes coincides with the acquisition of full trichromatic vision in primates. *PLoS Biol* 2:e5

24. Hayden S, Bekaert M, TA C et al (2010) Ecological adaptation determines functional mammalian olfactory subgenomes. *Genome Res* 20:1–9
25. Steiger SS, Kuryshv VY, Stensmyr MC et al (2009) A comparison of reptilian and avian olfactory receptor gene repertoires: species-specific expansion of group gamma genes in birds. *BMC Genomics* 10:446
26. Chiari Y, Cahais V, Galtier N et al (2012) Phylogenomic analyses support the position of turtles as the sister group of birds and crocodiles (Archosauria). *BMC Biol* 10:65
27. Wurm Y, Wang J, Riba-Grognuz O et al (2011) The genome of the fire ant *Solenopsis invicta*. *Proc Natl Acad Sci U S A* 108:5679–5684
28. Ji Y, Zhang Z, Hu Y (2009) The repertoire of G-protein-coupled receptors in *Xenopus tropicalis*. *BMC Genomics* 10:263
29. Hwang JI, Moon MJ, Park S et al (2013) Expansion of secretin-like G protein-coupled receptors and their peptide ligands via local duplications before and after two rounds of whole-genome duplication. *Mol Biol Evol* 30:1119–1130
30. Semyonov J, Park JI, Cl C et al (2008) GPCR genes are preferentially retained after whole genome duplication. *PLoS One* 3:e1903
31. Huang Y, Zheng Y, Su Z et al (2009) Differences in duplication age distributions between human GPCRs and their downstream genes from a network prospective. *BMC Genomics* 10(Suppl 1):S14
32. Balakirev ES, Ayala FJ (2003) Pseudogenes: are they “junk” or functional DNA? *Annu Rev Genet* 37:123–151
33. Pawson AJ, Morgan K, Maudsley SR et al (2003) Type II gonadotrophin-releasing hormone (GnRH-II) in reproductive biology. *Reproduction* 126:271–278
34. Hashiguchi Y, Nishida M (2007) Evolution of trace amine associated receptor (TAAR) gene family in vertebrates: lineage-specific expansions and degradations of a second class of vertebrate chemosensory receptors expressed in the olfactory epithelium. *Mol Biol Evol* 24:2099–2107
35. Jiang P, Josue J, Li X et al (2012) Major taste loss in carnivorous mammals. *Proc Natl Acad Sci U S A* 109:4956–4961
36. Hurst LD (2002) The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet* 18:486
37. Staubert C, Boselt I, Bohnkamp J et al (2010) Structural and functional evolution of the trace amine-associated receptors TAAR3, TAAR4 and TAAR5 in primates. *PLoS One* 5:e11133
38. Starback P, Wraith A, Eriksson H et al (2000) Neuropeptide Y receptor gene y6: multiple deaths or resurrections? *Biochem Biophys Res Commun* 277:264–269
39. Martin RD, Soligo C, Tavaré S (2007) Primate origins: implications of a cretaceous ancestry. *Folia Primatol (Basel)* 78:277–296
40. Rompler H, Schulz A, Pitra C et al (2005) The rise and fall of the chemoattractant receptor GPR33. *J Biol Chem* 280:31068–31075
41. Gaillard I, Rouquier S, Chavanieu A et al (2004) Amino-acid changes acquired during evolution by olfactory receptor 912–93 modify the specificity of odorant recognition. *Hum Mol Genet* 13:771–780
42. Rompler H, Yu HT, Arnold A et al (2006) Functional consequences of naturally occurring DRY motif variants in the mammalian chemoattractant receptor GPR33. *Genomics* 87:724–732
43. Torrents D, Suyama M, Zdobnov E et al (2003) A genome-wide survey of human pseudogenes. *Genome Res* 13:2559–2567
44. Pei B, Sisu C, Frankish A et al (2012) The GENCODE pseudogene resource. *Genome Biol* 13:R51
45. Muro EM, Mah N, Andrade-Navarro MA (2011) Functional evidence of post-transcriptional regulation by pseudogenes. *Biochimie* 93:1916–1921
46. Olsen MA, Schechter LE (1999) Cloning, mRNA localization and evolutionary conservation of a human 5-HT7 receptor pseudogene. *Gene* 227:63–69
47. Bard JA, Nawoschik SP, O’Dowd BF et al (1995) The human serotonin 5-hydroxytryptamine 1D receptor pseudogene is transcribed. *Gene* 153:295–296
48. Young JM, Shykind BM, Lane RP et al (2003) Odorant receptor expressed sequence tags demonstrate olfactory expression of over 400 genes, extensive alternate splicing and unequal expression levels. *Genome Biol* 4:R71
49. Abecasis GR, Auton A, Brooks LD et al (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature* 491:56–65
50. Gibbs RA, Taylor JF, Van Tassell CP et al (2009) Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science* 324:528–532
51. Yalcin B, Nicod J, Bhomra A et al (2010) Commercially available outbred mice for genome-wide association studies. *PLoS Genet* 6:e1001085
52. Grossman SR, Andersen KG, Shlyakhter I et al (2013) Identifying recent adaptations in large-scale genomic data. *Cell* 152:703–713
53. Hirschhorn JN, Gajdos ZK (2011) Genome-wide association studies: results from the first few years and potential implications for clinical medicine. *Annu Rev Med* 62:11–24

54. Yandell M, Ence D (2012) A beginner's guide to eukaryotic genome annotation. *Nat Rev Genet* 13:329–342
55. Bernardi G, Wiley EO, Mansour H et al (2012) The fishes of Genome 10K. *Mar Genomics* 7:3–6
56. Anonymous (2009) Genome 10K: a proposal to obtain whole-genome sequence for 10,000 vertebrate species. *J Hered* 100:659–674
57. Green RE, Krause J, Ptak SE et al (2006) Analysis of one million base pairs of Neanderthal DNA. *Nature* 444:330–336
58. Meyer M, Kircher M, Mt G et al (2012) A high-coverage genome sequence from an archaic Denisovan individual. *Science* 338: 222–226
59. Miller W, Drautz DI, Ratan A et al (2008) Sequencing the nuclear genome of the extinct woolly mammoth. *Nature* 456:387–390
60. Rohland N, Reich D, Mallick S et al (2010) Genomic DNA sequences from mastodon and woolly mammoth reveal deep speciation of forest and savanna elephants. *PLoS Biol* 8:e1000564
61. Noonan JP, Hofreiter M, Smith D et al (2005) Genomic sequencing of Pleistocene cave bears. *Science* 309:597–599
62. Grandy DK, Litt M, Allen L et al (1989) The human dopamine D2 receptor gene is located on chromosome 11 at q22-q23 and identifies a TaqI RFLP. *Am J Hum Genet* 45:778–785
63. Jensen S, Plaetke R, Holik J et al (1992) Linkage analysis of the D1 dopamine receptor gene and manic depression in six families. *Hum Hered* 42:269–275
64. Schoneberg T, Schulz A, Biebermann H et al (2004) Mutant G-protein-coupled receptors as a cause of human diseases. *Pharmacol Ther* 104:173–206
65. Lalueza-Fox C, Rompler H, Caramelli D et al (2007) A melanocortin 1 receptor allele suggests varying pigmentation among Neanderthals. *Science* 318:1453–1455
66. Rompler H, Rohland N, Lalueza-Fox C et al (2006) Nuclear gene indicates coat-color polymorphism in mammoths. *Science* 313:62
67. Kreuchwig A, Kleinau G, Kreuchwig F et al (2011) Research resource: update and extension of a glycoprotein hormone receptors web application. *Mol Endocrinol* 25:707–712
68. Pidasheva S, D'souza-LI L, Canaff L et al (2004) CASRdb: calcium-sensing receptor locus-specific database for mutations causing familial (benign) hypocalciuric hypercalcemia, neonatal severe hyperparathyroidism, and autosomal dominant hypocalcemia. *Hum Mutat* 24:107–111
69. Kazius J, Wurdinger K, Van Iterson M et al (2008) GPCR NaVa database: natural variants in human G protein-coupled receptors. *Hum Mutat* 29:39–44
70. Renaud G, Kircher M, Stenzel U et al (2013) freeIbis: an efficient basecaller with calibrated quality scores for Illumina sequencers. *Bioinformatics* 29:1208–1209
71. Small KM, Seman CA, Castator A et al (2002) False positive non-synonymous polymorphisms of G-protein coupled receptor genes. *FEBS Lett* 516:253–256
72. Tang CM, Insel PA (2005) Genetic variation in G-protein-coupled receptors—consequences for G-protein-coupled receptors as drug targets. *Expert Opin Ther Targets* 9:1247–1265
73. Frazer KA, Ballinger DG, Cox DR et al (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449:851–861
74. Nielsen R, Williamson S, Kim Y et al (2005) Genomic scans for selective sweeps using SNP data. *Genome Res* 15:1566–1575
75. Scheinfeldt LB, Biswas S, Madeoy J et al (2011) Clusters of adaptive evolution in the human genome. *Front Genet* 2:1–10
76. Sabeti PC, Schaffner SF, Fry B et al (2006) Positive natural selection in the human lineage. *Science* 312:1614–1620
77. Charlesworth B (2012) The effects of deleterious mutations on evolution at linked sites. *Genetics* 190:5–22
78. Alves I, Sramkova Hanulova A, Foll M et al (2012) Genomic data reveal a complex making of humans. *PLoS Genet* 8: e1002837
79. Schulz A, Schoneberg T (2003) The structural evolution of a P2Y-like G-protein-coupled receptor. *J Biol Chem* 278: 35531–35541
80. Boselt I, Rompler H, Hermsdorf T et al (2009) Involvement of the V2 vasopressin receptor in adaptation to limited water supply. *PLoS One* 4:e5573
81. Coster M, Wittkopf D, Kreuchwig A et al (2012) Using ortholog sequence data to predict the functional relevance of mutations in G-protein-coupled receptors. *FASEB J* 26: 3273–3281
82. Rennison DJ, Owens GL, Taylor JS (2012) Opsin gene duplication and divergence in ray-finned fish. *Mol Phylogenet Evol* 62: 986–1008
83. Cao J, Huang S, Qian J et al (2009) Evolution of the class C GPCR Venus flytrap modules involved positive selected functional divergence. *BMC Evol Biol* 9:67
84. Liebscher I, Muller U, Teupser D et al (2011) Altered immune response in mice deficient for the G protein-coupled receptor GPR34. *J Biol Chem* 286:2101–2110

85. Engel KM, Schrock K, Teupser D et al (2011) Reduced food intake and body weight in mice deficient for the G protein-coupled receptor GPR82. *PLoS One* 6:e29400
86. Herberger AL, Loretz CA (2013) Vertebrate extracellular calcium-sensing receptor evolution: selection in relation to life history and habitat. *Comp Biochem Physiol Part D Genomics Proteomics* 8:86–94
87. Cone RD, Lu D, Koppula S et al (1996) The melanocortin receptors: agonists, antagonists, and the hormonal control of pigmentation. *Recent Prog Horm Res* 51:287–318
88. Pointer MA, Mundy NI (2008) Testing whether macroevolution follows microevolution: are colour differences among swans (*Cygnus*) attributable to variation at the MC1R locus? *BMC Evol Biol* 8:249
89. Nadeau NJ, Burke T, Mundy NI (2007) Evolution of an avian pigmentation gene correlates with a measure of sexual selection. *Proc Biol Sci* 274:1807–1813
90. Nachman MW, Hoekstra HE, D'agostino SL (2003) The genetic basis of adaptive melanism in pocket mice. *Proc Natl Acad Sci U S A* 100:5268–5273
91. Briscoe AD, Bybee SM, Bernard GD et al (2010) Positive selection of a duplicated UV-sensitive visual pigment coincides with wing pigment evolution in *Heliconius* butterflies. *Proc Natl Acad Sci U S A* 107:3628–3633
92. Steiger SS, Fidler AE, Mueller JC et al (2010) Evidence for adaptive evolution of olfactory receptor genes in 9 bird species. *J Hered* 101:325–333
93. Moreno-Estrada A, Casals F, Ramirez-Soriano A et al (2008) Signatures of selection in the human olfactory receptor OR511 gene. *Mol Biol Evol* 25:144–154
94. Olender T, Waszak SM, Viavant M et al (2012) Personal receptor repertoires: olfaction as a model. *BMC Genomics* 13:414
95. Soranzo N, Bufo B, PC S et al (2005) Positive selection on a high-sensitivity allele of the human bitter-taste receptor TAS2R16. *Curr Biol* 15:1257–1265
96. Wooding S (2011) Signatures of natural selection in a primate bitter taste receptor. *J Mol Evol* 73:257–265
97. Campbell MC, Ranciaro A, Froment A et al (2012) Evolution of functionally diverse alleles associated with PTC bitter taste sensitivity in Africa. *Mol Biol Evol* 29:1141–1153
98. Spielman SJ, Wilke CO (2013) Membrane environment imposes unique selection pressures on transmembrane domains of G protein-coupled receptors. *J Mol Evol* 76:172–182
99. Deville J, Rey J, Chabbert M (2009) An indel in transmembrane helix 2 helps to trace the molecular evolution of class A G-protein-coupled receptors. *J Mol Evol* 68:475–489
100. Fatakia SN, Costanzi S, Chow CC (2011) Molecular evolution of the transmembrane domains of G protein-coupled receptors. *PLoS One* 6:e27813
101. Choi SS, Lahn BT (2003) Adaptive evolution of MRG, a neuron-specific gene family implicated in nociception. *Genome Res* 13:2252–2259
102. Yang S, Liu Y, Aa L et al (2005) Adaptive evolution of MRGX2, a human sensory neuron specific gene involved in nociception. *Gene* 352:30–35
103. Wang X, Grus WE, Zhang J (2006) Gene losses during human origins. *PLoS Biol* 4:e52
104. Rompler H, Staubert C, Thor D et al (2007) G protein-coupled time travel: evolutionary aspects of GPCR research. *Mol Interv* 7:17–25
105. Wang X, Thomas SD, Zhang J (2004) Relaxation of selective constraint and loss of function in the evolution of human bitter taste receptor genes. *Hum Mol Genet* 13:2671–2678
106. Zhang J, Webb DM (2003) Evolutionary deterioration of the vomeronasal pheromone transduction pathway in catarrhine primates. *Proc Natl Acad Sci U S A* 100:8337–8341
107. Harding RM, Healy E, Ray AJ et al (2000) Evidence for variable selective pressures at MC1R. *Am J Hum Genet* 66:1351–1361
108. John PR, Makova K, Li WH et al (2003) DNA polymorphism and selection at the melanocortin-1 receptor gene in normally pigmented southern African individuals. *Ann N Y Acad Sci* 994:299–306
109. Rees JL (2000) The melanocortin 1 receptor (MC1R): more than just red hair. *Pigment Cell Res* 13:135–140
110. Gross JB, Borowsky R, Tabin CJ (2009) A novel role for Mc1r in the parallel evolution of depigmentation in independent populations of the cavefish *Astyanax mexicanus*. *PLoS Genet* 5:e1000326
111. Hoekstra HE, Hirschmann RJ, Bunday RA et al (2006) A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* 313:101–104
112. Steiner CC, Rompler H, Boettger LM et al (2009) The genetic basis of phenotypic convergence in beach mice: similar pigment patterns but different genes. *Mol Biol Evol* 26:35–45
113. Rosenblum EB, Rompler H, Schoneberg T et al (2010) Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *Proc Natl Acad Sci U S A* 107:2113–2117

114. Rosenbaum DM, Rasmussen SG, Kobilka BK (2009) The structure and function of G-protein-coupled receptors. *Nature* 459: 356–363
115. Michino M, Abola E, Brooks CL et al (2009) Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nat Rev Drug Discov* 8: 455–463
116. Li B, Nowak NM, Kim SK et al (2005) Random mutagenesis of the M3 muscarinic acetylcholine receptor expressed in yeast: identification of second-site mutations that restore function to a coupling-deficient mutant M3 receptor. *J Biol Chem* 280: 5664–5675
117. Schlinkmann KM, Honegger A, Tureci E et al (2012) Critical features for biosynthesis, stability, and functionality of a G protein-coupled receptor uncovered by all-versus-all mutations. *Proc Natl Acad Sci U S A* 109: 9810–9815
118. Serrano-Vega MJ, Magnani F, Shibata Y et al (2008) Conformational thermostabilization of the betal-adrenergic receptor in a detergent-resistant form. *Proc Natl Acad Sci U S A* 105:877–882
119. Nichols CD, Roth BL (2009) Engineered G-protein coupled receptors are powerful tools to investigate biological processes and behaviors. *Front Mol Neurosci* 2:16
120. Conklin BR, Hsiao EC, Claeysen S et al (2008) Engineering GPCR signaling pathways with RASSLs. *Nat Methods* 5:673–678
121. Thor D, Schulz A, Hermsdorf T et al (2008) Generation of an agonistic binding site for blockers of the M(3) muscarinic acetylcholine receptor. *Biochem J* 412:103–112
122. Liu XG, Tan LJ, Lei SF et al (2009) Genome-wide association and replication studies identified TRHR as an important gene for lean body mass. *Am J Hum Genet* 84:418–423
123. Xi B, Chandak GR, Shen Y et al (2012) Association between common polymorphism near the MC4R gene and obesity risk: a systematic review and meta-analysis. *PLoS One* 7:e45731
124. Reiner AP, Lettre G, Nalls MA et al (2011) Genome-wide association study of white blood cell count in 16,388 African Americans: the continental origins and genetic epidemiology network (COGENT). *PLoS Genet* 7:e1002108
125. Tonjes A, Koriath M, Schleinitz D et al (2009) Genetic variation in GPR133 is associated with height: genome wide association study in the self-contained population of Sorbs. *Hum Mol Genet* 18:4662–4668
126. Albayrak O, Putter C, Al V et al (2013) Common obesity risk alleles in childhood attention-deficit/hyperactivity disorder. *Am J Med Genet B Neuropsychiatr Genet* 162:295–305
127. Antoni G, Morange PE, Luo Y et al (2010) A multi-stage multi-design strategy provides strong evidence that the BAI3 locus is associated with early-onset venous thromboembolism. *J Thromb Haemost* 8:2671–2679
128. Mells GF, Floyd JA, Morley KI et al (2011) Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. *Nat Genet* 43:329–332
129. Johnson AD, Yanek LR, Chen MH et al (2010) Genome-wide meta-analyses identifies seven loci associated with platelet aggregation in response to agonists. *Nat Genet* 42: 608–613

G Protein-Coupled Receptor Genetics

Research and Methods in the Post-Genomic Era

Stevens, C.W. (Ed.)

2014, XI, 423 p. 66 illus., 40 illus. in color., Hardcover

ISBN: 978-1-62703-778-5

A product of Humana Press