MOLECULAR EVOLUTION OF GPCRS

GLP1/GLP1 receptors

Jong-Ik Hwang, Seongsik Yun, Mi Jin Moon, Cho Rong Park and **Jae Young Seong**

Graduate School of Medicine, Korea University, Seoul 136-705, Republic of Korea

Correspondence
should be addressed to
J Y Seong
Email
jyseong@korea.ac.kr

Abstract

Glucagon-like peptide 1 (GLP1) is an intestinal incretin that regulates glucose homeostasis through stimulation of insulin secretion from pancreatic β-cells and inhibits appetite by acting on the brain. Thus, it is a promising therapeutic agent for the treatment of type 2 diabetes mellitus and obesity. Studies using synteny and reconstructed ancestral chromosomes suggest that families for GLP1 and its receptor (GLP1R) have emerged through two rounds (2R) of whole genome duplication and local gene duplications before and after 2R. Exon duplications have also contributed to the expansion of the peptide family members. Specific changes in the amino acid sequence following exon/gene/genome duplications have established distinct yet related peptide and receptor families. These specific changes also confer selective interactions between GLP1 and GLP1R. In this review, we present a possible macro (genome level)- and micro (gene/exon level)-evolution mechanisms of GLP1 and GLP1R, which allows them to acquire selective interactions between this ligand–receptor pair. This information may provide critical insight for the development of potent therapeutic agents targeting GLP1R.

Introduction

Glucagon-like peptide 1 (GLP1) is a gastrointestinal peptide hormone that stimulates insulin secretion from pancreatic β-cells and promotes insulin synthesis (Schmidt *et al.* 1985, Fehmann & Habener 1992, Moon *et al.* 2011). GLP1, encoded by the glucagon (*GCG*) gene, is produced by tissue-specific posttranslational proteolytic processing of a large precursor containing three independent peptides GCG, GLP1, and GLP2 (Mojsov *et al.* 1986, Kieffer & Habener 1999). Intestinal L-cells produce mature GLP1 and GLP2 but not GCG, while pancreatic α-cells produce mature GCG but not GLP1 and GLP2. Biologically active GLP1s consist of 30 or 31 amino acids, both of which have similar bioactivities and overall metabolism, although most GLP1 secreted from the gut is C-terminally amidated and 30 amino acids in length (Mojsov *et al.* 1990). GLP1, secreted from L-cells in response to food intake, regulates blood glucose level by inducing insulin secretion in a glucose concentration-dependent manner, while it suppresses GCG release from pancreatic α-cells (Fehmann *et al.* 1995, Hansotia & Drucker 2005). The glucose-dependent action of GLP1 has potential for the treatment of diabetes mellitus because it does not cause hypoglycemia under normal plasma glucose concentration conditions (Baggio & Drucker 2002, Nauck *et al.* 2002). In addition to its incretin activity, GLP1 stimulates proliferation and differentiation of pancreatic β-cells (Xu *et al.* 1999, Drucker 2003, Li *et al.* 2003). GLP1 is also produced in neurons of the brainstem and is transported through axonal networks to diverse brain regions such as the cerebral cortex, hypothalamus, and thalamus

(Vrang *et al.* 2007, Hisadome *et al.* 2010). In the brain, GLP1 is known to promote satiety by regulating appetite and food intake, resulting in body weight loss (Zander *et al.* 2002). GLP1 is also involved in neurite outgrowth and spatial learning ability in the central nervous system (During *et al.* 2003, Perry *et al.* 2007).

The physiological activities of GLP1 occur by interaction with the GLP1 receptor (GLP1R), which is expressed in pancreatic islets, brain, heart, kidney, and the gastrointestinal tract (Graziano *et al.* 1993, Wei & Mojsov 1995, Alvarez *et al.* 2005). Ligand-stimulated receptors increase intracellular cAMP by coupling to adenylate cyclase through the G_s protein (Drucker *et al.* 1987, Fehmann *et al.* 1995) and by elevating cytosolic calcium through activation of the G_q-mediated PLCβ pathway (Wheeler *et al.* 1993). However, one study reports that calcium responses may be secondary effects related to increased intracellular cAMP (Kang *et al.* 2001).

GLP1R is a member of the secretin receptor (SCTR)-like (class B1) family of G protein-coupled receptors (GPCRs) that comprise 15 human receptors for 20 corresponding peptide ligands (Harmar 2001, Hwang *et al.* 2013). As the SCTR-like receptors and their peptide ligands share similar structural features, they are likely to have emerged through gene/genome duplications from common ancestors. Duplicated receptors and peptides have undergone sequence changes during evolution, leading to diversification of their physiological functions (Acharjee *et al.* 2004, Kim *et al.* 2012). Specific diversification of peptides, i.e. conservation within orthologs but variation among paralogs, could confer selective interaction of a peptide with the cognate receptor, allowing discrimination by paralogous receptors (Wang *et al.* 2004, Li *et al.* 2005, Cho *et al.* 2007, Lee *et al.* 2009). This review describes the influence of genome/gene/exon duplications on the emergence of GLP1 and GLP1R, and the specific changes allowing selective interaction between this ligand–receptor pair.

Evolution of SCTR-like GPCRs and their peptide ligands

Since the first identification of SCTR that exhibits amino acid sequence distinct from that of other known GPCRs (Ishihara *et al.* 1991), the SCTR-like receptors including GLP1R have been identified in a variety of invertebrate and vertebrate species (Harmar 2001). SCTR-like GPCRs can be categorized into the following: i) corticotropin-releasing hormone receptor (CRHR) subfamily, ii) calcitonin receptor (CALCR) subfamily, iii) parathyroid hormone

receptor (PTHR) subfamily, iv) GCG receptor (GCGR) subfamily including GCGR, GLP1R, GLP2R, and glucose-dependent insulinotropic polypeptide receptor (GIPR), and v) growth hormone-releasing hormone receptor (GHRHR) subfamily including GHRHR, SCTR, vasoactive intestinal peptide receptor 1 (VIPR1), VIPR2, and pituitary adenylate cyclase-activating polypeptide receptor (ADCYAP1R1) (Harmar 2001, Fredriksson *et al.* 2003, Oh *et al.* 2006, Hwang *et al.* 2013).

Although GLP1 and GLP1R are found only in vertebrates (Sherwood *et al.* 2000, Irwin & Prentice 2011), the genome history of SCTR-like GPCRs and their peptide ligand genes date back to early bilaterian ancestors (Harmar 2001, Cardoso *et al.* 2006). Recent advances in bioinformatic tools and public genome databases for a wide range of animal kingdoms have allowed identification of additional members and exploration of evolutionary history of the SCTR-like family (Cardoso *et al.* 2006, Nordstrom *et al.* 2008, Irwin & Prentice 2011, Hwang *et al.* 2013, Park *et al.* 2013). CRHR and CALCR subfamilies are likely the earliest members of this family as they are present in most deuterostomes as well as in the protostomes *Caenorhabditis elegans* and *Caenorhabditis briggsae* (nematode) and *Drosophila melanogaster* and *Anopheles gambiae* (arthropod) (Harmar 2001, Cardoso *et al.* 2006). In addition, *D. melanogaster* contains genes that are phylogenetically closest to CRH and CALC out of all the peptide family members (Hewes & Taghert 2001, Hwang *et al.* 2013). Short-range linkage analysis between protostome and deuterostome homologous regions that contain SCTR-like receptor genes demonstrated the existence of common neighbor families for SCTR-like receptor genes (Cardoso *et al.* 2006). Recently, the presence of protostomian PTHR-like has been demonstrated (Mirabeau & Joly 2013). Thus, it is plausible that the ancestors for the *CRHR* (*CRHR1*), *CALCR*, and *PTHR* subfamilies have emerged during early metazoan evolution before the divergence of deuterostomes and protostomes. In the deuterostome lineage, besides the *CRHR*, *CALCR*, and *PTHR* gene subfamilies, the *GCGR* and *GHRHR* subfamily genes occurred during chordate and vertebrate evolution. The amphioxus (*Branchiostoma floridae*, a representative of the cephalochordates) genome contains the *CRHR*, *CALCR*, and *PTHR* subfamily genes, but lacks the *GCGR* and *GHRHR* subfamily genes (Nordstrom *et al.* 2008, Hwang *et al.* 2013). Blast search of the tunicate (*Ciona intestinalis*, a representative of the urochordates) genome shows the emergence of the *GCGR* subfamily genes together with the *CRHR*, *CALCR*, and *PTHR* subfamily genes (Cardoso *et al.* 2006,

Hwang *et al.* 2013). However, no *GHRHR* subfamily genes have been identified in tunicates. In lampreys (*Petromyzon marinus*, a basal vertebrate agnathan), a *VIPR*-like gene, a member of the *GHRHR* subfamily genes was identified (Ng *et al.* 2012, Hwang *et al.* 2013). Thus, lamprey has SCTR-like receptors belonging to all five subfamilies.

Unlike the receptor genes, the *SCT*-like peptide family genes were not retrieved from nonvertebrates (protostome and early deuterostome), likely due to the considerable variation in short peptide sequences of these species (Cardoso *et al.* 2010, Hwang *et al.* 2013). In basal vertebrate lamprey, however, several peptide genes such as adreno-medullin, *GHRH*, *VIP*, *ADCYAP1*, and *GCG* were observed (Cardoso *et al.* 2010, Hwang *et al.* 2013). Despite the presence of three *CRHR*-like genes, the corresponding *CRH*-like peptide genes were not found in lamprey. As in other vertebrates, the lamprey *VIP*, *ADCYAP1*, and *GCG* genes may encode two or more mature peptides (Ng *et al.* 2012, Hwang *et al.* 2013). As the mature peptides have similarities in their amino acid sequences on consecutive exons, these peptides are proposed to have emerged through exon duplication events during early vertebrate evolution (Irwin *et al.* 1997, Sherwood *et al.* 2000, Irwin & Prentice 2011).

Gene/genome duplications contribute to the emergence of GLP1 and GLP1R

Gene families with large numbers of paralogs are established through evolutionary processes, such as gene/genome duplications followed by gene alterations or losses (Abi-Rached *et al.* 2002, Larhammar *et al.* 2002, Holland 2003, Santini *et al.* 2003, Vienne *et al.* 2003, Kim *et al.* 2011, Hwang *et al.* 2013). In particular, two rounds (2R) of whole genome duplication during early vertebrate evolution are considered to mainly contribute to expansion of the family members. In addition, local gene duplications on a chromosome before or after 2R further diversify the family members (Hwang *et al.* 2013). Rapid and vast accumulation of genome sequence information of many representative taxa allows synteny-based genome comparisons and facilitates the analyses of origins and relationships of family members (Lee *et al.* 2009, Kim *et al.* 2011, 2012). Furthermore, entire genome comparisons between evolutionarily distinct taxa have led to recon-struction of hypothetical ancestral chromosomes of early vertebrates or chordates (Nakatani *et al.* 2007, Putnam *et al.* 2008). Nakatani *et al.* (2007), constructed hypo-thetical post-2R gnathostome ancestor chromosomes (*GACs A0–J1* linkage groups) that were derived from

pre-2R 10–13 vertebrate ancestral chromosomes (*VACs A–J*). Deducing locations of the gene family members to reconstructed chromosomes also provides fast and rela-tively accurate tools to trace gene origins (Yegorov & Good 2012, Hwang *et al.* 2013).

Emergence of GLP1R

When all *SCTR*-like genes were assigned on reconstructed pre-2R *VAC* as proposed by Nakatani *et al.* (2007), most genes, except for the *GCGR* subfamily, were found to be on the *VAC E* linkage group consisting of four paralogons as represented in human chromosomes 2 (*GAC E0*), 7 and 3 (*GAC E1*), 17 (*GAC E2*), and 12 (*GAC E3*) (Fig. 1). *GCGR* subfamily members are on different *GACs* (Hwang *et al.* 2013), which is likely due to some ancestral linkage groups being inaccurately reconstructed because of massive chromosome rearrangements in these regions, or alter-natively, the gene underwent a single gene translocation that caused it to move from its authentic chromosomal fragments before or after 2R. Synteny result suggests that consecutive local duplications of the *SCTR*-like ancestor on *VAC E* generated *GCGR* and *GHRHR* ancestral genes (Fig. 1). Neighbor gene analyses of *SCTR*-like-containing genome fragments of the amphioxus, tunicate, and lamprey may further support this possibility. For example, 15 amphioxus *SCTR*-like genes are located on ten different scaffolds. Many neighbor genes in these scaffolds well pair with those on human chromosomes having the SCTR family genes (Fig. 1). In tunicates, *CRHR* and *CALCR* are closely located on chromosome 3 with neighbors that are also found on *VAC E*. The *PTHR*-like on a small scaffold also shares a common neighbor gene on *VAC E*. Two *GCGR*-like genes reside closely on chromosome 3. Interestingly, their adjacent neighbors are shown on both *VAC E* and *GCGR* family gene-containing chromosomes (Fig. 1). In lamprey, although the sizes of scaffolds containing *SCTR*-like genes are small so as to have very limited numbers of neighbors or no neighbors, most neighbors in these scaffolds are mapped on *VAC E* (Fig. 1). Interestingly, one scaffold harboring the *GCGR* gene subfamily has a neighbor that localizes on *VAC E*, while another has a neighbor on *GCGR* family gene-containing chromosomes. Together, these results obtained from chordates and basal vertebrate lamprey indicate a consecutive emergence of SCTR-like GPCRs during deuterostome evolution.

The 2R and local duplications, followed by sequences change and gene loss, have established osteichthyan ancestral genes for SCTR-like GPCRs and their peptide ligand (Hwang *et al.* 2013). Regarding *GLP1R*, these
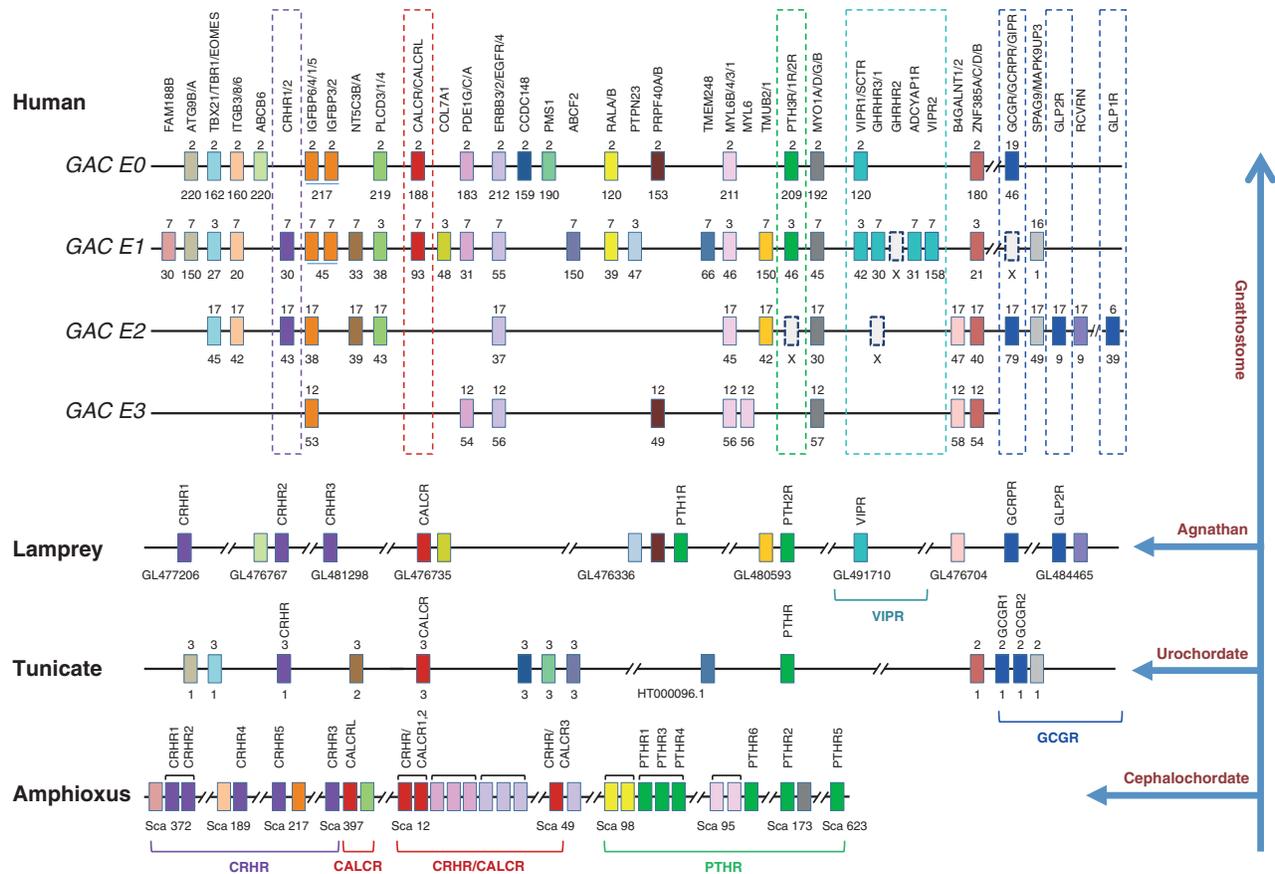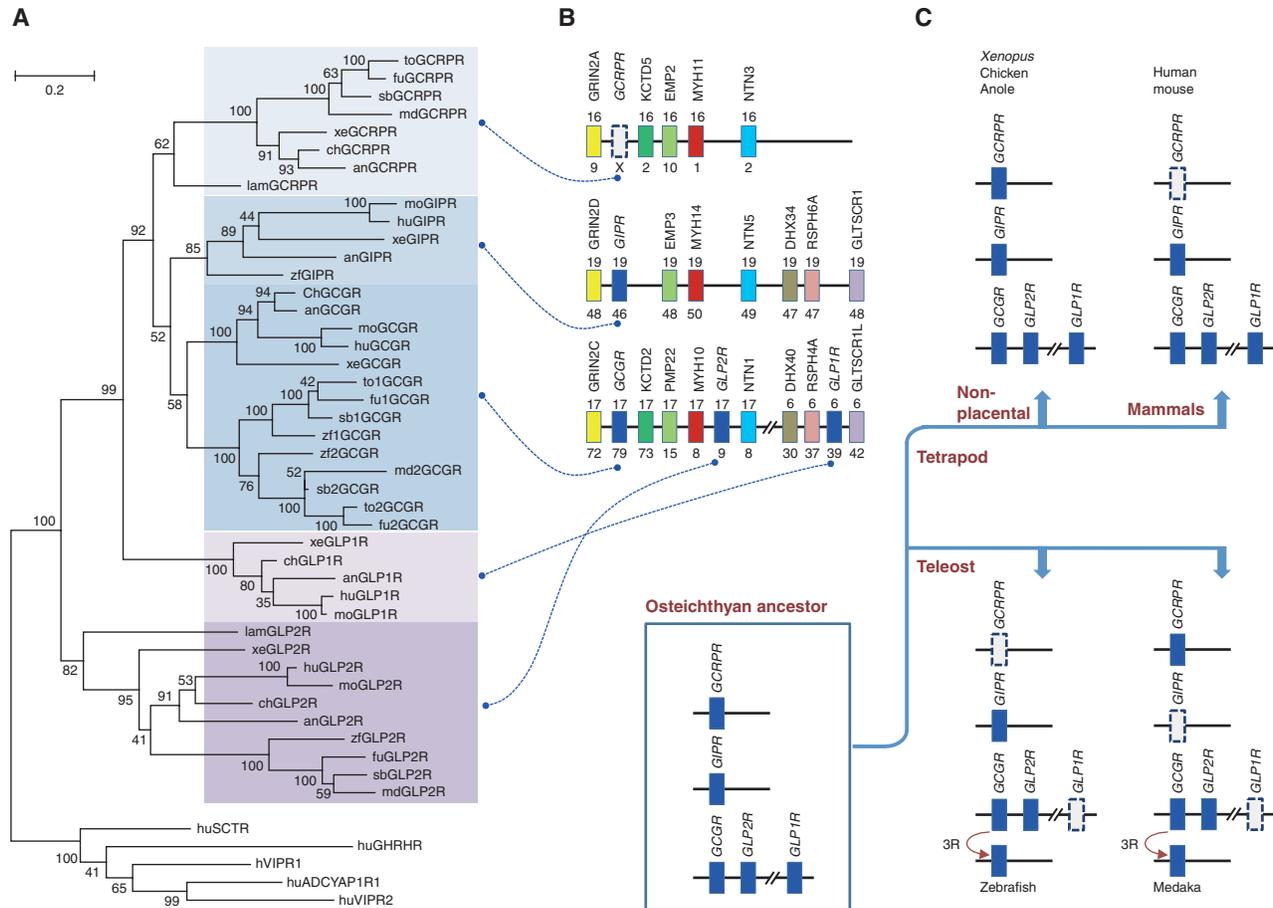
Figure 1

Conserved synteny for genome fragments containing the *SCTR*-like family genes. The genomic locations of the human *SCTR*-like family genes along with their neighbor genes were compared with those of orthologous genes in amphioxus (cephalochordate), tunicate (urochordate), and lamprey (basal vertebrate agnathan). Orthologs and paralogs are aligned on the same column with the same color. Chromosome numbers are indicated above the gene, and gene locations (megabase) are indicated below the gene in human and tunicate. In amphioxus and lamprey, scaffold numbers are shown below the genes. The absence of the gene in humans is indicated by white boxes with broken lines under which 'X' was labeled. Paralogous genes generated by 2R exist on human chromosomes 2, 7 (or 3), 17, and 12, which fall into reconstructed linkage groups *GAC E0*, *E1*, *E2*, and *E3* respectively. The *GCGR* subfamily genes were aligned on different chromosomes as they are on different linkage groups other than *GAC Es*. Annotations for amphioxus, tunicate, and lamprey *SCTR*-like family genes are shown in Hwang *et al*. (2013). In the case of lamprey, chromosome duplication events for generation of paralogous genes cannot be excluded although genes are on the same row.

gene/genome duplications contributed to the occurrence of *GCGR*, *GLP1R*, *GLP2R*, *GIPR*, and GCG-related peptide receptor (*GCRPR*) genes. *GCRPR* is a new member found in many vertebrate species but not in mammals (Irwin & Prentice 2011, Wang *et al*. 2012, Hwang *et al*. 2013, Park *et al*. 2013). The phylogenetic tree of the *GCGR* subfamily reveals two independent branches for *GLP2R* and *GLP1R*, and a branch comprising *GCGR*, *GIPR*, and *GCRPR* subbranches (Fig. 2A). Synteny analysis for *GCGR* subfamily gene-containing genome fragments of vertebrate species shows that *GCGR*, *GIPR*, and *GCRPR* are on three different paralogons that share common neighbor gene families. Although *GCRPR* is not found in humans, neighbor gene analysis of other vertebrate species suggests

that *GCRPR* was likely lost in human chromosome 16 (Fig. 2B). *GLP2R* is on the same chromosome containing *GCGR*. Although *GLP1R* is on different chromosome, it shares neighbor gene family with the *GIPR*-containing genome fragment. The phylogenetic tree of the *GCGR* subfamily gene suggests that *GLP2R* is likely an earliest form of the subfamily, followed by subsequent emergences of *GLP1R* and an ancestor for *GCGR/GIPR/GCRPR*. Combined with synteny data, this result raises a possibility that *GLP2R*, *GLP1R*, and an ancestor for *GCGR/GIPR/GCRPR* arose through local duplication before 2R while *GCGR*, *GIPR*, and *GCRPR* have emerged through 2R (Fig. 2C). Interestingly, this receptor duplication scheme is well correlated with that of corresponding peptide

Figure 2

Phylogeny and synteny for the *GCGR* subfamily genes. (A) Maximum likelihood phylogenetic tree for the GCGR subfamily of human (hu), mouse (mo), chicken (ch), anole lizard (an), *Xenopus* (xe), zebrafish (zf), medaka (md), fugu (fu), stickleback (sb), tetraodon (to), and lamprey (lam) along with human SCTR, GHRHR, VIPRs, and ADCYAP1R1. The receptor sequences were aligned by using MUSCLE and a tree was constructed with MEGA 5.05. Bootstrap number indicates 100 replicates. (B) Synteny for human genome fragments having the *GCGR* subfamily genes. Chromosome numbers are indicated above the gene, and gene locations (megabase) are indicated below the gene. Paralogous genes are aligned on the same column with the same color. (C) Proposed evolutionary history of the *GCGR* subfamily genes in vertebrates. *GLP2R*, *GLP1R*, and *GCGR* have emerged by local duplication before 2R while *GIPR* and *GCRPR* arose by 2R in an osteichthyan ancestor. After divergence of teleosts and tetrapods, *GLP1R* disappeared and *GCGR* were doubled through teleost-specific 3R in teleosts. The absence of the gene in each species is indicated by white boxes with broken lines.

genes. In the peptide genes, GCG-, GLP1-, and GLP2-coding exons are duplicated in a single gene *GCG*. *GCG*, *GIP*, and *GCRP* arose through 2R (Irwin & Prentice 2011, Wang *et al.* 2012, Park *et al.* 2013). After divergence of teleosts and tetrapods, in teleosts *GLP1R* was lost and *GCGR* was doubled due to the teleost-specific third round (3R). There are also species-specific losses of genes. For instance, *GCRPR* in zebrafish and mammals and *GIPR* in medaka disappeared (Fig. 2C).

Emergence of GLP1

The majority of genes for SCT-like peptide ligands are located on *VAC D*. The *GHRH* and *CRH/UCN* ancestors are aligned on *VAC B* where *GLP1R* is located, while the *GCG* ancestor is on *VAC E* where most SCTR-like GPCRs are found. Because the ancestors of this family, the *UNC2/3*, *CALC*, and *PTH*, are on *VAC D*, it is postulated that all these peptide members have emerged on *VAC D* through local duplications. Subsequently, some members may have been translocated to *VAC B* or *E* (Hwang *et al.* 2013). The time of *GCG* ancestor emergence remains unclear. Although tunicates have two *GCGR*-like genes, conventional Blast search tools failed to identify any peptide ligand genes for SCTR-like receptors in chordates (Cardoso *et al.* 2010). This is likely due to considerable amino acid variation within a short sequence of the mature peptide (Cardoso *et al.* 2010, Hwang *et al.* 2013).

Because the lamprey has two *GCG*-like genes, the *GCG* ancestor is likely to have emerged before the divergence of agnathans and gnathostomes (Hwang *et al.* 2013).

With regard to the *GLP1* family peptides, *GCG* (GCG–GLP1–GLP2), *GIP*, and a novel *GCRP* have emerged through 2R (Irwin & Prentice 2011, Hwang *et al.* 2013, Park *et al.* 2013). GCRP was classified as a member of the exendin-4 family in a previous study (Irwin & Prentice 2011). However, a recent study using phylogenetic analysis and biochemical ligand–receptor activity assays has revealed that GCRP was distinct from exendin-4 as well as other GLP1-related peptides (Park *et al.* 2013). Synteny analysis of *GCG*-, *GIP*-, and *GCGR*-containing genome fragments reveals similar neighbor genes, flanking these peptide family genes (Irwin & Prentice 2011). In accordance with this observation, the *GCG*, *GIP*, and *GCGR* genes are located on *GAC E0*, *E2*, and *E3* respectively (Hwang *et al.* 2013, Park *et al.* 2013), suggesting that these three genes are ohnologous to one another. However, the fourth gene was lost before the divergence of tetrapods and teleosts.

In the teleost lineage, the 3R of genome duplication further contributed to emergence of the second form of

GCG and *GCGR*. However, *GLP1R* disappeared in this lineage. The pattern of gene loss in teleosts differs by species. For example, zebrafish have lost *GCRP* and *GCRPR*, while the medaka lacks *GIP* and *GIPR* (Hwang *et al.* 2013, Park *et al.* 2013). In the tetrapod lineage, mammals have lost *GCRP* and *GCRPR* (Fig. 2C). The absence of *GLP1R* in teleost fish, in spite of the presence of the GLP1 peptide, remains enigmatic. With regard to the concept of coevolution of the peptide and the receptor, the possibility remains that existing receptors recognize paralogous peptides in place of the missing orthologous partners. For example, fish GLP1 binds fish GCGRs with high affinity (Irwin & Wong 2005). Thus, it is possible that GLP1 has gained a GCG-like function in teleost fish (Nguyen *et al.* 1994).

Exon duplication events for GLP1 and related peptides

Because most vertebrate *GCG* genes encode GCG, GLP1, and GLP2 on consecutive exons, these three peptides are likely to have emerged through exon duplication events (Fig. 3A). However, the time of exon duplication for GLP1
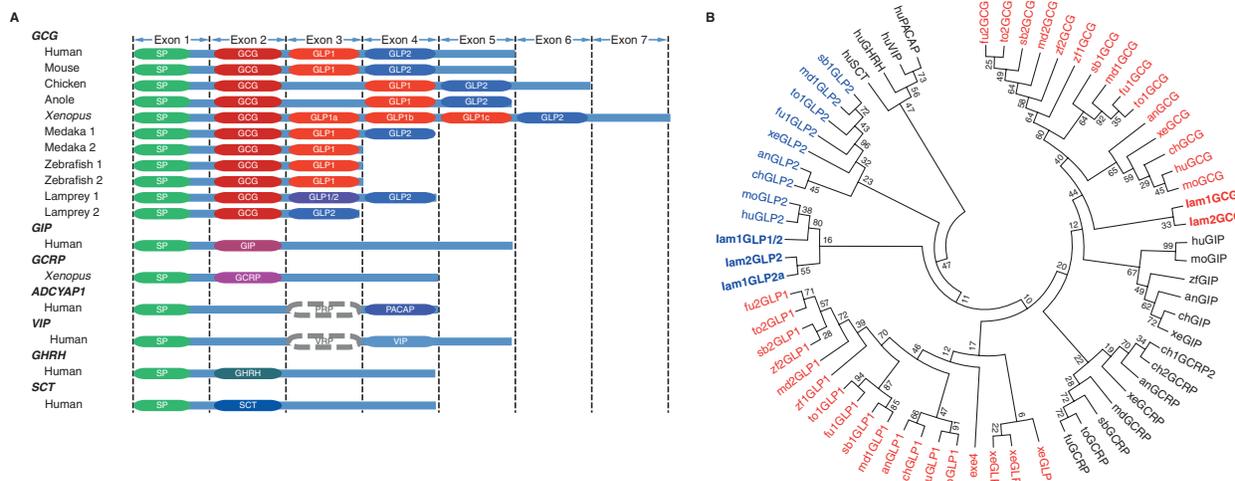

Figure 3

Phylogeny and exon structures of *GCG* and related peptide genes. (A) Exon structures of the *GCG* and related genes from various vertebrates and lamprey. In this diagram, as only the open reading frame-encoding exons are counted, the first exon contains the signal peptide (SP) sequence. The exons that do not encode the mature peptide sequence are also shown as simple lines. GCG, GLP1, and GLP2 are encoded by three different exons of the human *GCG* gene. The *Xenopus GCG* gene encodes three independent GLP1 peptides as well as GCG and GLP2. The second forms of medaka and zebrafish *GCG* genes lack GLP2. Location of the mature peptides on the exons of other peptide genes such as *GIP*, *GCRP*, *ADCYAP1*, *VIP*, *SCT*, and *GHRH* is also shown. Nonmammalian *ADCYAP1* and *VIP* encode

PACAP-related peptide (PRP) and VIP-related peptide (VRP), respectively, while those sequences in human genes are dysfunctional. The vestiges that remained in the human genes are indicated by spaces covered with gray dashed lines. (B) Maximum likelihood phylogenetic tree for GLP1-related peptides of human (hu), mouse (mo), chicken (ch), anole lizard (an), *Xenopus* (xe), zebrafish (zf), medaka (md), fugu (fu), stickleback (sb), tetraodon (to), and lamprey (lam) along with human SCT, GHRH, VIP, and PACAP are shown. The mature peptide sequences were aligned using MUSCLE, and a bootstrap consensus tree was constructed using MEGA 5.05. Bootstrap numbers represent 100 replicates.

is unclear (Sherwood *et al.* 2000). The basal vertebrate lamprey has two *GCG*-like genes (Irwin *et al.* 1999). One (*lam1GCG*) encodes GCG and two GLP2-like peptides (GLP1/2 and GLP2), and the other (*lam2GCG*) encodes GCG and GLP2. The phylogenetic position of lamGLP1/2 is currently unclear. A previous study using Clustal X-2.1 alignment followed by the neighbor-joining phylogenetic tree indicated that lam1GLP1/2 is more similar to GLP1 than to other related peptides (Hwang *et al.* 2013), while this study using multiple sequence alignment with high accuracy (MUSCLE) alignment with a maximum likelihood phylogenetic tree shows a close relationship of lam1GLP1/2 with mammalian GLP2 (Fig. 3B). Despite the confusion, this observation suggests that exon duplication events date back to early vertebrate evolution before the divergence of agnathans and gnathostomes. Owing to the short nucleotide sequences of scaffolds (GL481103 and GL477955) harboring the lamprey *GCGs*, it is still uncertain whether the two lamprey *GCGs* were generated by genome duplication or local gene duplication. However, as very close relationships between peptides encoded by *lam1GCG* and *lam2GCG* are evident from the phylogenetic tree (Fig. 3B), these two genes may have emerged recently, probably through local duplication. Interestingly, the *ADCYAP1* and *VIP* genes in vertebrates and basal vertebrate lamprey, which are believed to have emerged next to the *GCG* gene, also encode two mature peptides, and one (PACAP-related peptide (PRP) or VIP-related peptide (VRP)) of these two peptides is similar to GHRH in amino acid sequence (Laburthe *et al.* 1996, Sherwood *et al.* 2000, Ng *et al.* 2012, Hwang *et al.* 2013). Although the physiological roles of PRP and VRP remain incompletely understood, their interactions with specific receptors have been described in zebrafish and chicken (Wang *et al.* 2010). In addition, some vertebrate *GHRH* genes also encode PACAP- or VIP-like peptide sequences, although these peptides are likely nonfunctional. It is important to note that the PACAP, VIP, GHRH, and SCT peptide sequences are relatively more closely related to the GLP2 peptide when compared with other related peptides (Fig. 3B). This observation raises a possibility that the ancestor for *ADCYAP1*, *VIP*, *GHRH*, and *SCT* emerged through local gene duplication of the *GLP2*-containing ancestral gene before 2R, and that exon duplication events in this ancestor may influence the presence of the two peptides in the *ADCYAP1*, *VIP*, and *GHRH* genes (Fig. 3B). The exon duplication event before 2R was followed by additional exon duplication in *Xenopus* of the tetrapod lineage, or exon degeneration after 3R in the teleost lineage. The *GCG* gene of *Xenopus*, a representative amphibian, contains two

more exons encoding GLP1s. A copy of teleost-specific duplicated *GCGs* has lost the GLP2 sequence. In zebrafish, both *GCGs* lack the GLP2 sequence (Fig. 3A).

Specific amino acid sequence changes determine selective ligand–receptor interactions

GLP1R and its paralogous receptors exhibit high degrees of amino acid sequence identity with one another and share a relatively long N-terminal extracellular domain (ECD) beginning with an α-helix followed by four β-strands forming two antiparallel sheets that are stabilized by disulfide bonds and a salt bridge (Grace *et al.* 2004, Parthier *et al.* 2007, Runge *et al.* 2008, Underwood *et al.* 2010). Likewise, all peptide ligands for these receptors have typical three-dimensional structures to facilitate receptor binding. GLP1 and its family of peptides consist of 30–40 amino acids which share similarities in amino acid sequence and secondary structure. The structures of the peptides demonstrated by nuclear magnetic resonance and crystallography reveal that the N-terminal six or seven amino acid residues of peptides form a random coil structure where they share more than 70% amino acid sequence identity, and this is followed by an α-helical structure with moderate sequence similarity among the peptides (Thornton & Gorenstein 1994, Neidigh *et al.* 2001, Alana *et al.* 2007, Parthier *et al.* 2007, Underwood *et al.* 2010). Some residues with different lengths are extended after the α-helix to form flexible regions in the C-terminus, although this region is not necessary for receptor recognition and activation (Hinke *et al.* 2003, Moon *et al.* 2010). These peptides form a common local structure, called a helix N-capping motif, to stabilize the helical structure (Neumann *et al.* 2008). This motif in GLP1 is formed between hydrophobic amino acids Phe-6 and Val-10, and the polar amino acid Thr-7 through hydrophobic interaction and hydrogen bonding. These interactions are believed to facilitate receptor activation upon ligand–receptor binding (Fig. 4A). While the N-terminus and the first half of the central α-helical structures are responsible for interaction with the receptor core domain, the second half of the α-helical region binds directly to the ECD of the receptor (Parthier *et al.* 2007, Underwood *et al.* 2010, Moon *et al.* 2012*a*).

Interaction between peptide N-terminus and the receptor core domain

Although the GLP1 peptide and receptor families exhibit relatively high sequence similarities and conserved
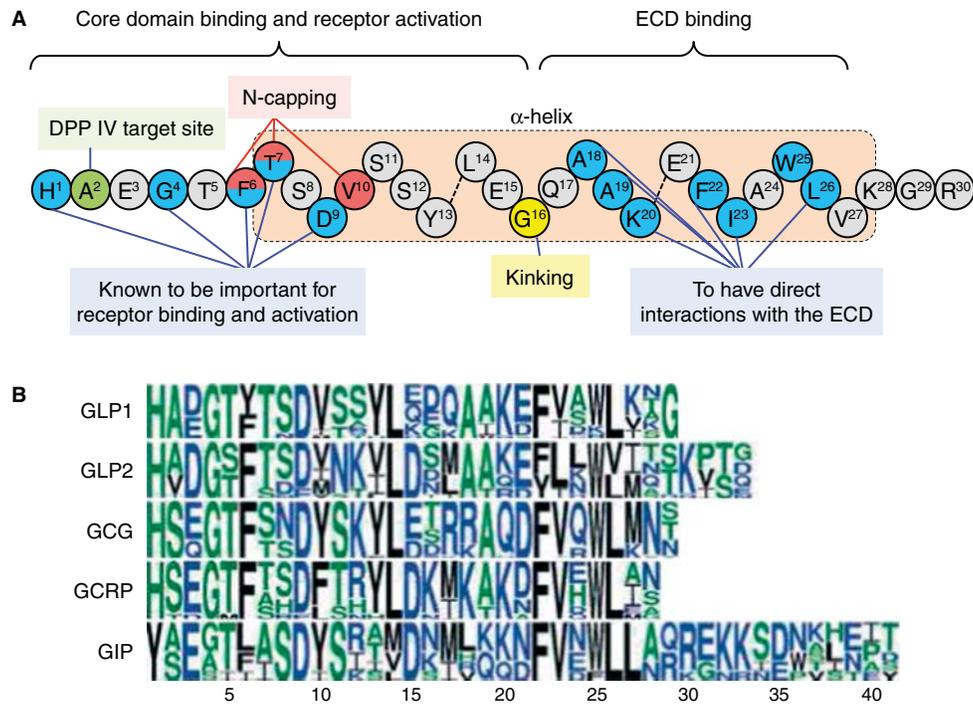
Figure 4

General structure of GLP1 and sequence alignment of GLP1 and related peptides. (A) General structure of GLP1 and functional residues. The biologically active GLP1 consists of a random coiled N-terminus with six residues followed by an α-helix starting from Thr-7 to Val-27. The residues, His-1, Gly-4, Phe-6, Thr-7, and Asp-9 are known to be important for receptor core domain binding and activation. Residues Phe-6, Thr-7, and Val-10 form a helix N-capping structure. The residues, Ala-18, Ala-19, Lys-20, Phe-22, Ile-23, and Leu-26 in the second half of the α-helix are demonstrated to have direct interaction with the N-terminal ECD of the receptor. Ala-2 is the target site of the dipeptidyl peptidase IV (DPPIV) protease and Gly-16 is involved in kinking of the α-helix in the ECD-bound structure. (B) Sequence logo of GLP1 and its related peptides. The amino acid sequences of the orthologous peptides were analyzed using a WebLogo program (http://weblogo.threeplusone.com/create.cgi). Sequences were retrieved from human, mouse, chicken, anole lizard, *Xenopus*, zebrafish, medaka, stickleback, tetraodon, and fugu.

structures (Sherwood *et al.* 2000), each peptide ligand selectively interacts with its own cognate receptor, indicating that this diversification is particularly important for selective interaction of the ligand–receptor pair, resulting in the specific physiological functions of each peptide-receptor pair (Moon *et al.* 2012*b*, Park *et al.* 2013).

In the N-terminal part of GPL1, residues His-1, Gly-4, Phe-6, Thr-7, and Asp-9 are crucial either for maintaining secondary structure of the peptide or for interaction with the receptor (Adelhorst *et al.* 1994, Gallwitz *et al.* 1994). Particularly, Gly-4, Phe-6, and Asp/Glu-9 are highly conserved for all GLP1 and related peptides (Fig. 4B), such that these residues in the related peptides are commonly important for recognition of their cognate receptors, particularly the core domain of the receptor (Perret *et al.* 2002, Hinke *et al.* 2003, Runge *et al.* 2003*a*, Yaqub *et al.* 2010). His-1 and Thr/Ser-7 are conserved in the GLP1 family of peptides (GLP1, GLP2, GCG, and GCRP) except for GIP that has Tyr-1 and Ile-7 in the corresponding sites (Fig. 4B), raising a possibility that

His-1 and Thr-7 of GLP1 are likely important for selective interaction with GLP1R, but discriminate GIPR. Indeed, introducing His-1 and Thr-7 into GIP allows this chimeric peptide to bind and activate GLP1R with a relatively high affinity and potency (Moon *et al.* 2010, 2012*a*). In contrast, Tyr-1 and Ile-7 in GIPR are important for selective interactions with its receptor. In addition to Tyr-1 and Ile-7, alanine scanning or amino acid modification studies reveal that highly conserved residues Glu-3, Gly-4, Asp-9, and Asp-15 are found to be important for receptor activation (Hinke *et al.* 2003, Yaqub *et al.* 2010). In GCG, besides the highly common residues Gly-4, Phe-6, The-7, and Asp/Glu-9 and GCG-specific residues Ser-2, Gln-3, Tyr-10, and Lys-12 (Fig. 4B) are found to be important for receptor binding and activation (Perret *et al.* 2002, Runge *et al.* 2003*a*). Interestingly, introducing these individual residues into the equivalent positions of GLP1 did not significantly alter peptide affinity and potency towards GLP1R. However, the replacement of these residues in GCG with those of GLP1 greatly reduced GCGR affinity

Published by Bioscientifica Ltd

(Runge et al. 2003a), indicating that positions 2, 3, 10, and 12 are only important for GCG. This result is consistent with the observation that wild-type (WT) or chimeric peptides having either the GCG or GLP1 N-terminus interact with receptors having the GLP1R core domain while WT or chimeric peptide containing the GLP1 N-terminus poorly binds or activates receptors containing the GCGR core domain (Runge et al. 2003b). The structure–activity relationships of GLP2 have not been extensively analyzed. However, an alanine scanning analysis revealed that residues at positions 2, 5, 6, and 17 were important for receptor activation (DaCambra et al. 2000).

A recent observation, after comparing amino acid sequences of the GLP1 peptide family including a novel GCRP, indicates that residues at positions 16–18 exhibited ortholog-specific amino acid changes, such that the Lys-16–Met/Ile-17–Lys-18, Ser/Thr-16–Arg-17–Arg-18, and Gly/Asp-16–Gln-17–Ala-18 motifs were conserved in GCRP, GCG, and GLP1 respectively (Park et al. 2013). Interestingly, placing the Gly/Asp-16–Gln-17–Ala-18 motif in GCRP increased peptide potency toward the WT GLP1R and GCRPR chimeric receptor that has the GLP1R core domain (Park et al. 2013). This observation raises the possibility that the motif at positions 16–18 may contribute to selective interactions with their cognate receptors. This motif resides at the border between residues that interact with the ECD and residues that contact the core domain. Particularly, Gly-16 of GLP1 contributes to the formation of a kink in the peptide helix (Underwood et al. 2010). Thus, this motif can be proposed to be involved in interaction between ECD and the core domain under ligand-bound conditions. A recent study has proposed that the ECD of GCGR is likely to negatively regulate GCGR activity through interaction with extracellular loop 3 (ECL3) of the core domain (Koth et al. 2012).

As ligand-bound crystal structures of the core domain of the GLP1R family are not available, the residues in the core domains responsible for ligand binding or receptor activation are poorly understood, although many approaches, such as alanine scanning, photoaffinity labeling, and molecular docking modeling have been attempted to identify these specific amino acid residues (Xiao et al. 2000, Lopez de Maturana & Donnelly 2002, Lopez de Maturana et al. 2004, Chen et al. 2009, 2010, Lin & Wang 2009). Mutation mapping studies suggest that charged amino acids (Lys-197, Asp-198, Lys-202, Asp-215, Arg-227, and Lys-288) and conserved residues (Met-204, Tyr-205, and Trp-306) are likely important for GLP1-induced receptor activation (Xiao et al. 2000, Lopez de Maturana & Donnelly 2002, Al-Sabah & Donnelly 2003,

Lopez de Maturana et al. 2004, Koole et al. 2012). However, these studies do not explain how individual residues of the peptide interact with residues in the receptor. Our recent study reveals that His-1 and Thr-7 were found to interact with Asn-302 at ECL2 and Ile-196 at the upper half of transmembrane helix 2 (TMH2) in GLP1R respectively (Moon et al. 2012a). Interestingly, Asn-302 is conserved for all vertebrate GLP1R, and at corresponding positions in GCGR (Asn-300 for human), GLP2R (Asn-336 for human), and GCRPR (Asn-301 for chicken), that are activated by peptide ligands containing His-1. GIPR, however, has hydrophobic Val-293 at this position (Moon et al. 2012b). Likewise, all receptors interacting with Thr-7-containing peptides have conserved hydrophobic Ile or Val at the corresponding position of Ile-196 in GLP1R, while GIPR has Ser-189 in this position. The study using photolabile probes of GLP1 provided only partial information on spatial approximation of interacting residues of the peptide and receptor (Chen et al. 2010). Molecular docking modeling with biochemical analyses has proposed the ligand-binding pockets in the GLP1R core domain (Lin & Wang 2009, Coopman et al. 2011, Kirkpatrick et al. 2012). However, these modeling approaches failed to provide a common consensus of the ligand-binding pocket such that ligand contact residues in the binding pocket are all different from one another. Thus, stricter biochemical analyses with an appropriate strategy may be helpful to construct a more reliable ligand-bound receptor model.

Interaction between the peptide C-terminus and the receptor ECD

Interactions between the peptide C-terminus and the receptor ECD are well defined for GLP1R and GIPR as demonstrated by ligand-bound ECD crystal structures (Parthier et al. 2007, Runge et al. 2008, Underwood et al. 2010). In addition, these crystal structures were useful in understanding interactions between GCG and the GCGR ECD and between GLP2 and the GLP2R ECD by applying molecular docking models (Venneti & Hewage 2011, Koth et al. 2012, Siu et al. 2013). The crystal structure of the ligand-bound ECD of the receptors shows that only residues in the second half of the α-helix (ranging from Ala-18 to Val-29 for GLP1) interact with the ECD (Underwood et al. 2010). The α-helices of all peptide hormones are partly amphipathic allowing hydrophobic and hydrophilic interactions with ECD of their receptor. The hydrophobic face (Ala-18, Ala-19, Phe-22, Ile-23, Trp-25, and Leu-26) of GLP1 is toward the hydrophobic binding cavity of the ECD that is formed by residues

Leu-32, Thr-35, Val-36, and Trp-39 in the N-terminal α-helix, residues Tyr-69 at the end of the β2 strand, and residues Try-88, Leu-89, Trp-91, and Pro-90 in the loop between the β3 and β4 strands (Runge *et al.* 2008, Underwood *et al.* 2010). Many of the residues in the binding pocket are highly conserved in the equivalent positions of GIPR, GLP2R, and GCGR and they are found to be important for peptide binding (Parthier *et al.* 2007, Venneti & Hewage 2011, Koth *et al.* 2012). Considering that residues Ala-19, Phe-22, Ile-23, and Leu-26 in the hydrophobic face of the peptides are also highly conserved, this result suggests that these evolutionarily conserved residues in the peptides and receptors are likely to contribute to primary binding between the α-helical domain of the peptides and the ECD of the receptors (Parthier *et al.* 2007, Underwood *et al.* 2010, Venneti & Hewage 2011, Koth *et al.* 2012). This primary interaction may explain cross-interactions of peptides with other receptors in the same family. For example, the GLP1 chimeric peptides having the α-helical region of GIP or GCG are able to activate GLP1R with a relatively high potency (Runge *et al.* 2003*a,b*, Moon *et al.* 2010). The hydrophilic face of GLP1 comprises residues Gln-17, Lys-20, Glu-21, and Lys-28. Of these residues, only Lys-20 was found to make a direct interaction with the ECD residue Glu-128 (Underwood *et al.* 2010). At positions 17–20, including the hydrophilic residues, GLP1 and its related peptides exhibited marked variations in amino acid sequence, raising a possibility that selective recognition of the peptide toward the cognate ECD can be achieved through interactions of these ortholog-specific residues of the α-helix with specific residues in the ECD of the cognate receptor (Parthier *et al.* 2007, Underwood *et al.* 2010).

Conclusions

GLP1 and GLP1R have emerged through 2R and local duplications before and after 2R. The exon duplication events also contributed to occurrence of the GLP1 peptide. Because the GLP1 peptide and receptor families arose from a common ancestor, they share similarities in their amino acid sequences and basic structures. Their conserved residues and structural similarities across the family members suggest a common mechanism underlying ligand–receptor binding followed by receptor activation. The comparison of ligand–receptor interaction mechanisms reveals that the conserved residues among GLP1 and related peptides tend to interact with conserved residues of GLP1R and its paralogous receptors (Moon *et al.* 2012*b*). Specific change after genome/gene/exon duplication may allow selective interaction of each ligand–

receptor pair, and could provide for diversified functions. Our recent studies have demonstrated that interactions between residues evolved from ortholog-specific peptides and receptors enhanced recognition of a peptide for the cognate receptor, while maintaining discrimination by other receptors (Moon *et al.* 2010, 2012*a*). During the evolutionary process, GLP1, through its specific interaction with GLP1R, acquired a particular function in glucose homeostasis and appetite regulation. Although GLP1 has gained great attention due to its beneficial effects on type 2 diabetes and obesity, molecular interaction with its receptor is not fully understood. Particularly, the specific amino acid residues within the TMH and ECL of receptors that may confer ligand binding and receptor activation remain poorly investigated. Understanding the evolutionary history of GLP1/GLP1R can provide clues to explore molecular interaction properties between GLP1 and GLP1R. Elucidation of the structure of the ligand-binding pocket in the core domain of GLP1R may facilitate development of potent peptide agonists or small molecules capable of activating GLP1R, which could greatly contribute to the treatment of diabetes and obesity.

References

Abi-Rached L, Gilles A, Shiina T, Pontarotti P & Inoko H 2002 Evidence of en bloc duplication in vertebrate genomes. *Nature Genetics* **31** 100–105. (doi:10.1038/ng855)

Acharjee S, Do-Rego JL, Oh DY, Ahn RS, Choe H, Vaudry H, Kim K, Seong JY & Kwon HB 2004 Identification of amino acid residues that direct differential ligand selectivity of mammalian and nonmammalian V1a type receptors for arginine vasopressin and vasotocin. Insights into molecular coevolution of V1a type receptors and their ligands. *Journal of Biological Chemistry* **279** 54445–54453. (doi:10.1074/jbc.M408909200)

Adelhorst K, Hedegaard BB, Knudsen LB & Kirk O 1994 Structure–activity studies of glucagon-like peptide-1. *Journal of Biological Chemistry* **269** 6275–6278.

Alana I, Malthouse JP, O'Harte FP & Hewage CM 2007 The bioactive conformation of glucose-dependent insulinotropic polypeptide by NMR and CD spectroscopy. *Proteins* **68** 92–99. (doi:10.1002/prot.21372)

Al-Sabah S & Donnelly D 2003 A model for receptor-peptide binding at the glucagon-like peptide-1 (GLP-1) receptor through the analysis of truncated ligands and receptors. *British Journal of Pharmacology* **140** 339–346. (doi:10.1038/sj.bjp.0705453)

Alvarez E, Martinez MD, Roncero I, Chowen JA, Garcia-Cuartero B, Gispert JD, Sanz C, Vazquez P, Maldonado A, de Caceres J et al. 2005 The expression of GLP-1 receptor mRNA and protein allows the effect of GLP-1 on glucose metabolism in the human hypothalamus and brainstem. *Journal of Neurochemistry* **92** 798–806. (doi:10.1111/j.1471-4159.2004.02914.x)

Baggio LL & Drucker DJ 2002 Harnessing the therapeutic potential of glucagon-like peptide-1: a critical review. *Treatments in Endocrinology* **1** 117–125. (doi:10.2165/00024677-200201020-00005)

Cardoso JC, Pinto VC, Vieira FA, Clark MS & Power DM 2006 Evolution of secretin family GPCR members in the metazoa. *BMC Evolutionary Biology* **6** 108. (doi:10.1186/1471-2148-6-108)

Cardoso JC, Vieira FA, Gomes AS & Power DM 2010 The serendipitous origin of chordate secretin peptide family members. *BMC Evolutionary Biology* **10** 135. (doi:10.1186/1471-2148-10-135)

Chen Q, Pinon DI, Miller LJ & Dong M 2009 Molecular basis of glucagon-like peptide 1 docking to its intact receptor studied with carboxyl-terminal photolabile probes. *Journal of Biological Chemistry* **284** 34135–34144. (doi:10.1074/jbc.M109.038109)

Chen Q, Pinon DI, Miller LJ & Dong M 2010 Spatial approximations between residues 6 and 12 in the amino-terminal region of glucagon-like peptide 1 and its receptor: a region critical for biological activity. *Journal of Biological Chemistry* **285** 24508–24518. (doi:10.1074/jbc.M110.135749)

Cho HJ, Acharjee S, Moon MJ, Oh DY, Vaudry H, Kwon HB & Seong JY 2007 Molecular evolution of neuropeptide receptors with regard to maintaining high affinity to their authentic ligands. *General and Comparative Endocrinology* **153** 98–107. (doi:10.1016/j.ygcen.2006.12.013)

Coopman K, Wallis R, Robb G, Brown AJ, Wilkinson GF, Timms D & Willars GB 2011 Residues within the transmembrane domain of the glucagon-like peptide-1 receptor involved in ligand binding and receptor activation: modelling the ligand-bound receptor. *Molecular Endocrinology* **25** 1804–1818. (doi:10.1210/me.2011-1160)

DaCambra MP, Yusta B, Sumner-Smith M, Crivici A, Drucker DJ & Brubaker PL 2000 Structural determinants for activity of glucagon-like peptide-2. *Biochemistry* **39** 8888–8894. (doi:10.1021/bi000497p)

Drucker DJ 2003 Glucagon-like peptides: regulators of cell proliferation, differentiation, and apoptosis. *Molecular Endocrinology* **17** 161–171. (doi:10.1210/me.2002-0306)

Drucker DJ, Philippe J, Mojsov S, Chick WL & Habener JF 1987 Glucagon-like peptide I stimulates insulin gene expression and increases cyclic AMP levels in a rat islet cell line. *PNAS* **84** 3434–3438. (doi:10.1073/pnas.84.10.3434)

During MJ, Cao L, Zuzga DS, Francis JS, Fitzsimons HL, Jiao X, Bland RJ, Klugmann M, Banks WA, Drucker DJ et al. 2003 Glucagon-like peptide-1 receptor is involved in learning and neuroprotection. *Nature Medicine* **9** 1173–1179. (doi:10.1038/nm919)

Fehmann HC & Habener JF 1992 Insulinotropic hormone glucagon-like peptide-I(7–37) stimulation of proinsulin gene expression and pro-insulin biosynthesis in insulinoma β TC-1 cells. *Endocrinology* **130** 159–166. (doi:10.1210/en.130.1.159)

Fehmann HC, Goke R & Goke B 1995 Cell and molecular biology of the incretin hormones glucagon-like peptide-I and glucose-dependent insulin releasing polypeptide. *Endocrine Reviews* **16** 390–410.

Fredriksson R, Lagerstrom MC, Lundin LG & Schioth HB 2003 The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Molecular Pharmacology* **63** 1256–1272. (doi:10.1124/mol.63.6.1256)

Gallwitz B, Witt M, Paetzold G, Morys-Wortmann C, Zimmermann B, Eckart K, Folsch UR & Schmidt WE 1994 Structure/activity characterization of glucagon-like peptide-1. *European Journal of Biochemistry* **225** 1151–1156. (doi:10.1111/j.1432-1033.1994.1151b.x)

Grace CR, Perrin MH, DiGruccio MR, Miller CL, Rivier JE, Vale WW & Riek R 2004 NMR structure and peptide hormone binding site of the first extracellular domain of a type B1 G protein-coupled receptor. *PNAS* **101** 12836–12841. (doi:10.1073/pnas.0404702101)

Graziano MP, Hey PJ, Borkowski D, Chicchi GG & Strader CD 1993 Cloning and functional expression of a human glucagon-like peptide-1 receptor. *Biochemical and Biophysical Research Communications* **196** 141–146. (doi:10.1006/bbrc.1993.2226)

Hansotia T & Drucker DJ 2005 GIP and GLP-1 as incretin hormones: lessons from single and double incretin receptor knockout mice. *Regulatory Peptides* **128** 125–134. (doi:10.1016/j.regpep.2004.07.019)

Harmar AJ 2001 Family-B G-protein-coupled receptors. *Genome Biology* **2** REVIEWS3013. (doi:10.1186/gb-2001-2-12-reviews3013)

Hewes RS & Taghert PH 2001 Neuropeptides and neuropeptide receptors in the *Drosophila melanogaster* genome. *Genome Research* **11** 1126–1142. (doi:10.1101/gr.169901)

Hinke SA, Gelling R, Manhart S, Lynn F, Pederson RA, Kuhn-Wache K, Rosche F, Demuth HU, Coy D & McIntosh CH 2003 Structure–activity relationships of glucose-dependent insulinotropic polypeptide (GIP). *Biological Chemistry* **384** 403–407. (doi:10.1515/BC.2003.046)

Hisadome K, Reimann F, Gribble FM & Trapp S 2010 Leptin directly depolarizes preproglucagon neurons in the nucleus tractus solitarius: electrical properties of glucagon-like peptide 1 neurons. *Diabetes* **59** 1890–1898. (doi:10.2337/db10-0128)

Holland PW 2003 More genes in vertebrates? *Journal of Structural and Functional Genomics* **3** 75–84. (doi:10.1023/A:1022656931587)

Hwang JI, Moon MJ, Park S, Kim DK, Cho EB, Ha N, Son GH, Kim K, Vaudry H & Seong JY 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. *Molecular Biology and Evolution* **30** 1119–1130. (doi:10.1093/molbev/mst031)

Irwin DM & Prentice KJ 2011 Incretin hormones and the expanding families of glucagon-like sequences and their receptors. *Diabetes, Obesity & Metabolism* **13** (Suppl 1) 69–81. (doi:10.1111/j.1463-1326.2011.01444.x)

Irwin DM & Wong K 2005 Evolution of new hormone function: loss and gain of a receptor. *Journal of Heredity* **96** 205–211. (doi:10.1093/jhered/esi024)

Irwin DM, Satkunarajah M, Wen Y, Brubaker PL, Pederson RA & Wheeler MB 1997 The *Xenopus* proglucagon gene encodes novel GLP-1-like peptides with insulinotropic properties. *PNAS* **94** 7915–7920. (doi:10.1073/pnas.94.15.7915)

Irwin DM, Huner O & Youson JH 1999 Lamprey proglucagon and the origin of glucagon-like peptides. *Molecular Biology and Evolution* **16** 1548–1557. (doi:10.1093/oxfordjournals.molbev.a026067)

Ishihara T, Nakamura S, Kaziro Y, Takahashi T, Takahashi K & Nagata S 1991 Molecular cloning and expression of a cDNA encoding the secretin receptor. *EMBO Journal* **10** 1635–1641.

Kang G, Chepurny OG & Holz GG 2001 cAMP-regulated guanine nucleotide exchange factor II (Epac2) mediates Ca^{2+}-induced Ca^{2+} release in INS-1 pancreatic β-cells. *Journal of Physiology* **536** 375–385. (doi:10.1111/j.1469-7793.2001.0375c.xd)

Kieffer TJ & Habener JF 1999 The glucagon-like peptides. *Endocrine Reviews* **20** 876–913. (doi:10.1210/er.20.6.876)

Kim DK, Cho EB, Moon MJ, Park S, Hwang JI, Kah O, Sower SA, Vaudry H & Seong JY 2011 Revisiting the evolution of gonadotropin-releasing hormones and their receptors in vertebrates: secrets hidden in genomes. *General and Comparative Endocrinology* **170** 68–78. (doi:10.1016/j.ygcen.2010.10.018)

Kim DK, Cho EB, Moon MJ, Park S, Hwang JI, Do Rego JL, Vaudry H & Seong JY 2012 Molecular coevolution of neuropeptides gonadotropin-releasing hormone and kisspeptin with their cognate G protein-coupled receptors. *Frontiers in Neuroscience* **6** 3. (doi:10.3389/fnins.2012.00003)

Kirkpatrick A, Heo J, Abrol R & Goddard WA III 2012 Predicted structure of agonist-bound glucagon-like peptide 1 receptor, a class B G protein-

coupled receptor. *PNAS* **109** 19988–19993. (doi:10.1073/pnas.1218051109)

Koole C, Wootten D, Simms J, Miller LJ, Christopoulos A & Sexton PM 2012 Second extracellular loop of human glucagon-like peptide-1 receptor (GLP-1R) has a critical role in GLP-1 peptide binding and receptor activation. *Journal of Biological Chemistry* **287** 3642–3658. (doi:10.1074/jbc.M111.309328)

Koth CM, Murray JM, Mukund S, Madjidi A, Minn A, Clarke HJ, Wong T, Chiang V, Luis E, Estevez A et al. 2012 Molecular basis for negative regulation of the glucagon receptor. *PNAS* **109** 14393–14398. (doi:10.1073/pnas.1206734109)

Laburthe M, Couvineau A, Gaudin P, Maoret JJ, Rouyer-Fessard C & Nicole P 1996 Receptors for VIP, PACAP, secretin, GRF, glucagon, GLP-1, and other members of their new family of G protein-linked receptors: structure–function relationship with special reference to the human VIP-1 receptor. *Annals of the New York Academy of Sciences* **805** 94–109 (discussion 110–111). (doi:10.1111/j.1749-6632.1996.tb17476.x)

Larhammar D, Lundin LG & Hallbook F 2002 The human Hox-bearing chromosome regions did arise by block or chromosome (or even genome) duplications. *Genome Research* **12** 1910–1920. (doi:10.1101/gr.445702)

Lee YR, Tsunekawa K, Moon MJ, Um HN, Hwang JI, Osugi T, Otaki N, Sunakawa Y, Kim K, Vaudry H et al. 2009 Molecular evolution of multiple forms of kisspeptins and GPR54 receptors in vertebrates. *Endocrinology* **150** 2837–2846. (doi:10.1210/en.2008-1679)

Li Y, Hansotia T, Yusta B, Ris F, Halban PA & Drucker DJ 2003 Glucagon-like peptide-1 receptor signaling modulates β cell apoptosis. *Journal of Biological Chemistry* **278** 471–478. (doi:10.1074/jbc.M209423200)

Li JH, Choe H, Wang AF, Maiti K, Wang C, Salam A, Chun SY, Lee WK, Kim K, Kwon HB et al. 2005 Extracellular loop 3 (EL3) and EL3-proximal transmembrane helix 7 of the mammalian type I and type II gonadotropin-releasing hormone (GnRH) receptors determine differential ligand selectivity to GnRH-I and GnRH-II. *Molecular Pharmacology* **67** 1099–1110. (doi:10.1124/mol.104.004887)

Lin F & Wang R 2009 Molecular modeling of the three-dimensional structure of GLP-1R and its interactions with several agonists. *Journal of Molecular Modeling* **15** 53–65. (doi:10.1007/s00894-008-0372-2)

Lopez de Maturana R & Donnelly D 2002 The glucagon-like peptide-1 receptor binding site for the N-terminus of GLP-1 requires polarity at Asp198 rather than negative charge. *FEBS Letters* **530** 244–248. (doi:10.1016/S0014-5793(02)03492-0)

Lopez de Maturana R, Treece-Birch J, Abidi F, Findlay JB & Donnelly D 2004 Met-204 and Tyr-205 are together important for binding GLP-1 receptor agonists but not their N-terminally truncated analogues. *Protein and Peptide Letters* **11** 15–22. (doi:10.2174/0929866043478491)

Mirabeau O & Joly JS 2013 Molecular evolution of peptidergic signaling systems in bilaterians. *PNAS* **110** E2028–E2037. (doi:10.1073/pnas.1219956110)

Mojsov S, Heinrich G, Wilson IB, Ravazzola M, Orci L & Habener JF 1986 Preproglucagon gene expression in pancreas and intestine diversifies at the level of post-translational processing. *Journal of Biological Chemistry* **261** 11880–11889.

Mojsov S, Kopczynski MG & Habener JF 1990 Both amidated and nonamidated forms of glucagon-like peptide I are synthesized in the rat intestine and the pancreas. *Journal of Biological Chemistry* **265** 8001–8008.

Moon MJ, Kim HY, Kim SG, Park J, Choi DS, Hwang JI & Seong JY 2010 Tyr1 and Ile7 of glucose-dependent insulinotropic polypeptide (GIP) confer differential ligand selectivity toward GIP and glucagon-like peptide-1 receptors. *Molecules and Cells* **30** 149–154. (doi:10.1007/s10059-010-0100-5)

Moon MJ, Kim HY, Park S, Kim DK, Cho EB, Hwang JI & Seong JY 2011 Insulin contributes to fine-tuning of the pancreatic β-cell response to glucagon-like peptide-1. *Molecules and Cells* **32** 389–395. (doi:10.1007/s10059-011-0157-9)

Moon MJ, Kim HY, Park S, Kim DK, Cho EB, Park CR, You DJ, Hwang JI, Kim K, Choe H et al. 2012a Evolutionarily conserved residues at glucagon-like peptide-1 (GLP-1) receptor core confer ligand-induced receptor activation. *Journal of Biological Chemistry* **287** 3873–3884. (doi:10.1074/jbc.M111.276808)

Moon MJ, Park S, Kim DK, Cho EB, Hwang JI, Vaudry H & Seong JY 2012b Structural and molecular conservation of glucagon-like peptide-1 and its receptor confers selective ligand–receptor interaction. *Frontiers in Endocrinology* **3** 141. (doi:10.3389/fendo.2012.00141)

Nakatani Y, Takeda H, Kohara Y & Morishita S 2007 Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome Research* **17** 1254–1265. (doi:10.1101/gr.6316407)

Nauck MA, Heimesaat MM, Behle K, Holst JJ, Nauck MS, Ritzel R, Hufner M & Schmiegel WH 2002 Effects of glucagon-like peptide 1 on counter-regulatory hormone responses, cognitive functions, and insulin secretion during hyperinsulinemic, stepped hypoglycemic clamp experiments in healthy volunteers. *Journal of Clinical Endocrinology and Metabolism* **87** 1239–1246. (doi:10.1210/jc.87.3.1239)

Neidigh JW, Fesinmeyer RM, Prickett KS & Andersen NH 2001 Exendin-4 and glucagon-like-peptide-1: NMR structural comparisons in the solution and micelle-associated states. *Biochemistry* **40** 13188–13200. (doi:10.1021/bi010902s)

Neumann JM, Couvineau A, Murail S, Lacapere JJ, Jamin N & Laburthe M 2008 Class-B GPCR activation: is ligand helix-capping the key? *Trends in Biochemical Sciences* **33** 314–319. (doi:10.1016/j.tibs.2008.05.001)

Ng SY, Chow BK, Kasamatsu J, Kasahara M & Lee LT 2012 Agnathan VIP, PACAP and their receptors: ancestral origins of today's highly diversified forms. *PLoS ONE* **7** e44691. (doi:10.1371/journal.pone.0044691)

Nguyen TM, Mommsen TP, Mims SM & Conlon JM 1994 Characterization of insulins and proglucagon-derived peptides from a phylogenetically ancient fish, the paddlefish (*Polyodon spathula*). *Biochemical Journal* **300** 339–345.

Nordstrom KJ, Fredriksson R & Schioth HB 2008 The amphioxus (*Branchiostoma floridae*) genome contains a highly diversified set of G protein-coupled receptors. *BMC Evolutionary Biology* **8** 9. (doi:10.1186/1471-2148-8-9)

Oh DY, Kim K, Kwon HB & Seong JY 2006 Cellular and molecular biology of orphan G protein-coupled receptors. *International Review of Cytology* **252** 163–218.

Park CR, Moon MJ, Park S, Kim D-K, Cho EB, Millar RP, Hwang J-I & Seong JY 2013 A novel glucagon-related peptide (GCRP) and its receptor GCRPR account for coevolution of their family members in vertebrates. *PLoS ONE* **8** e65420. (doi:10.1371/journal.pone.0065420)

Parthier C, Kleinschmidt M, Neumann P, Rudolph R, Manhart S, Schlenzig D, Fanghanel J, Rahfeld JU, Demuth HU & Stubbs MT 2007 Crystal structure of the incretin-bound extracellular domain of a G protein-coupled receptor. *PNAS* **104** 13942–13947. (doi:10.1073/pnas.0706404104)

Perret J, Van Craenenbroeck M, Langer I, Vertongen P, Gregoire F, Robberecht P & Waelbroeck M 2002 Mutational analysis of the glucagon receptor: similarities with the vasoactive intestinal peptide (VIP)/pituitary adenylate cyclase-activating peptide (PACAP)/secretin receptors for recognition of the ligand's third residue. *Biochemical Journal* **362** 389–394. (doi:10.1042/0264-6021:3620389)

Perry T, Holloway HW, Weerasuriya A, Mouton PR, Duffy K, Mattison JA & Greig NH 2007 Evidence of GLP-1-mediated neuroprotection in an animal model of pyridoxine-induced peripheral sensory neuropathy. *Experimental Neurology* **203** 293–301. (doi:10.1016/j.expneurol.2006.09.028)

Putnam NH, Butts T, Ferrier DE, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi M, Shoguchi E, Terry A, Yu JK et al. 2008 The amphioxus genome and the evolution of the chordate karyotype. *Nature* **453** 1064–1071. (doi:10.1038/nature06967)

Runge S, Gram C, Brauner-Osborne H, Madsen K, Knudsen LB & Wulff BS 2003*a* Three distinct epitopes on the extracellular face of the glucagon receptor determine specificity for the glucagon amino terminus. *Journal of Biological Chemistry* **278** 28005–28010. (doi:10.1074/jbc.M301085200)

Runge S, Wulff BS, Madsen K, Brauner-Osborne H & Knudsen LB 2003*b* Different domains of the glucagon and glucagon-like peptide-1 receptors provide the critical determinants of ligand selectivity. *British Journal of Pharmacology* **138** 787–794. (doi:10.1038/sj.bjp.0705120)

Runge S, Thogersen H, Madsen K, Lau J & Rudolph R 2008 Crystal structure of the ligand-bound glucagon-like peptide-1 receptor extracellular domain. *Journal of Biological Chemistry* **283** 11340–11347. (doi:10.1074/jbc.M708740200)

Santini S, Boore JL & Meyer A 2003 Evolutionary conservation of regulatory elements in vertebrate Hox gene clusters. *Genome Research* **13** 1111–1122. (doi:10.1101/gr.700503)

Schmidt WE, Siegel EG & Creutzfeldt W 1985 Glucagon-like peptide-1 but not glucagon-like peptide-2 stimulates insulin release from isolated rat pancreatic islets. *Diabetologia* **28** 704–707. (doi:10.1007/BF00291980)

Sherwood NM, Krueckl SL & McRory JE 2000 The origin and function of the pituitary adenylate cyclase-activating polypeptide (PACAP)/glucagon superfamily. *Endocrine Reviews* **21** 619–670. (doi:10.1210/er.21.6.619)

Siu FY, He M, de Graaf C, Han GW, Yang D, Zhang Z, Zhou C, Xu Q, Wacker D, Joseph JS *et al.* 2013 Structure of the human glucagon class B G-protein-coupled receptor. *Nature* **499** 444–449. (doi:10.1038/nature12393)

Thornton K & Gorenstein DG 1994 Structure of glucagon-like peptide (7–36) amide in a dodecylphosphocholine micelle as determined by 2D NMR. *Biochemistry* **33** 3532–3539. (doi:10.1021/bi00178a009)

Underwood CR, Garibay P, Knudsen LB, Hastrup S, Peters GH, Rudolph R & Reedtz-Runge S 2010 Crystal structure of glucagon-like peptide-1 in complex with the extracellular domain of the glucagon-like peptide-1 receptor. *Journal of Biological Chemistry* **285** 723–730. (doi:10.1074/jbc.M109.033829)

Venneti KC & Hewage CM 2011 Conformational and molecular interaction studies of glucagon-like peptide-2 with its N-terminal extracellular receptor domain. *FEBS Letters* **585** 346–352. (doi:10.1016/j.febslet.2010.12.011)

Vienne A, Rasmussen J, Abi-Rached L, Pontarotti P & Gilles A 2003 Systematic phylogenomic evidence of en bloc duplication of the ancestral 8p11.21–8p21.3-like region. *Molecular Biology and Evolution* **20** 1290–1298. (doi:10.1093/molbev/msg127)

Vrang N, Hansen M, Larsen PJ & Tang-Christensen M 2007 Characterization of brainstem preproglucagon projections to the paraventricular and dorsomedial hypothalamic nuclei. *Brain Research* **1149** 118–126. (doi:10.1016/j.brainres.2007.02.043)

Wang C, Yun O, Maiti K, Oh DY, Kim KK, Chae CH, Lee CJ, Seong JY & Kwon HB 2004 Position of Pro and Ser near Glu7.32 in the extracellular loop 3 of mammalian and nonmammalian gonadotropin-releasing hormone (GnRH) receptors is a critical determinant for differential ligand selectivity for mammalian GnRH and chicken GnRH-II. *Molecular Endocrinology* **18** 105–116. (doi:10.1210/me.2003-0101)

Wang Y, Li J, Wang CY, Kwok AY, Zhang X & Leung FC 2010 Characterization of the receptors for chicken GHRH and GHRH-related peptides: identification of a novel receptor for GHRH and the receptor for GHRH-LP (PRP). *Domestic Animal Endocrinology* **38** 13–31. (doi:10.1016/j.domaniend.2009.07.008)

Wang Y, Meng F, Zhong Y, Huang G & Li J 2012 Discovery of a novel glucagon-like peptide (GCGL) and its receptor (GCGLR) in chickens: evidence for the existence of GCGL and GCGLR genes in non-mammalian vertebrates. *Endocrinology* **153** 5247–5260. (doi:10.1210/en.2012-1586)

Wei Y & Mojsov S 1995 Tissue-specific expression of the human receptor for glucagon-like peptide-I: brain, heart and pancreatic forms have the same deduced amino acid sequences. *FEBS Letters* **358** 219–224. (doi:10.1016/0014-5793(94)01430-9)

Wheeler MB, Lu M, Dillon JS, Leng XH, Chen C & Boyd AE III 1993 Functional expression of the rat glucagon-like peptide-I receptor, evidence for coupling to both adenylyl cyclase and phospholipase-C. *Endocrinology* **133** 57–62. (doi:10.1210/en.133.1.57)

Xiao Q, Jeng W & Wheeler MB 2000 Characterization of glucagon-like peptide-1 receptor-binding determinants. *Journal of Molecular Endocrinology* **25** 321–335. (doi:10.1677/jme.0.0250321)

Xu G, Stoffers DA, Habener JF & Bonner-Weir S 1999 Exendin-4 stimulates both β-cell replication and neogenesis, resulting in increased β-cell mass and improved glucose tolerance in diabetic rats. *Diabetes* **48** 2270–2276. (doi:10.2337/diabetes.48.12.2270)

Yaqub T, Tikhonova IG, Lattig J, Magnan R, Laval M, Escrieut C, Boulegue C, Hewage C & Fourmy D 2010 Identification of determinants of glucose-dependent insulinotropic polypeptide receptor that interact with N-terminal biologically active region of the natural ligand. *Molecular Pharmacology* **77** 547–558. (doi:10.1124/mol.109.060111)

Yegorov S & Good S 2012 Using paleogenomics to study the evolution of gene families: origin and duplication history of the relaxin family hormones and their receptors. *PLoS ONE* **7** e32923. (doi:10.1371/journal.pone.0032923)

Zander M, Madsbad S, Madsen JL & Holst JJ 2002 Effect of 6-week course of glucagon-like peptide 1 on glycaemic control, insulin sensitivity, and β-cell function in type 2 diabetes: a parallel-group study. *Lancet* **359** 824–830. (doi:10.1016/S0140-6736(02)07952-7)

Journal of Molecular Endocrinology