GPR15–C10ORF99 functional pairing initiates colonic Treg homing in amniotes

Jingjing Song[1,2,†], Huaping Zheng[1,†], Jingwen Xue[2], Jian Liu[2], Qian Sun[1], Wei Yang[2], Fang Liu[2], Xiangyin Xiang[1], Kai He[3], Younan Chen[1], Jingqiu Cheng[1], Wei Li[1], Jin Jin[4], Juergen Brosius[1] & Cheng Deng[1,2,*]

Abstract

Regulatory T lymphocyte (Treg) homing reactions mediated by G protein-coupled receptor (GPCR)–ligand interactions play a central role in maintaining intestinal immune homeostasis by restraining inappropriate immune responses in the gastrointestinal tract. However, the origin of Treg homing to the colon remains mysterious. Here, we report that the C10ORF99 peptide (also known as CPR15L and AP57), a cognate ligand of GPR15 that controls Treg homing to the colon, originates from a duplication of the flanking CDHR1 gene and is functionally paired with GPR15 in amniotes. Evolutionary analysis and experimental data indicate that the GPR15–C10ORF99 pair is functionally conserved to mediate colonic Treg homing in amniotes and their expression patterns are positively correlated with herbivore diet in the colon. With the first herbivorous diet in early amniotes, a new biological process (herbivorous diet short-chain fatty acid-C10ORF99/GPR15-induced Treg homing colon immune homeostasis) emerged, and we propose an evolutionary model whereby GPR15–C10ORF99 functional pairing has initiated the first colonic Treg homing reaction in amniotes. Our findings also highlight that GPCR–ligand pairing leads to physiological adaptation during vertebrate evolution.

Introduction

G protein-coupled receptors (GPCRs) comprise the largest membrane protein family (with > 800 members encoded in vertebrate genomes) and play normal physiological roles in cell proliferation, survival and motility, and are associated with multiple diseases (Kroeze, 2003). Dysfunctional GPCRs contribute to many human diseases and are targeted, directly or indirectly, by 50–60% of all therapeutic agents (Pierce *et al*, 2002; Kroeze, 2003). Chemokines are small secreted proteins that direct the systemic trafficking and microenvironmental homing of various cell types in health and disease (Kroeze, 2003; Zlotnik & Yoshie, 2012; Zabel *et al*, 2015). Chemokines signal through class-A GPCRs and participate in homeostatic and inflammation-induced immune cell trafficking (Zlotnik & Yoshie, 2012; Zabel *et al*, 2015). Particularly, chemokines and their cognate GPCRs control the recruitment of various types of lymphocytes from the blood, contributing to the systemic organization of the immune system (Kim *et al*, 2013; Ocon *et al*, 2017). For example, the homing of CCR9-positive regulatory T lymphocytes (Tregs) to the small intestine is activated by CCL25; Treg homing to CXCR4-positive bone marrow cells is activated by CXCL12; and CCR4 activated by CCL17 is associated with Treg homing in the heart (Ding *et al*, 2012; Castan *et al*, 2017).

GPR15 has been considered an orphan GPCR and an HIV co-receptor with sequence similarity to leucocyte-chemoattractant receptors (Adamczyk, 2017; Suply *et al*, 2017). Recently, GPR15 was highlighted as a colon-specific Treg-homing receptor (Kim *et al*, 2013; Nguyen *et al*, 2014). A GPR15–GFP knock-in model revealed selective GPR15 expression by colon Tregs under homeostatic conditions and GPR15-mediated Treg recruitment to the colon (Kim *et al*, 2013). CCR9 functions similarly by promoting Treg recruitment to the small intestine (Johnson *et al*, 2010; Perrigoue *et al*, 2014). GPR15L (encoded by human C10ORF99 and mouse 2610528A11Rik) was identified as the cognate ligand of GPR15 (Yang *et al*, 2015; Suply *et al*, 2017). C10ORF99 is mainly expressed by epithelial cells in the gastrointestinal tract (particularly the colon and rectum) and attracts Treg subsets in a GPR15-dependent manner (Kim *et al*, 2013). Moreover, colonic C10ORF99 expression is developmentally determined and affected by inflammation and the microbiota (Ocon

et al, 2017). Therefore, GPR15–C10ORF99 signalling is important for recruiting specialized T lymphocytes to the colon and sites of cutaneous inflammation (Ocon et al, 2017; Suply et al, 2017), and the colon exhibits a different immune regulation than the small intestine.

Animal microbiota in the gastrointestinal tract have coevolved with the host (Mazmanian et al, 2005; Ley et al, 2008; Kamada et al, 2012), and their coexistence reflects an equilibrium established with the host immune system (Hooper et al, 2012; Kim et al, 2013). The colon harbours significantly more microbiota than the small intestine (Geuking Markus et al, 2011; McGuckin et al, 2011; Kuhn & Stappenbeck, 2013) and higher Treg frequencies (Atarashi et al, 2016). Humans and mice rely on colonic bacteria to break down undigestible dietary components such as fibres (Remely et al, 2013). Short-chain fatty acids (SCFAs) are bacterial fermentation products with variable concentrations (50–100 mM) in the colonic lumen (Cummings et al, 1987). In mice, SCFAs and the SCFA receptor (GPR43) may help induce colonic Treg homing (Perrigoue et al, 2014). Tregs utilize GPR15 (inducible by SCFAs) for colonic homing, and SCFAs and GPR43 are required to generate and expand colonic Tregs (Perrigoue et al, 2014). Moreover, SCFAs induce GPR15 expression, hence contributing to colonic Treg homing (Perrigoue et al, 2014). Interestingly, an herbivorous diet increases SCFA contents and microbial metabolites in the colon (Perrigoue et al, 2014). These data suggest that an herbivorous diet promotes SCFA- and GPR15-dependent colonic Treg homing in humans and mice.

GPCR–ligand pairs typically show high binding affinities and conserved functions in vertebrates and invertebrates (Strotmann et al, 2011; Vogel et al, 2013). Consequently, polypeptide ligands (i.e. apelin, angiotensin, cholecystokinin) from different vertebrates (even humans and fishes) can cross-stimulate GPCRs, and these orthologous receptor–ligand pairs control similar physiological functions across all vertebrates (de Mendoza et al, 2014; Hu et al, 2017). The small intestine is conserved across all vertebrates, and receptor–ligand (CCR9–CCL25)-dependent Treg homing is conserved to maintain physiological functions in vertebrates (Devries et al, 2006; Perrigoue et al, 2014). However, the emergence of the colon in tetrapods and the origination of colonic Treg homing remain mysterious. Here, we studied the origin and functional evolution of the GPR15–C1ORF99 pairing in vertebrates to gain insight into early colonic immune function in tetrapods.

Results

GPR15 evolved differently in fishes and amniotes

Vertebrate amino acid sequences of the GPR15, GPR25, apelin receptor (APJ), angiotensin (AGT) receptor 1 (AGTR1) and bradykinin receptor B2 (BDKRB2) were downloaded from the NCBI and Ensemble databases. A consensus neighbour-joining tree was built for the apelin/angiotensin/bradykinin receptor family (including mammals, birds, reptiles, amphibians and fishes), using MEGA 7.0.26 (JTT + G + I; bootstraps, 500; cut-off for the condensed tree, 20%) (Fig 1A). An ancestral gene, AGTR-like, from Ciona intestinalis served as an outgroup (Fournier et al, 2012). A similar tree was derived using maximum likelihood analysis, implemented in MEGA 7.0.26 (Fig EV1A). Only GPR25 was still considered as an orphan GPCR; bradykinin (KNG1), AGT, apelin and C10ORF99 are endogenous peptide

ligands for BDKRB2, AGTR1, APJ and GPR15 respectively (Fig 1A; Yang et al, 2015; Suply et al, 2017). GPR25, AGTR1 and BDKRB2 were highly conserved among all vertebrates tested (Fig 1A and B). However, GPR15 orthologues were detected in mammals, birds, three kinds of fishes and reptiles, but was presumably deleted in amphibians and most fishes. Regarding fish GPR15, Latimeria chalumnae belongs to the Sarcopterygii clade, and Lepisosteus oculatus and Scleropages formosus belong to the Actinopterygii clade, which did not undergo a third round of whole-genome duplication (Bian et al, 2016). GPR15 was not detected in other fishes, particularly those who underwent a third round of whole-genome duplication.

After removing gaps, indels and stop codons from the alignment, we obtained 981-base pair sequences for the GPR15 open-reading frame (ORF) sequences. We used the same orthologous genes and tree topology to identify rapidly evolving genes (REGs) in the GPR15 receptor family. The fish GPR15 genes were found to evolve rapidly (Fig 1C). The same REGs were identified in other receptors in the same family (Fig EV1B). No other GPCRs from the same family underwent rapid evolution (Fig EV1B). REGs can evolve neutrally or under positive selection. A branch-site model was utilized to assess whether GPR15 homologues underwent positive selection in different species. The GPR15 sequences of two lineages of fishes (Latimeria chalumnae and Lepisosteus oculatus/Scleropages formosus) had a large nonsynonymous (dN)/synonymous (dS) substitution rate ratio (branch site dN/dS of $\omega \gg 1$; Fig 1A and D), which was highly significant (likelihood ratio tests [LRTs], $P < 0.05$; Fig 1D). No other branches exhibited this ratio (Fig 1A and D). Residues that underwent positive selection differ in both fish lineages (Fig 1D, Dataset EV1). The same branch-site model was utilized to detect whether other receptors in the same family underwent positive selection as well (Fig EV1C). No other GPCRs from the same family underwent positive selection (Fig EV1C).

The C10ORF99 ligand originated from CDHR1 gene duplication in amniotes

To determine the origination of the GPR15–C10ORF99 receptor–ligand relationship, the gene synteny of vertebrate chromosomes containing these genes was analysed (Fig 2A). GPR15 exists in various vertebrate species (but not amphibians and most fishes) and shows conservation with its flanking gene. GPR15 localizes closely with CLDND1 on the same chromosomes of all vertebrates, although the corresponding amphibian and most fish chromosomes lack both GPR15 and C10ORF99 (Fig 2A). Similarly, the gene encoding the C10ORF99 ligand, and the flanking genes CDHR1, GHITM, LRIT2 and LRIT1, are highly conserved in amniotes, but C10ORF99 is absent from both fishes and amphibians (Fig 2A). BLAST searches detected the C10ORF99 gene among several non-protein-coding RNAs of reptiles and birds, and these sequences were conserved, particularly in regions encoding mature peptides and both cystine linkages of C10ORF99 (Dataset EV2–EV4). Interestingly, the signal peptide of CDHR1 in amniotes, but not of other vertebrates (fishes and amphibians), share similarity with the signal peptide encoded by C10ORF99 (an upstream gene; (Fig 2A–C and Dataset EV5 and EV6). Moreover, aligning C10ORF99 and CDHR1 of Chelonia mydas showed that CDHR1 genes shared large nucleotide similarities with C10ORF99, including 56% similarity between their 5′-untranslated regions (UTRs); 61% similarity between the protein coding region of
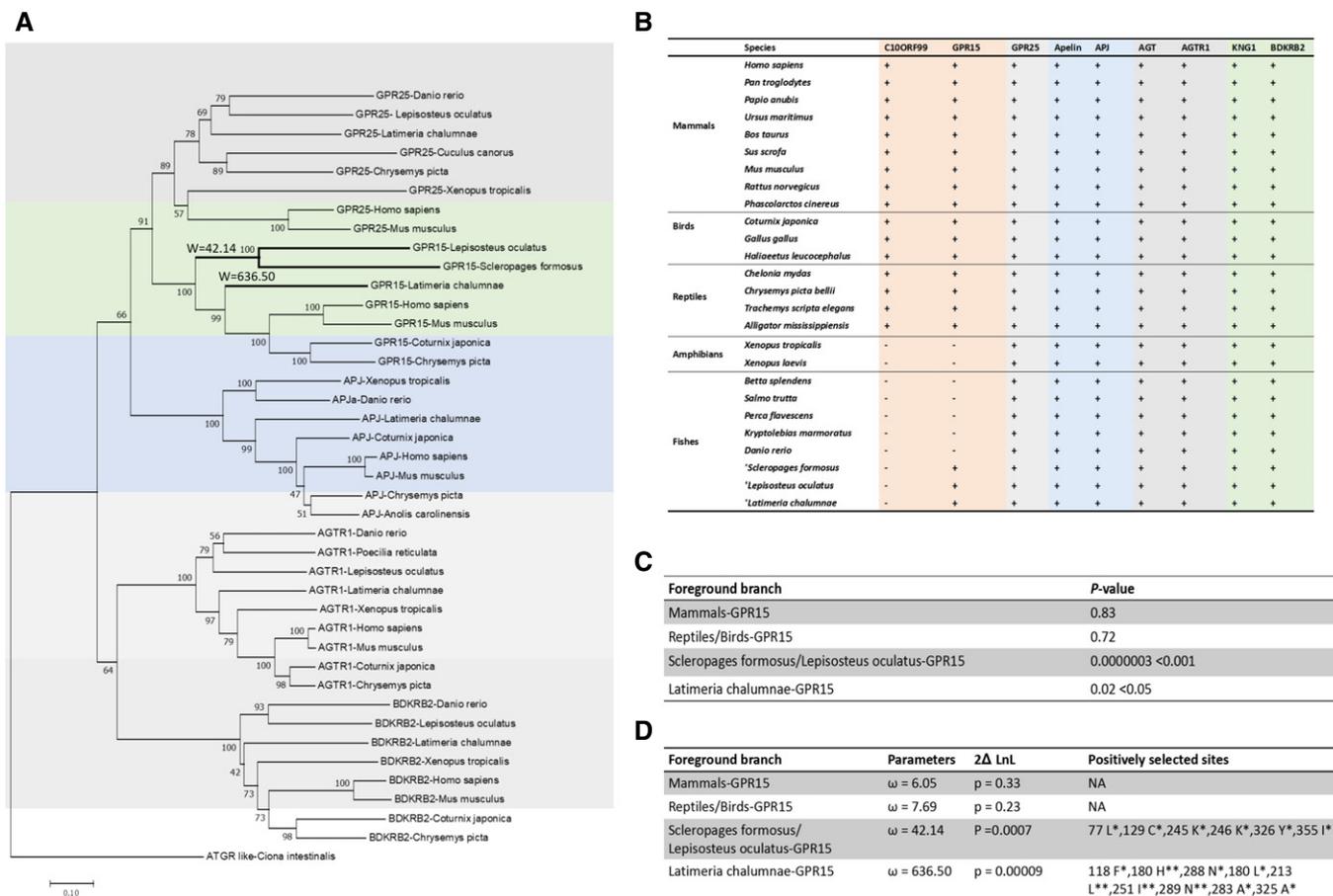
Figure 1. Evolution of the apelin/angiotensin/bradykinin peptide receptor subfamily.

A Phylogenetic tree of the apelin, angiotensin and bradykinin receptor, Scale (0.1). The *Ciona intestinalis* AGTR-like gene served as an outgroup. The ω parameters of two fish branches are shown. Scale: The unit of evolutionary distance is the number of amino acid substitutions at each position.

B Conservation of the apelin/angiotensin/bradykinin peptide receptor subfamily among all vertebrates except for the GPR15–C10ORF99 pair. *Indicates that these three fish species did not undergo a third round of whole-genome duplication during fish evolution.

C Result of rapidly evolving genes (REGs). This calculates the χ^2 critical value and *P*-value for conducting the likelihood ratio test. The genes with $P < 0.05$ were considered.

D GPR15 underwent positive selection in fishes. The parameters for each branch are shown; the dN/dS ratio was calculated with the whole protein-coding region. **$P < 0.01$, *$P < 0.05$, NA: not available.

exon 1, which encodes the signal peptide for both C10ORF99 and CDHR1; 59% similarity between intron 1; 81% similarity between CDHR1 intron 6 and C10ORF99 intron 1, 45% similarity between CDHR1 exon 7 and C10ORF99 exon 2, 80% similarity between CDHR1 intron 8 and C10ORF99 3′-UTRs, 93% similarity between CDHR1 intron 11 and C10ORF99 3′-UTRs respectively (Fig 2C, Dataset EV5). But C10ORF99 shared no significant similarity with CDHR1 of fishes or amphibians (Fig 2C). The CDHR1 genes also shared high similarity (5′-UTR, the protein coding region of exon1 and intron 1) with C10ORF99 in other reptiles, birds and mammals including humans, based on alignments with sequences from *Homo sapiens, Mus musculus, Phascolarctos cinereus, Coturnix japonica, Nothoprocta perdicaria, Dromaius novaehollandiae, Alligator mississippiensis, Pelodiscus sinensis* and *Chelonia mydas* (Fig 2B and Dataset EV6). The protein coding region of exon 1 of CDHR1 and C10ORF99, encoding their signal peptide, shared relatively high conservation of protein sequences and nucleic acid sequences (Fig 2B and C, Dataset EV5), and CDHR1 and C10ORF99 also had similar 5′-UTRs, which was consistent with the high colonic expression of both CDHR1 (expression data from NCBI database) and C10ORF99. In contrast, exon 2 of C10ORF99 exhibits a different ORF compared with CDHR1 exon 7, although they share similarity in nucleic acid sequences (Dataset EV5). Therefore, we conclude that C10ORF99 originated from reptile CDHR1 gene. This process encompassed exon deletions, co-option of a novel exon (3) and coding sequence changes by frame-shift yielding a new gene product, namely the C10ORF99 mature peptide (encoded by exon 2 and exon 3, Fig 2C).

GPR15–C10ORF99 functional pairing in amniotes

Five prototypical vertebrate GPR15 homologues were selected for functional pairing experiments, including human and mouse GPR15
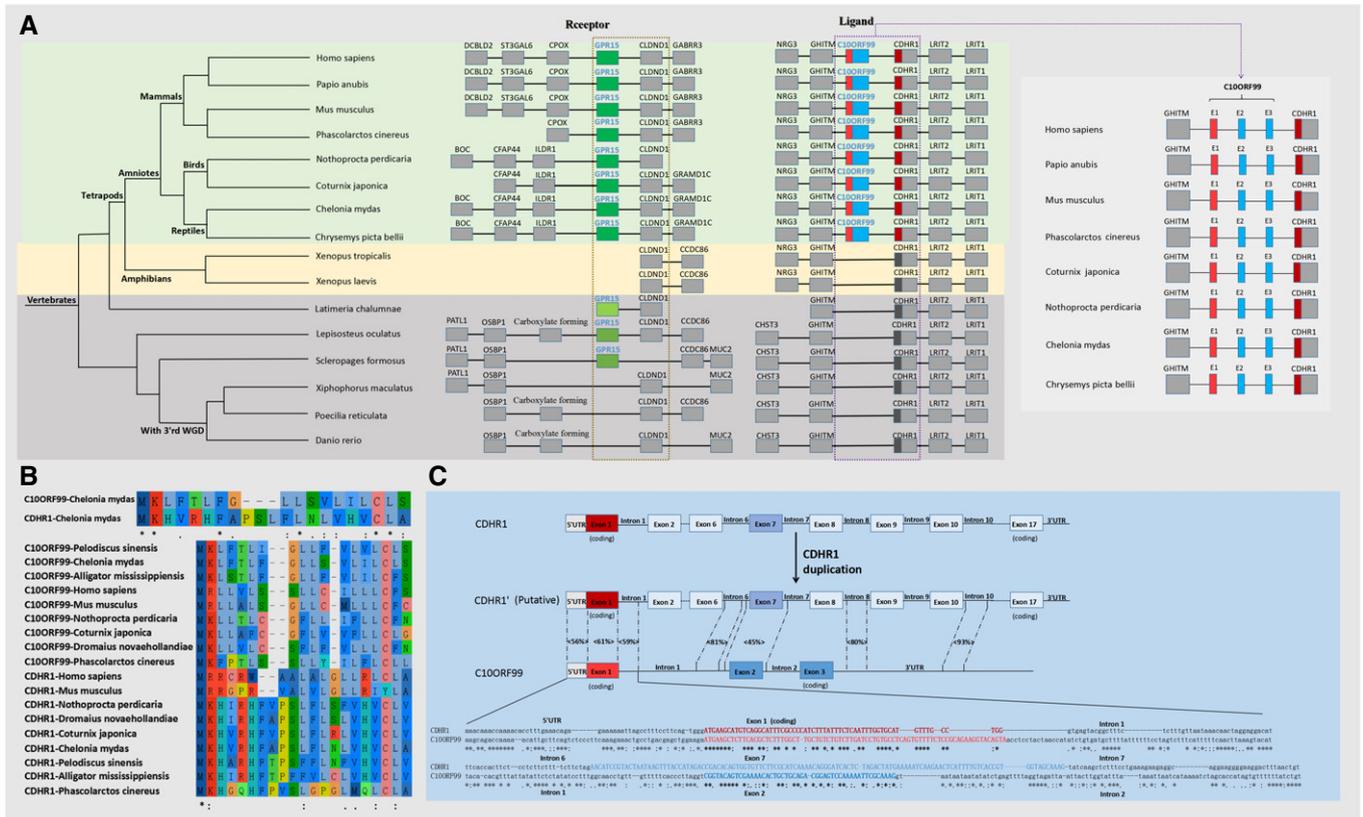
Figure 2. The C10ORF99 ligand originated in amniotes via CDHR1 gene duplication.

A   Synteny for chromosomal regions containing the GPR15–C10ORF99 pair. The pair and their neighbouring genes in vertebrate genome fragments are shown. Deep red shading represents exon 1 of the CDHR1 gene of amniotes. Deep grey colour represents exon 1 of the CDHR1 gene of amphibians and fishes. Light red represents exon 1 of the C10ORF99 gene. 3'rd WGD indicates the fishes that underwent a third round of whole-genome duplication.

B   Amino acid sequence alignment of C10ORF99 and CDHR1 signal peptides.

C   Molecular evolution of C10ORF99 from CDHR1. Regions in CDHR1 corresponding to regions in the C10ORF99 gene and nucleotide sequence similarities are given. The locations of the protein-coding regions of exons and as well as some intron locations are marked, and protein-coding regions of exon sequences are bold and capitalized. Coding means protein coding region in exon.

(hGPR15 and mGPR15 respectively), Japanese quail GPR15 (qGPR15), painted turtle GPR15 (tGPR15), *Latimeria chalumnae*/coelacanth GPR15 (cGPR15) and *Scleropages formosus*/Asian bony tongue GPR15 (aGPR15). Stimulation of intracellular calcium by GPR15 receptors requires ligand binding to stimulate Gq signalling (Yang *et al*, 2015; Suply *et al*, 2017). Human C10ORF99 (hC10ORF99) and Japanese quail C10ORF99 (qC10ORF99), featuring conserved mature peptides of C10ORF99 in mammals and reptiles/birds, respectively (Dataset EV2), were synthesized for receptor–ligand function assays (Suply *et al*, 2017; Fig 3A–G). Both hC10ORF99 and qC10ORF99 can interchange to stimulate hGPR15, mGPR15, qGPR15 and tGPR15 (Fig EV2A). The hC10ORF99 was utilized to activate cells transfected with the empty vector plasmid pcDNA3.1-V5-His (Fig 3A) or the pcDNA3.1-V5-His containing a gene encoding one of the mammalian GPR15, and qC10ORF99 was utilized to activate cells transfected with the pcDNA3.1-V5-His containing a gene encoding one of the reptile/bird/fish GPR15. The neurotensin (NTS) peptide was used to stimulate the human neurotensin receptor 1 (NTSR1) as a positive control for Gq-signalling activation (Fig 3A; Slosky *et al*, 2020). A remarkable increase in

Ca^{2+} ions was observed in cells expressing hGPR15 (highest stimulation, 10 nM), mGPR15 (highest stimulation, 10 nM), tGPR15 (highest stimulation, 100 nM) and qGPR15 (highest stimulation, 10 nM) (Fig 3B–E), but not cGPR15 or aGPR15 (Fig 3F and G), when stimulating with qC10ORF99 (even with high dose-1 μM). YM.254890, a well-known inhibitor of the Gq-signalling pathway, was used to confirm GPR15 receptor–ligand stimulation (Schrage *et al*, 2015; Tang *et al*, 2019). The Gq-signalling pathway was stimulated by hGPR15, mGPR15, tGPR15 and qGPR15, but inhibited by YM.254890 (Fig 3H).

Several point mutations of fish GPR15 underwent positive selection (Fig 1D). Among them, cGPR15 has several point mutations (F118G, S224T, H180L, N288L, A325Y and N289R) in the *Latimeria chalumnae* lineage and aGPR15 has several point mutations (T326Y, A245K, N77L, S355I and V129C) in the *Lepisosteus oculatus/Scleropages formosus* lineage that exhibited no stimulation after induction with the C10ORF99 ligand (Fig EV2 and sequence alignment for point mutants in Dataset EV1). Only the cGPR15-A283L from reverse mutants of *Latimeria chalumnae* and the cGPR15-P246K from reverse mutants of *Lepisosteus oculatus/Scleropages formosus* were activated
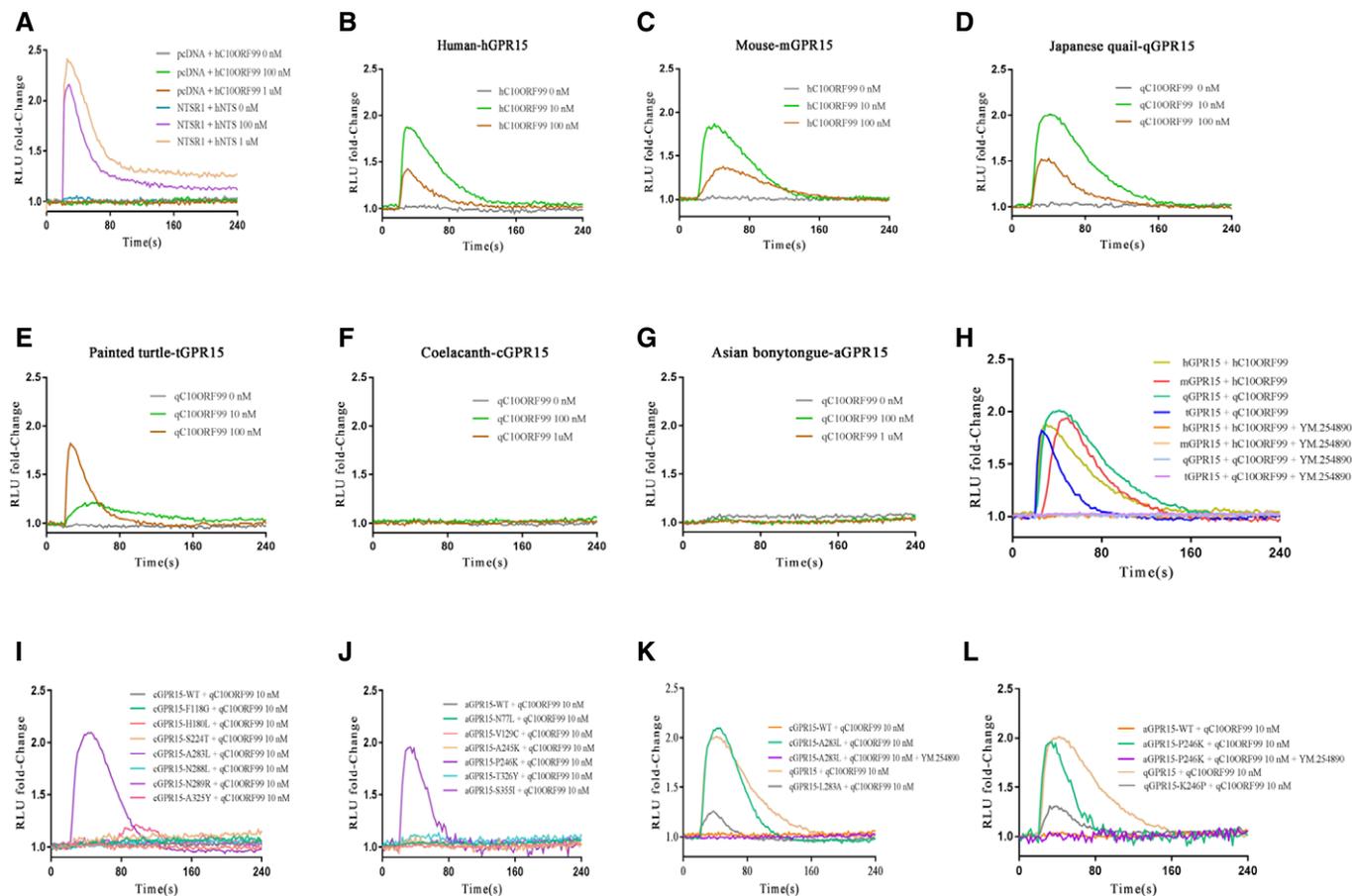
Figure 3. GPR15–C10ORF99 functional pairing in amniotes.

A–G The cellular levels of Ca^{2+} ions after ligand (C10ORF99) stimulation of the indicated GPR15 receptors (human, mouse, Japanese quail, painted turtle, coelacanth and Asian bony tongue GPR15). A plasmid-encoding human NTSR1 and the parental pcDNA plasmid were transfected as positive and negative controls respectively. The fold-change in calcium levels was calculated based on the fluorescence intensity (excitation and emission wavelengths: 490 and 520 nm respectively). Calcium RLU (relative light unit) fold was calculated using no stimulation data as standard value.

H The cellular levels of Ca^{2+} ions after ligand (C10ORF99) stimulation of GPR15 in human, mouse, Japanese quail, painted turtle. YM.254890 inhibits Ca^{2+} ion levels in cells transfected with GPR15 receptors of vertebrates. The fold-change in calcium levels was calculated based on the fluorescence intensity (excitation and emission wavelengths: 490 and 520 nm respectively).

I, J The cellular levels of Ca^{2+} ions after ligand (C10ORF99) stimulation of GPR15 in coelacanth, Asian bony tongue and their mutants.

K, L The cellular levels of Ca^{2+} ions after ligand (C10ORF99) stimulating GPR15 in Japanese quail and its mutants. The fold-change in calcium levels was calculated based on the fluorescence intensity (excitation and emission wavelengths: 490 and 520 nm respectively).

by qC10ORF99 (10 nM) based on Ca^{2+} measurements (Figs 3I and J, and EV2B and C). We also constructed point mutants for these key residues involved in qGPR15 activation (Dataset EV1). Both point mutants (L283A and K246P) exhibited decreased stimulation of qGPR15 and these two residues appeared to play critical roles in GPR15 activation by C10ORF99 (Fig 3K and L). Our results showed GPR15–C10ORF99 pairing in amniotes (but not during vertebrate origination) and, furthermore, indicated that GPR15 homologues of fishes and amniotes may serve different physiological functions.

Functional conservation of GPR15–C10ORF99 pairing in amniotes

The GPR15–C10ORF99 pairing is important for recruiting specialized Tregs in the colon (Kim *et al*, 2013; Nguyen *et al*, 2014; Ocon *et al*, 2017). Therefore, we designed chemotaxis assays to study the role of GPR15 and its ligand (C10ORF99) in migration. We also

constructed an adenovirus-based gene delivery vector to overexpress GPR15. We determined the migration activities of HEK293 and CD4[+] T cells *in vitro* by performing standard migration assays (Vinet *et al*, 2013; Adamczyk *et al*, 2017).

HEK293 cells expressing mGPR15, qGPR15 and tGPR15 showed significant C10ORF99-dependent migration (Fig 4A–D), whereas cGPR15 and aGPR15 (representing two different fish lineages) did not (Fig 4E–G). The cGPR15-A283L and aGPR15-P246K mutants showed restored migration ability, whereas wild-type cGPR15 and wild-type aGPR15 did not (Fig 4E–H). CD4[+] T cells were isolated from the peripheral blood of healthy 6-week-old GPR15-knockout mice, and an *in vitro* transwell migration assay towards CXCL12 was performed as a positive control (Adamczyk, 2017; Ocon *et al*, 2017). Cells expressing mGPR15, qGPR15 or tGPR15 (but not cGPR15 or aGPR15) (Fig 4M–O) showed significant C10ORF99-dependent migration (Fig 4I–L). The cGPR15-A283L and aGPR15-

P246K mutants showed rescued migration ability (Fig 4N–P), consistent with the calcium stimulation results (Fig 3).

CCR9 in chickens and turtles was highly expressed in the small intestine, but not in the colon, consistent with the CCR9 expression pattern observed in mammals (Fig 5A and B; Perrigoue *et al*, 2014). Moreover, frog CCR9 showed no significant expression differences in the small intestine and colon (Fig 5C; Devries *et al*, 2006). Opposite results were found in chicken and turtle GPR15, which was highly expressed in the colon, but not the small intestine, especially in turtles (Fig 5D and E). Consistently, C10ORF99 was also highly expressed in the colon, but not in the small intestine (Fig 5F and G). The C10ORF99 expression pattern was consistent with the GPR15

expression pattern, especially in turtles (Fig 5F and G), whereas neither C10ORF99 nor GPR15 was detected in frogs (Fig 5H). The relative expression of C10ORF99 in turtles was significantly higher than that in birds and mammals, presumably due to tGPR15 being relatively insensitive to ligand stimulation and requiring a higher C10ORF99 concentration (Fig 3E). Moreover, $CD4^+$ T cells of colon and small intestine were isolated from the healthy 6-week-old chickens, and the quantitative RT-PCR showed GPR15 to be highly expressed in the colonic $CD4^+$ T cells in comparison to small intestine (Fig 5I). Also, the *in vitro* transwell migration assay showed that colonic $CD4^+$ T cells exhibited much stronger C10ORF99-dependent migration compared to the $CD4^+$ T cells from the small intestine of chickens (Fig 5J).
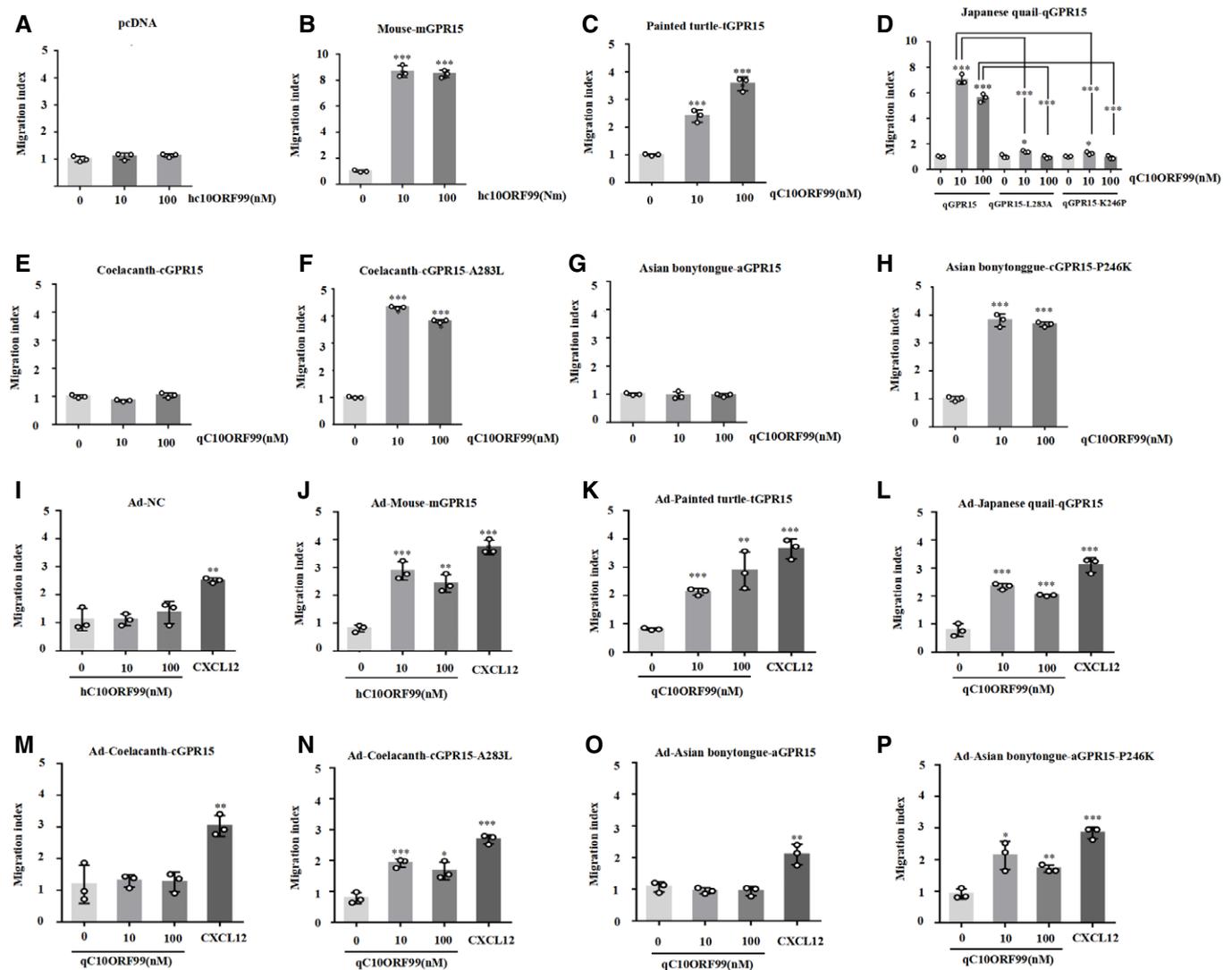
Figure 4. Functional conservation of GPR15–C10ORF99 pairing in amniotes.

A–H   The migration of HEK293 cells expressing mGPR15, qGPR15, tGPR15, cGPR15, aGPR15 and their mutants induced by ligand C10ORF99.

I–P   The migration of $CD4^+$ cells from GPR15 knockout mice, engineered to express mGPR15, qGPR15, tGPR15, cGPR15, aGPR15 and their mutants induced by ligand C10ORF99.

Data information: All data shown are generated from at least three technical replicates. Error bars indicate SD ($n = 3$). ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$ (*t*-test). Ad-NC represents adenovirus-based gene delivery vector as negative control.
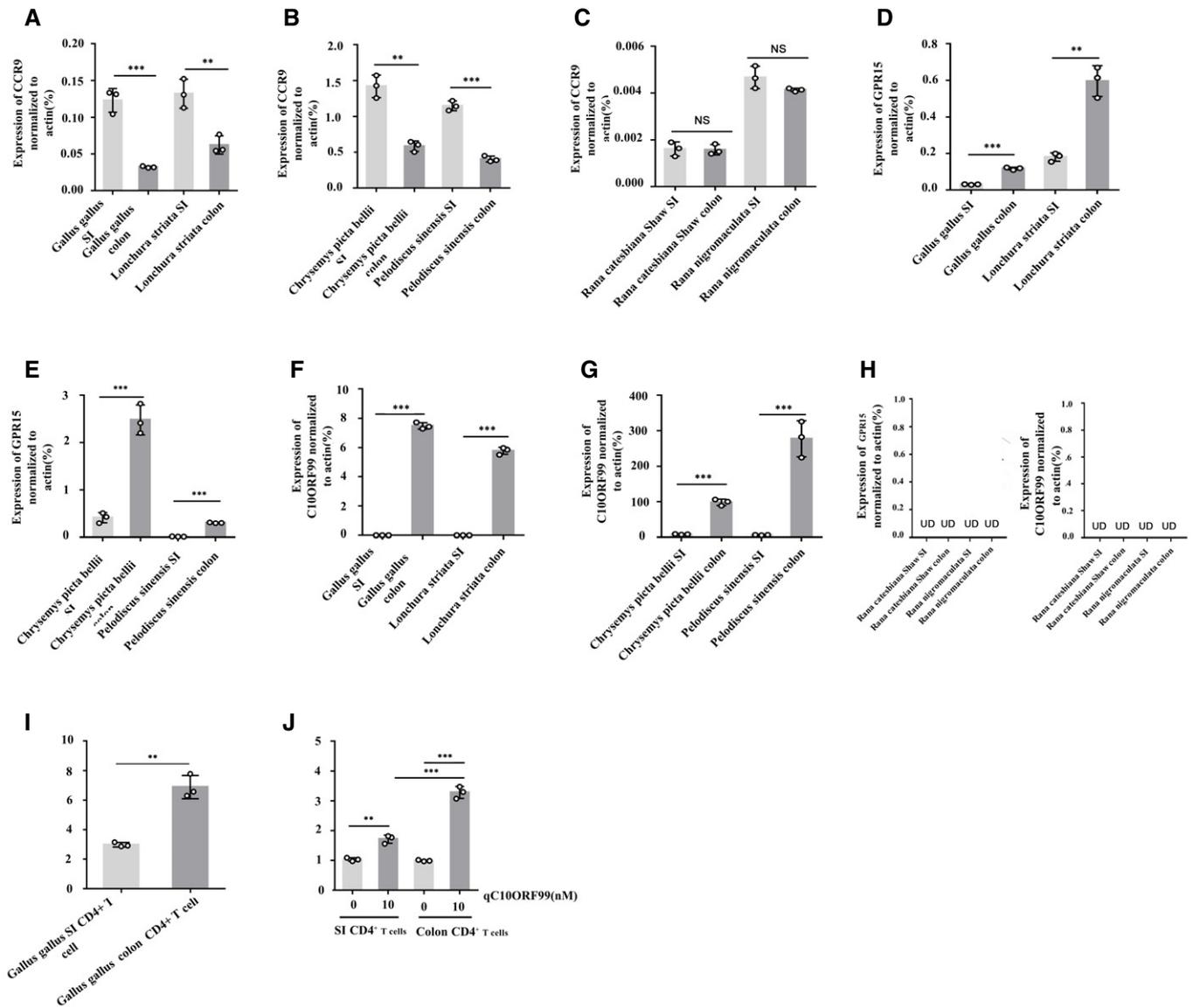
Figure 5. Conserved expression of GPR15–C10ORF99 pairing in amniotes.

A CCR9 expression in the small intestine (SI) and colon of *Gallus gallus* and *Lonchura striata*.
B CCR9 expression in the small intestine (SI) and colon of *Chrysemys picta bellii* and *Pelodiscus sinensis*.
C CCR9 expression in the small intestine (SI) and colon of *Rana catesbeiana Shaw* and *Rana nigromaculata*.
D GPR15 expression in the small intestine (SI) and colon of *Gallus gallus* and *Lonchura striata*.
E GPR15 expression in the small intestine (SI) and colon of *Chrysemys picta bellii* and *Pelodiscus sinensis*.
F C10ORF99 expression in the small intestine (SI) and colon of *Gallus gallus* and *Lonchura striata*.
G C10ORF99 expression in the small intestine (SI) and colon of *Chrysemys picta bellii* and *Pelodiscus sinensis*.
H GPR15 and C10ORF99 expression in the small intestine (SI) and colon of *Rana catesbeiana Shaw* and *Rana nigromaculata*.
I GPR15 expression in CD4[+] T cell from chicken small intestine (SI) and colon.
J The migration of CD4[+] cells from chicken small intestine (SI) and colon induced by ligand C10ORF99.

Data information: All data shown are representative of at least three independent technical replicates. Error bars indicate SD ($n = 3$). UD, undetermined, ***$P < 0.001$;
**$P < 0.01$; NS: no significance (*t*-test).

Correlation between diet and GPR15- and C10ORF99 expression levels in animal colons

In animal digestion systems, the diet affects the contents of SCFAs and microbial metabolites in the colon (Smith *et al*, 2013; Perrigoue

et al, 2014). SCFAs in the colon represent the main cause of colonic Treg homing and immune reactions mediated by GPR15 (Smith *et al*, 2013; Perrigoue *et al*, 2014). A similar process of herbivorous diet-induced SCFAs promoting Treg homing was also identified in birds (Selvaraj, 2013; Lee *et al*, 2018). Three batches of just-weaned

mice (18 mice per batch) were fed with a chicken, corn or mixture of corn and chicken diet for 5.5 months, representing a carnivore diet, an herbivore diet and an omnivore diet respectively (Fig 6A). As a control, wild-type mice were fed commercial food chow (composition is shown in Appendix Table S1, as well as the carnivore and herbivore diet) within or without propionate (a common SCFA that can induce GPR15 expression in animal colons), as a positive control (Smith *et al*, 2013; Adamczyk, 2017). Significantly higher GPR15 expression was induced in mice fed with corn (high), mixed corn and chicken (moderate) and commercial food chow (less than moderate but significantly higher than that in the chicken-only fed group). Mice on a diet containing propionate (positive control) showed the highest GPR15 expression and mice fed chicken showed the lowest GPR15 expression (less than the group obtaining commercial food chow; Fig 6B). GPR15 expression in mice on

commercial food chow was lower than that in mice on mixed corn and chicken, probably because in dietary experiment, mice on corn plus chicken preferred corn as their food (Appendix Table S2), and crude fibre only occupied a small percentage ($\leq 5\%$) in commercial food chow (Appendix Table S1). C10ORF99 showed high expression and was similarly induced in mice on corn, propionate, commercial food chow, mixed corn plus chicken (with highest expression in corn-fed mice), but not in mice on chicken alone (Fig 6C). C10ORF99 expression declined in all experimental groups after week 12 (peaking in week 10), and C10ORF99 expression in mice fed an herbivore and omnivore diet was still similar with that in the herbivore (e.g. goat, horse) and the omnivore (e.g. dog), and higher than that in mice fed a carnivore diet (Fig EV3).

To exclude a possible immune reaction in our dietary experiments, we also examined the expression of FOXP3, a Treg-
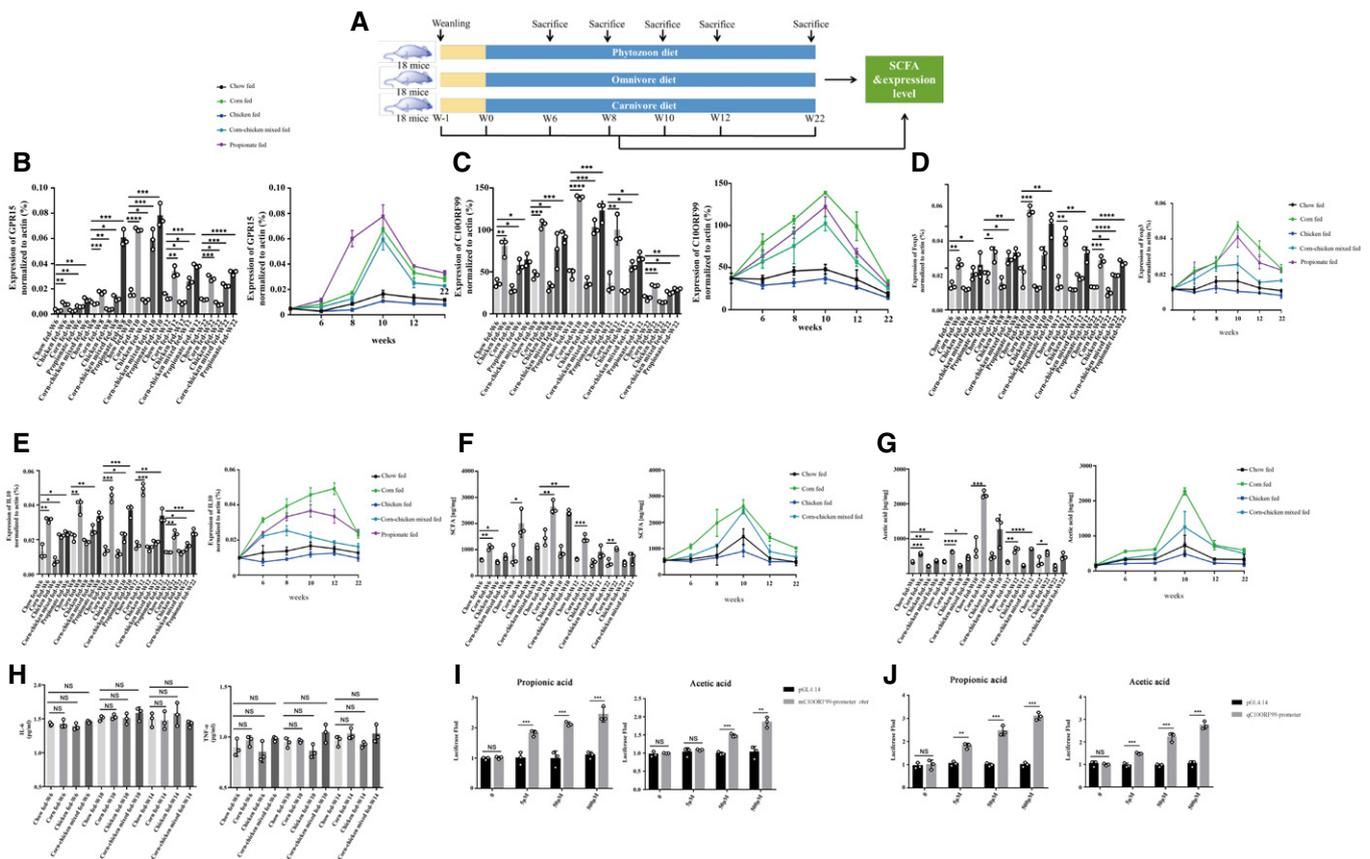
Figure 6. The correlation of GPR15 and C10ORF99 expression and diet in the mouse colon.

A   The scheme of three batches of just-weaned mice (18 mice per batch) were fed with a chicken, corn or mixture of corn and chicken diet for 5.5 months, representing a carnivore diet, an herbivore diet and an omnivore diet respectively

B–E The expression levels of GPR15, C10ORF99, FOXP3 and IL-10 in mice fed with corn, chicken and mixture of corn plus chicken, commercial food chow and a positive control (propionate added to commercial food chow) at different times (weanling mice were defined as w-1) in colonic tissue were determined by qRT-PCR. Each symbol or bar represents pooled RNA from colon tissues from three mice.

F, G SCFAs and acetic acid content in dietary experiments. Faecal samples from mice on the indicated diets were analysed for SCFAs by gas chromatography.

H   TNF-α and IL-6 levels in faecal samples from mice on the indicated diet demonstrated no colitis in all mouse groups. IL-6 and TNF-α protein production levels, as determined by performing enzyme-linked immunosorbent assays.

I, J (I) Luciferase activities of constructs with the mouse or (J) Japanese quail C10ORF99 promoter in HEK293 cells.

Data information: All data shown are representative of at least three independent biological replicates. Error bars indicate SD ($n = 3$). ****$P < 0.0001$ ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$; NS: no significance (one-way ANOVA).

associated transcriptional repressor, and interleukin-10 (IL-10), a key cytokine in Treg-mediated suppression (Smith *et al*, 2013). FOXP3 and IL-10 showed the highest, moderate, less than moderate and lowest expression levels in mice fed corn/propionate, mixed corn plus chicken, commercial food chow and chicken, respectively (Fig 6D and E). The frequency of Foxp3$^+$ Treg cells was similar in mice fed corn/propionate, mixed corn plus chicken, commercial food chow and chicken, respectively (Fig EV4A and C), although CD4$^+$ cell numbers showed no significant difference among the groups (Fig EV4B and D). Similar to C10ORF99 expression, this increasing trend continued with a peak occurring during week 10, with a decrease that was maintained until week 22 (Fig 6B–E). Thus, Treg homing was suddenly induced to a high level at the beginning and subsequently maintained at a relatively moderate level (Fig 6B–E). To exclude the side effects of intestinal inflammation (e.g. colitis, which can induce GPR15 expression; Adamczyk, 2017; Llewellyn *et al*, 2018) on GPR15 and C10ORF99 expression, we also examined the expression levels of inflammatory factors IL-6 and tumour necrosis factor (TNF)-α, which are important mediators of inflammatory reactions in patients with intestinal inflammation (Kim *et al*, 2013; Xiao *et al*, 2016; Llewellyn *et al*, 2018). No significant differences were found in IL-6 and TNF-α levels among mice during all different feeding plans (Fig 6H). Moreover, the colon length did not differ among mice fed an omnivore diet, herbivore diet or a carnivore diet (Appendix Fig S1).

To further investigate GPR15 upregulation in the colon, we also analysed the faecal SCFA contents of the experimental mice. Overall, the total levels of SCFAs were stimulated to varying degrees in mice fed corn, mixed corn plus chicken or commercial food chow, but not in mice on chicken (Fig 6F). Acetic acid, the most abundant SCFA in the colon that comprises over half of all SCFAs detected in faeces (Rios-Covian *et al*, 2016), was induced by similar amounts (Fig 6G), which was consistent with previous findings (Smith *et al*, 2013; Adamczyk, 2017). In order to test whether promoter activity was underlying the regulation, the upstream sequences (3 kb) of the mC10ORF99 and qC10ORF99 genes (i.e. potential transcription promoter regions) were cloned into the PGL4.14 plasmid for use in ORF luciferase reporter assays (Appendix Fig S2A–C) (Yang *et al*, 2018; Tian *et al*, 2020). We treated HEK293 cells expressing the chimeric plasmid with propionate or acetate. Luciferase activities significantly increased after stimulation, compared with the luciferase activity of C10ORF99 in control cells transfected with the parental PGL4.14 plasmid (Fig 6I and J). Colonic GPR15 expression was induced by an herbivore diet and SCFA treatment (Adamczyk, 2017). Colonic C10ORF99 expression was significantly induced by an herbivore diet and SCFA treatment, which can promote colonic Treg homing (Fig 6I and J).

Discussion

GPR15–C10ORF99 pairing mediates Treg homing to the colon, whereas CCR9–CCL25 pairing regulates Treg homing to the small intestine (Devries *et al*, 2006; Johnson *et al*, 2010; Perrigoue *et al*, 2014). The small intestine is a common and conserved organ across vertebrates, and the corresponding CCR9–CCL25 receptor–ligand pair directs a conserved physiological function in vertebrates (Devries *et al*, 2006; Perrigoue *et al*, 2014). The colon emerged in

tetrapods (Lindgren *et al*, 2010) and an herbivorous diet was a key "innovation" of amniotes during tetrapod evolution (Sumida & Martin, 1997; Sues & Reisz, 1998). An evolutionary model for the first colonic Treg homing reaction in amniotes is proposed in Fig 7. According to this model, GPR15 originated from the apelin/angiotensin/bradykinin peptide receptor family after two rounds of whole-genome duplication during vertebrate origination, but was lost in amphibians and almost all fishes, due to the lack of colonic Treg homing reactions (Figs 1 and 2). The colon first emerged in tetrapods, but underwent a major change when ancestral amniotes began to feed on plants, dating from the Late Carboniferous period (Sues & Reisz, 1998; Fastovsky, 2001). An herbivorous diet results in the sudden and massive delivery of high-fibre plant materials to the colon of amniotes, which harbours the largest population of microorganisms and their fermentation products, which is important because bacteria can break down undigestible dietary components such as fibres (Remely *et al*, 2013). A much stronger colonic immune reaction compared to the small intestine was specifically required in the colon of amniotes. Thus, the ligand (the mature C10ORF99 peptide) originating *de novo* after CDHR1 duplication adapted to a binding pocket in GPR15 to mediate specific Treg homing to the colon and to maintain colonic immune homeostasis. SCFAs resulting from bacterial fermentation stimulate both colonic GPR15 and C10ORF99 expression.

Thus, a novel biological process (herbivorous diet, leading to SCFA production and C10ORF99/GPR15-Treg homing for colon immune homeostasis) emerged due to the first herbivorous diet and the first GPR15–C10ORF99 signalling in the colon of amniotes (Fig 7). This process differs from vitamin A metabolite (retinoic acid)-dependent CCR9–CCL25-induced Treg homing, which is important for small intestine immune homeostasis and is functionally conserved in all vertebrates (Iwata *et al*, 2004; Perrigoue *et al*, 2014). Frog CCR9 functions in both the small intestine and colon (Fig 5; Devries *et al*, 2006) because there is no massive bacterial fermentation as a consequence of an herbivorous diet. Because of the lack of customized antibodies for species-specific isoforms, we could not determine the protein levels of CCR9, GPR15 and C10ORF99 in chicken, turtle and frog. SCFA-treated HEK293 cells expressing the chimeric plasmid showed increased expression of C10ORF99, which can promote colonic Treg homing. But how the regulation of the C10orf99 gene promoter/enhancer by SCFAs is occuring in the HEK293 system remains unknown. As GPR43 is involved as the primary step of acetate recognition, we analysed mRNA levels of GPR43 in HEK 293 cells. The results showed that its expression is undetectable in 293 cells by qPCR (Appendix Fig S4A), consistent with protein atlas data sets (Appendix Fig S4B).

A strictly carnivorous common shrew (*Sorex araneus*, a small mammal), with a diet consisting of insects, spiders, worms, amphibians and small rodents (Love *et al*, 2010), was dissected revealing a nearly negligible colon in its digestion system (Fig EV5). Consistently, by comparing the coding region of C10ORF99 with available genomes of *Sorex araneus*, we found that C10ORF99 in this strict carnivore has a missense mutation leading to a stop codon in this phylogenetically close species (Fig EV5A), although GPR15 is intact in the *Sorex araneus* genome (Dataset EV7). Interestingly, *Sorex araneus*-sGPR15 exhibits low stimulation induced by hC10ORF99, compared with the conserved function of its most closely related species *Erinaceus europaeus*-eGPR15 and *Tupaia chinensis*-tcGPR15
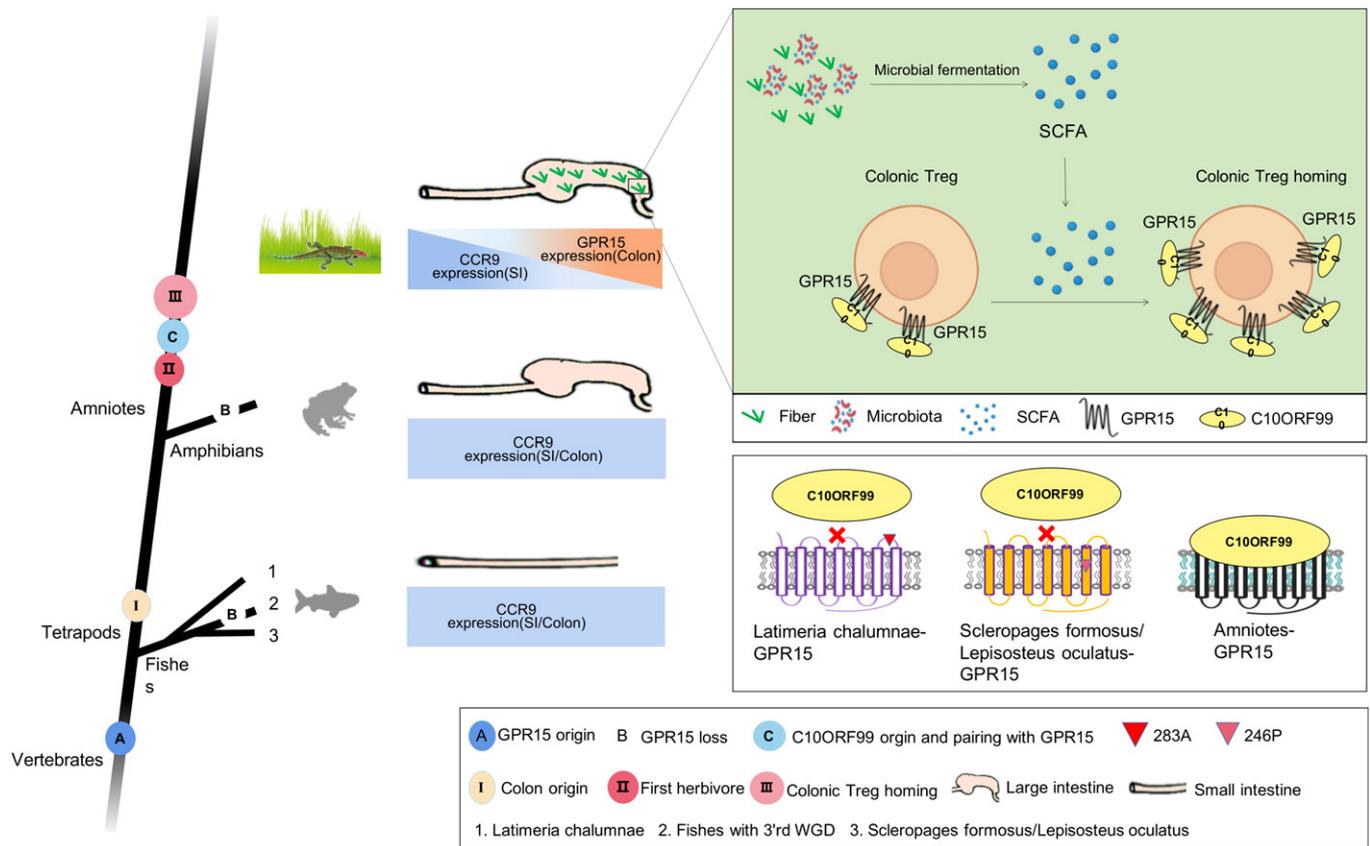
Figure 7. A proposed model where GPR15–C10ORF99 signalling initiates the first colonic Treg homing reaction in herbivorous amniotes.

(Fig EV5B). Two opposite biological processes (herbivorous diet leading to GPR15–C10ORF99 signalling and colonic Treg homing, and no herbivorous diet over a long time period, leading to the loss of colonic C10ORF99 expression) suggested that colonic Treg homing reactions can be promoted by an herbivorous diet, but can be disrupted by the lack of an herbivorous diet over a long time period, leading to disappearance of the colon and the loss of a functional C10ORF99 gene. In addition, the mRNA expression of C10ORF99 was also induced in skin in response to epithelial damage and TLR signalling (Kiene *et al*, 2014; Rioux *et al*, 2020), and is highly upregulated in skin during inflammatory diseases such as psoriasis (Suply *et al*, 2017; Chen *et al*, 2018).

The colon serves many important physiological functions, not only in absorbing water, sodium and other minerals but also in harbouring the largest population of microorganisms and collection of bacterial fermentation products in the digestive tract (Cummings, 1969). Feeding on plant tissues requires numerous anatomical and physiological changes to facilitate the efficient digestion of these materials. Thus, the colon morphology varies between species with different diets, for example, carnivores generally have short and smooth colons (Alvarez, 1948; David *et al*, 1980; Kararli, 2010; Feldhamer, 2011). In contrast, herbivores tend to have long and sacculated colons (Alvarez, 1948; David *et al*, 1980; Kararli, 2010; Feldhamer, 2011). The diet affects the gut microbiota composition and activity and the microbiota in the animal gastrointestinal tract has coevolved with the host (Mazmanian *et al*, 2005; Ley *et al*,

2008; Kamada *et al*, 2012), and their coexistence is mainly due to an equilibrium with the host immune system (Hooper *et al*, 2012; Kim *et al*, 2013). Furthermore, the colon harbours the largest population of microorganisms in the digestive tract, particularly in organisms with an herbivorous diet, and bacteria can break down undigestible dietary components, such as fibres (Remely *et al*, 2013). SCFAs are bacterial fermentation products that range in concentration from 50 mM to 100 mM in the colonic lumen (Cummings *et al*, 1987). Emerging data in mice indicated that SCFAs and the SCFA receptor (GPR43) play critical roles in colonic Treg homing (Natarajan & Pluznick, 2014; Perrigoue *et al*, 2014). Tregs utilize GPR15 (which is upregulated by SCFAs) to home to the colon, and SCFAs and GPR43 are required to generate and expand colonic Tregs (Perrigoue *et al*, 2014). Moreover, SCFAs induce GPR15 expression, hence contributing to colonic Treg homing (Perrigoue *et al*, 2014). Interestingly, when we expressed a C10orf99 gene promoter/enhancer construct in HEK 293 cells we observed upregulation of expression in response to SCFA, despite the fact that we could not find evidence for the presence of GPR43 in HEK cells (Appendix Fig S4A), consistent with protein atlas data sets (Appendix Fig S4B). This indicates that another yet undefined mechanism mediates the SCFA induction of C10orf99 in this cell system.

Treg cells play central roles in maintaining intestinal homeostasis by restraining inappropriate immune responses in the gastrointestinal tract under healthy and pathological conditions (Campbell, 2015). The CCR9/CCL25-chemokine axis controls Treg homing to

the small intestine (Perrigoue *et al*, 2014), and CCR9–CCL25 pairing maintains sequence conservation and functional conservation in vertebrates (Devries *et al*, 2006; Galindo-Villegas *et al*, 2013; Su *et al*, 2019), which is also consistent with the conservation of the small intestine in vertebrates (Devries *et al*, 2006; Bird & Tafalla, 2015). Notably, the gastrointestinal tracts of zebrafish and medaka express colon-specific genes (Wang *et al*, 2010; Aghaallaei *et al*, 2016). Because zebrafish Foxp3-expressing Treg cells can be found in the gastrointestinal tract (Hui *et al*, 2017), it is conceivable that the CCR9/CCL25 axis plays a more ancient role in T-cell homing to the gastrointestinal tract. In contrast, the GPR15–C10ORF99 receptor–ligand pair was found to selectively mediate Treg cell migration to the colon (Kim *et al*, 2013; Vinet *et al*, 2013), and we suggest this system developed after vertebrate origination in tetrapods. C10ORF99 ligand is expressed in colonic epithelial cells, enabling GPR15 high expressing Treg cell influx into the colon and thereby regulating local immune homeostasis. Here, we found that the C10ORF99 peptide originated via duplication of the flanking CDHR1 gene and functionally paired with GPR15 in amniotes. However, C10ORF99 exon1 and exon2 are corresponding to the regions of CDHR1, but the activating peptide encoded by exon3 is absent in CDHR1, suggesting that CDHR1 is not the agonist for GPR15 (Figs 2 and 3). By performing evolutionary analysis and generating experimental data, we found that GPR15–C10ORF99 pairing was functionally conserved and promoted colonic Treg homing in amniotes and their expression pattern exhibited a positive correlation with an herbivorous diet in animal colons (Figs 3 and 5). We propose a novel evolutionary model, wherein an herbivorous diet leads to SCFA production and C10ORF99/GPR15-dependent Treg homing and colon immune homeostasis, which first evolved in conjunction with the first herbivorous diet and the first GPR15–C10ORF99 signalling in the colon of ancestral amniotes (Fig 7).

During the course of evolution, GPR15 was presumably deleted in amphibians and most fishes, presumably because it was not exapted to serve a physiological function. However, two lineages of fish represented by *Latimeria chalumnae, Lepisosteus oculatus, Scleropages formosus* and the cartilage fish represented by *Carcharodon carcharias, Scyliorhinus canicula, Rhincodon typus* (Appendix Fig S3A–C) exhibited different evolutionary processes and may have employed other ligand interactions for other physiological functions or became a potential or genuine orphan GPCR (Tang *et al*, 2019). Future studies should be performed to investigate the different functions of the fish GPR15 genes in both fish lineages. Another exception is that the C10ORF99 gene was lost in the strictly carnivorous shrew, *Sorex araneus*, which also lost the colon (Fig EV5). According to the divergency time of *Sorex araneus* (Fig EV5), the closest species are *Erinaceus europaeus* and *Tupaia chinensis*, which have an intact C10ORF99–GPR15-signalling pathway and a colon in their digestive system, because both of them kept their omnivorous diet (Castex *et al*, 1978; Yu *et al*, 2016). The sequence of the *Sorex araneus* GPR15 gene still is intact, although some sensitivity to C10ORF99 has been lost, indicating that *Sorex araneus* GPR15 is in the process of undergoing inactivation or acquiring alternative functions, due to ligand loss (Fig EV5).

G protein-coupled receptors represent the largest family of membrane receptor proteins in vertebrates and play important roles in human physiology and disease (Kroeze, 2003). GPCR–ligand pairs have particularly high binding affinities and have maintained many conserved functions during vertebrate evolution (Kroeze, 2003). Particularly, the GPCR polypeptide family includes 16 subfamilies and almost all receptor–ligand functional pairs were formed during vertebrate origination, especially during two rounds of whole-genome duplication events (Cañestro, 2012). Polypeptide ligands from different vertebrates, for example human and fish, can interchange to stimulate their corresponding GPCRs, and these receptor–ligand pairs control conserved physiological processes and similar diseases in vertebrates (Nomiyama & Yoshie, 2015; Foster *et al*, 2019). However, after vertebrate origination, relatively few GPCR–ligand pairs were exapted for new organ origination and the corresponding physiological functions. Previous findings suggested that several GPCRs, for example, the relaxin receptor, BRS3, and GPR39, underwent differential molecular evolution in terms of GPCR–ligand pairing, leading to various potential physiological functions during vertebrate evolution (Sharma *et al*, 2018; Tang *et al*, 2019; Zhang *et al*, 2020). Our study also provides novel evidence that the origination and functional pairing of GPR15–C10ORF99 signalling participated in the origination of colonic immune functions in amniotes.

Materials and Methods

Plasmid construction, point mutations of GPR15 and construction of adenovirus-based gene delivery vector

GPR15 cDNA from various species were synthesized from General Biosystems Corporation Limited (Anhui, China), and GPR15 point mutants were generated by introducing point mutations using a QuikChange II site-directed mutagenesis kit (Agilent Technologies). Briefly, overlapping primers with the desired point mutations were used to amplify wild-type GPR15. The parental plasmid was digested using the DpnI enzyme, and the newly synthesized plasmid was used as a template for PCR amplification of GPR15 mutants before subcloning into the pcDNA3.1-V5-His plasmid. Adenovirus-based gene delivery vector was constructed by Hanbio Biotechnology Co., Ltd (http://www.hanbio.net). PGL4.14 plasmid was kindly provided by Dr. Ziqiang Zhu's lab.

Peptides of C10ORF99

As for ligand peptides, C-terminal peptide of human C10ORF99, containing the activating peptide for GPR15 (EPEPRLWVVP-GALPQV, as previously described (Ocon *et al*, 2017), and C-terminal peptide of Japanese quail C10ORF99 peptides (AAPSPLFVPGPLPQL) were synthesized from Biotech Bioengineering (Shanghai, China) for functional assays.

Phylogenetic analysis and positive selection analysis of apelin/angiotensin/bradykinin peptide receptor subfamily

The coding sequences of apelin/angiotensin/bradykinin peptide receptor subfamily were downloaded from NCBI, then translated into amino acid sequence and aligned with Clustal Omega (Sievers & Higgins, 2018), using default settings. The phylogenetic tree based on multiple alignments of amino acid sequences was obtained using

the neighbour-joining or maximum likelihood method in MEGA 7.0.26 (Sudhir *et al*, 2016). We used the branch-site model of PAML (Phylogenetic Analysis by Maximum Likelihood) version 4.4 (Yang, 2007) to test for positive selections on relevant amino acid positions. To identify the rapidly evolving genes in apelin/angiotensin/brady-kinin peptide receptor subfamily, we ran the Codeml module in the PAML software package with a free-ratio model (model = 1) and estimated lineage-specific evolutionary rates for each branch. Rapidly evolving genes (REGs), having a higher ω (*Ka/Ks*) than the rest of the lineage in the tree, were detected using a branch model (model = 2; Yang, 2007; Lin *et al*, 2019).

Cell culture and transfection

Human embryonic kidney 293-HEK293 and Chinese hamsters Ovary-CHO cells were used for the functional experiments. Both cell lines were maintained in Dulbecco's modified Eagle's medium (DMEM, purchased from Thermo) supplemented with 10% fetal bovine serum (FBS, Gibco), streptomycin and gentamicin (Gibco). The peripheral blood mononuclear cells of GPR15-knockout mice were prepared by density gradient separation on Ficoll-Paque Plus (Amersham Biosciences). CD4$^+$ T cells from mice were enriched using the CD4 (L3T4) MicroBeads (Miltenyi Biotec). CD3e (R&D systems, mouse CD3ε antibody, #145-2C11), CD28 (R&D systems, mouse CD28 antibody, #794716) antibodies and rmIL-2 (R&D systems, recombinant mouse IL-2, 402-ML) were used to activate the T cells. CD4$^+$ T cells from healthy 6-week-old chicken colon and small intestine were enriched using the anti-chicken CD4$^+$ antibody (Cat. No. Serotec MICA 2164F) as described previously (Rohe *et al*, 2017). Cells were grown in a humidified atmosphere at 37°C and 5% CO_2. Transfections of HEK293 and CHO cells were performed using Lipofectamine-3000 reagent (Invitrogen). Transfection of adenovirus-based gene delivery vector to HEK293 and CD4$^+$ cells was performed using Flat-angle centrifugation method from Hanbio Biotechnology Co., Ltd according to the manufacturer's instructions: Cells were pipetted into a centrifuge tube and centrifuged at low speed (200 × g) for 20 min. After removing most of the supernatant, the virus solution was added with multiplicity of infection (MOI) = 20, and incubated at room temperature for 15 min. Finally, the cells and virus fluid were simultaneously aspirated and transferred to a Petri dish to continue virus infection. Fresh medium was replaced after 24 h for further culturing for subsequent experiments.

Intracellular calcium assay

Intracellular calcium was measured using the non-wash calcium assay Fluo-4 kit (ab112129, Abcam) according to the manufacturer's instructions. Briefly, CHO cells were seeded in 24-well plates and were transfected with hGPR15, mGPR15, qGPR15, tGPR15, cGPR15, aGPR15, sGPR15, tcGPR15, eGPR15, GPR15 with various point mutants or the empty vector plasmid pcDNA3.1-V5-His (200 ng), respectively, and incubated overnight at 37°C with 5% CO_2. Cells were reseeded the following day into a 96-well assay plate (black plate and clear bottom). After 12 h, growth media were aspirated, and calcium dye was added. Following incubation for 30 min at 37°C and 10 min at room temperature, hC10ORF99 or qC10ORF99 with different concentrations (1 nM–1 μM) was added, and assay plates were placed into a fluorescence kinetic plate reader (Hamamatsu) immediately. Receptor-expressing CHO cells were pre-treated with the inhibitor (YM.254890) for 5 min at a concentration of 100 nM. Subsequently, hC10ORF99 and qC10ORF99 were added to the cells at a concentration of 10 nM or 100 nM, prior to the assay. The basal fluorescence intensity was recorded 15 times at 1 Hz for 20 s. The results were normalized to the average basal fluorescence intensity in ratio, and the peak response was used for the result calculation. Calcium RLU (relative light unit) fold was calculated using no stimulation data as standard value. Results were presented as mean ± SD from at least three independent experiments.

Luciferase assay

To detect whether SCFAs directly promote the C10ORF99 expression, HEK293 cells were seeded in 24-well plates and transfected with empty vector-pGL4.14 or qC10ORF99 promoter-pGL4.14 (300 ng) using Lipofectamine-3000 (Beyotime, Shanghai, China). After 36 h, sodium propionate or sodium acetate in serum-free medium was used to stimulate cells for 12 h. Luciferase activities were determined 48 h after transfection using luciferase assay kits (Beyotime, Shanghai, China) and normalized to β-galactosidase activity.

Calculation of gene expression levels

In order to quantify the expression levels of C10ORF99 in various animals, a total of nine RNA sequence data were downloaded from Sequence Read Archive (SRA), these data were generated from colon tissues of horse (Equus caballus, SRA accession code SRP224090) and dog (*Canis lupus familiaris*, SRA accession code SRP181218). Firstly, we used FastQC (http://www.bioinformatics. babraham.ac.uk/projects/fastqc/) for quality control checks with the above data, and according to the above results to trim low-quality reads using Trimmomatic (Anthony *et al*, 2014). Hisat2 (Kim *et al*, 2015) was applied with the default command options to align the trimmed reads against the reference genome (EquCab3.0 and CanFam3.1). Finally, featureCounts (Yang *et al*, 2014) was used to count the reads aligned to each gene. The TPM (Transcripts per million reads) was calculated in our study to quantify the gene expression levels.

Animals

C57BL/6 (wild-type) mice were purchased from the Model Animal Research Center of Nanjing University (Nanjing, China). Mice were housed under conventional conditions in filter top–protected cages and cared for in accordance with the Experimental Animal Center of Nanjing Normal University and were provided with food and water ad libitum. GPR15-knockout mice on a C57BL/6 background were purchased from Njbiogene Co., Ltd (http://njbiogene.com/). Specific-pathogen-free (SPF) chickens were purchased from Njbiogene Co., Ltd (http://njbiogene.com/). *Lonchura striata*, *Chrysemys picta bellii*, *Pelodiscus sinensis*, *Rana catesbeiana Shaw*, *Rana nigromaculata* were purchased from Taobao online shop (https://www. taobao.com/). All animal care and experimental procedures complied with the Animal Management Regulations of the Ministry of Health, China (Document No. 55, 2001) and were approved by the Science and Technology Department of Jiangsu Province [SYXK (SU) 2020–0047].

Migration assays

In vitro migration assays for HEK293 cells and CD4$^+$ T cells were conducted as described (Vinet *et al*, 2013; Adamczyk *et al*, 2017). 1–2 × 10^6 CD4$^+$ lymphocytes were added to the top wells of 5 μm pore, polycarbonate 24-well tissue culture inserts (Costar) in 100 μl, with 600 μl of chemokine solution in the bottom chamber. All migrations were conducted in chemotaxis medium at 37°C/5% CO$_2$ for 3 h. The chemokines used were recombinant mouse CXCL12/SDF-1 alpha protein (100 nM) (R&D Systems, 460-SD-050) and C10ORF99 mature peptide. The migration index was determined as the ratio of migrated cells towards CXCL12 in comparison to cell migrated to media alone.

Dietary experiments

Just-weaned C57BL/6 mice were used in diet feeding assays. The weanling mice were defined as w-1, and w0 indicates the first week when we started to provide different foods as below: Mice were provided with corn, chicken, corn-chicken mixed, commercial food chow (food was changed daily) for 5.5 months and with/without propionate (150 mM; Smith *et al*, 2013) in drinking water (water changed weekly).

RNA isolation and quantitative real-time PCR

RNA was obtained from mouse colon using the RNeasy Fibrous Tissue Kit (Qiagen). Following DNase digestion (Qiagen), cDNA was synthesized with M-MLV reverse transcriptase (Promega) and oligo (dT) mixed with random hexamer primers (Invitrogen). Real-time PCR was performed using the SYBR Green PCR kit (Vazyme Biotech Co., Ltd, Nanjing, China) and specific primers for mouse GPR15 (5′-CCATTGT-GTGGCCAGTCGTA-3′ and 5′-GAAGAGTAGGCAACC CCAGC-3′), mouse FoxP3 (5′-GAACGCCATCCGCCACAACCTGA-3′ and 5′-CCCTGCCCCCACCACCTCTGC-3′), mouse C10ORF99 (5′-CTT CTAGCCCTTTCCGGTCT-3′ and 5′-CACCACCCATGACTTGACTG-3′), mouse IL10 (5′-TGCTATGCTGCCTGCTCTTA-3' and 5′-TCATTTC CGATAAGGCTTGG-3'), mouse beta-actin (5'-GGTGGGAATGGGT CAGAAGG-3' and 5'-AATGCCTGGGTACATGGTGG-3'). To determine CCR9, GPR15 and C10ORF99 expression pattern in reptiles/birds/amphibians' gastrointestinal tract, the adult animals were sacrificed to isolate the small intestine and colon. RT-PCR protocol was described above and specific primers from conserved sequences of bird GPR15 (5′-ACCTTTTACTGCTCCATCAC-3′ and 5′-AATGA AAGGGGTTGGCACAGC-3′), turtle GPR15 (5'-GCCAGGAAAGTCA GAACAAG-3' and 5'-TAGATGAAAGGGTTGGCACA-3'), bird C10ORF99 (5'-TTGTCTTTCTTTTGTGCCTC-3' and 5'-ACAGGGCTT GCATTTCCTTA-3'), turtle C10ORF99 (5'-ATGAAGCTCTTCACGC TC-3' and 5'-GAGGAAGCGGTCCTGGTTTC-3'), bird CCR9 (5'-AT GTATAAGATCAACTTCTA-3'; 5'-GATCATGGTGATGATCTT-3'), turtle CCR9 (5'-TTGTCAATAGCATGTACAAGATCAA-3'; 5'-TAGGA CAGTAACTTTGATGGT-3'), frog CCR9 (5'-AGAAAGCTAAAAAC AATGAC-3' and 5'-ACATGTTAACTTGCTGATAA-3'), bird beta-actin (5'-CATCACCATTGGCAATGAGA-3' and 5'-ATGCCAGGGTACATT GTGGT-3'), turtle beta-actin (5'-CTTCAACTCCATCATGAAGTG and 5'-GCAATGATCTTAATTTTCAT-3'), frog beta-actin (5'-CATGGACT CAGGTGATGGTG-3' and 5'-GCTGTGGTGGTGAAGCTGTA-3'). Conserved sequence alignment is shown in Dataset EV8.

Immunohistochemistry, microscopy and image analysis

Mouse colon tissues were fixed in 4% paraformaldehyde in PBS, embedded in paraffin, sectioned and the fixed sections were incubated in 3% H$_2$O$_2$ solution in PBS at room temperature for 10 min. Nonspecific antibody binding was blocked by incubation with 5% normal goat serum in PBS for 1 h at room temperature. Slides were stained overnight at 4°C with the following primary antibodies: CD4 (Themo Fisher Scientific, PA5-87425; 1:200 dilution), Foxp3 (Cell Signaling Technology, D6O8R; 1:200 dilution), HRP-conjugated Goat Anti-Rabbit IgG (H + L) (Invitrogen, 31210) were used as secondary reagents. The sections were developed using CY3-Tyramide and nuclear counterstaining was performed with 4′,6-diamidino-2-phenylindole (DAPI; Sigma- Aldrich, D9542). Images were analysed with a Leica DM RXA2 confocal microscope controlled by Leica Microsystems confocal software (version 2.61 Build 1537; all from Leica Microsystems, Wetzlar, Germany). ImageJ (National Institutes of Health) was used for further quantification of the fluorescence and intensities of the images.

Measurement of faecal SCFA levels

Faecal samples from mice under different diets were analysed for SCFA by gas chromatography. For this, 100 mg faeces was suspended in 400 μl saline solution. After homogenization and centrifugation, the supernatants were intensely mixed with saline and diethyl ether and spiked with methylacrylate. After freezing of the mixture for at least 6 h at –25°C, the organic phase was isolated and analysed. This was performed with a Shimadzu GC-2010 Plus gas chromatograph fitted with a 25 m × 0.20 mm i.d. PEG column (BP 21, SEG) and a flame ionization detector. Ultra-high purity helium was used as carrier gas.

Statistical analysis

Experiments were repeated independently at least three times. Results were analysed using GraphPad Prism 7. Differences between the two groups were compared using two-tailed Student *t*-test. Statistical significance was defined as a $0.01 \leq P < 0.05$ (*), $0.001 \leq P < 0.01$ (**), $0.0001 \leq P < 0.001$(***) and $P < 0.0001$ (****).

Data availability

No primary data sets have been generated or deposited.

Expanded View for this article is available online.

Author contributions

CD and JB designed the study. CD, JS and HZ wrote the paper. JS, HZ, QS and WY performed and analysed the experiments. CD, YC, JC, WL, JL, JJ, XX, JX, KH

and FL analysed the data. All authors reviewed the results and approved the final version of the manuscript. The authors declare that they have no conflicts of interest with the contents of this article.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Adamczyk A, Gageik D, Frede A, Pastille E, Hansen W, Rueffer A, Buer J, Buning J, Langhorst J, Westendorf AM (2017) Differential expression of GPR15 on T cells during ulcerative colitis. *JCI Insight* 2: e90585

Aghaallaei N, Gruhl F, Schaefer CQ, Wernet T, Weinhardt V, Centanin L, Loosli F, Baumbach T, Wittbrodt J (2016) Identification, visualization and clonal analysis of intestinal stem cells in fish. *Development* 143: 3470–3480

Alvarez WC (1948) An introduction to gastro-enterology. *Cal Med* 68: 404

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.

Atarashi K, Tanoue T, Shima T, Imaoka A, Kuwahara T, Momose Y, Cheng G, Yamasaki S, Saito T, Ohba Y (2016) Induction of colonic regulatory T cells by indigenous species. *Science* 331: 337–341

Bian C, Hu Y, Ravi V, Kuznetsova IS, Shen X, Mu X, Sun Y, You X, Li J, Li X *et al* (2016) The Asian arowana (*Scleropages formosus*) genome provides new insights into the evolution of an early lineage of teleosts. *Sci Rep* 6: 24501

Bird S, Tafalla C (2015) Teleost chemokines and their receptors. *Biology* 4: 756–784

Campbell DJ (2015) Control of regulatory T cell migration, function, and homeostasis. *J Immunol* 195: 2507–2513

Cañestro C (2012) Two rounds of whole-genome duplication: evidence and impact on the evolution of vertebrate innovations. *Semin Cell Dev Biol* 24: 83–94

Castan L, Magnan A, Bouchaud G (2017) Chemokine receptors in allergic diseases. *Allergy* 72: 682–690

Castex C, Hoo-Paris R, Donnio R, More J, Sutter BCJ (1978) Determination of insulin levels of the common dormouse (glis glis) and hedgehog (*Erinaceus europaeus*) with reference to the rat. *Comp Biochem Physiol A Physiol* 59: 285–289

Chen C, Wu N, Duan Q, Yang H, Wang X, Yang P, Zhang M, Liu J, Liu Z, Shao Y *et al* (2018) C10orf99 contributes to the development of psoriasis by promoting the proliferation of keratinocytes. *Sci Rep* 8: 8590

Cummings JH (1969) Absorption and secretion by the colon. *Gastroenterology* 16: 966–971

Cummings JH, Pomare EW, Branch WJ, Naylor CPE, Macfarlane GT (1987) Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut* 28: 1221–1227

David J, Chivers C, Hladik M (1980) Morphology of the gastrointestinal tract in primates: Comparisons with other mammals in relation to diet. *J Morphol* 166: 337–386.

DeVries ME, Kelvin AA, Xu L, Ran L, Robinson J, Kelvin DJ (2006) Defining the origins and evolution of the chemokine/chemokine receptor system. *J Immunol* 176: 401–415

Ding Y, Xu J, Bromberg JS (2012) Regulatory T cell migration during an immune response. *Trends Immunol* 33: 174–180

Fastovsky DE (2001) Evolution of herbivory in terrestrial vertebrates: perspectives from the fossil record. *Palaios* 16: 528–529

Feldhamer GA (2011) Mammalogy: adaptation, diversity, ecology. *Q Rev Biol* 83: 322–323

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791

Foster SR, Hauser AS, Vedel L, Strachan RT, Huang XP, Gavin AC, Shah SD, Nayak AP, Haugaard-Kedstrom LM, Penn RB *et al* (2019) Discovery of human signaling systems: pairing peptides to G protein-coupled receptors. *Cell* 179: 895–908

Fournier D, Luft FC, Bader M, Ganten D, Andrade-Navarro MA (2012) Emergence and evolution of the renin–angiotensin–aldosterone system. *J Mol Med* 90: 495–508

Galindo-Villegas J, Mulero I, García-Alcazar A, Muñoz I, Peñalver-Mellado M, Streitenberger S, Scapigliati G, Meseguer J, Mulero V (2013) Recombinant TNFα as oral vaccine adjuvant protects European sea bass against vibriosis: insights into the role of the CCL25/CCR9 axis. *Fish Shellfish Immunol* 35: 1260–1271

Geuking MB, Cahenzli J, Lawson MA, Ng DC, Slack E, Hapfelmeier S, McCoy KD, Macpherson AJ (2011) Intestinal bacterial colonization induces mutualistic regulatory T cell responses. *Immunity* 34: 794–806

Hooper LV, Littman DR, Macpherson AJ (2012) Interactions between the microbiota and the immune system. *Science* 336: 1268–1273

Hu GM, Mai TL, Chen CM (2017) Visualizing the GPCR network: classification and evolution. *Sci Rep* 7: 15495

Hui SP, Sheng DZ, Sugimoto K, Gonzalez-Rajal A, Nakagawa S, Hesselson D, Kikuchi K (2017) Zebrafish regulatory T cells mediate organ-specific regenerative programs. *Dev Cell* 43: 659–672

Iwata M, Hirakiyama A, Eshima Y, Kagechika H, Kato C, Song SY (2004) Retinoic acid imprints gut-homing specificity on T cells. *Immunity* 21: 527–538

Johnson EL, Singh R, Singh S, Johnson-Holiday CM, Grizzle WE, Partridge EE, Lillard Jr JW (2010) CCL25-CCR9 interaction modulates ovarian cancer cell migration, metalloproteinase expression, and invasion. *World J Surg Oncol* 8: 62

Kamada N, Kim YG, Sham HP, Vallance BA, Puente JL, Martens EC, Nunez G (2012) Regulated virulence controls the ability of a pathogen to compete with the gut microbiota. *Science* 336: 1325–1329

Kararli TT (2010) Comparison of the gastrointestinal anatomy, physiology, and biochemistry of humans and commonly used laboratory animals. *Biopharm Drug Dispos* 16: 351–380

Kiene M, Rethi B, Jansson M, Dillon S, Lee E, Lantto R, Wilson C, Pohlmann S, Chiodi F (2014) Toll-like receptor 3 signalling up-regulates expression of the HIV co-receptor G-protein coupled receptor 15 on human CD4$^+$ T cells. *PLoS One* 9: e88195

Kim D, Langmead B, Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 12: 357–360

Kim SV, Xiang WV, Kwak C, Yang Y, Lin XW, Ota M, Sarpel U, Rifkin DB, Xu R, Littman DR (2013) GPR15-mediated homing controls immune homeostasis in the large intestine mucosa. *Science* 340: 1456–1459

Kroeze WK, Sheffler DJ, Roth BL (2003) G-protein-coupled receptors at a glance. *J Cell Sci* 116: 4867–4869

Kuhn KA, Stappenbeck TS (2013) Peripheral education of the immune system by the colonic microbiota. *Semin Immunol* 25: 364–369

Lee IK, Gu MJ, Ko KH, Bae S, Kim G, Jin GD, Kim EB, Kong YY, Park TS, Park BC *et al* (2018) Regulation of CD4(+)CD8(-)CD25(+) and CD4(+)CD8(+)CD25(+) T cells by gut microbiota in chicken. *Sci Rep* 8: 8627

Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R *et al* (2008) Evolution of mammals and their gut microbes. *Science* 320: 1647–1651

Lin Z, Chen L, Chen X, Zhong Y, Yang Y, Xia W, Liu C, Zhu W, Wang H, Yan B *et al* (2019) Biological adaptations in the *Arctic cervid*, the reindeer (*Rangifer tarandus*). *Science* 364: eaav6312

Lindgren J, Caldwell MW, Konishi T, Chiappe LM, Farke AA (2010) Convergent evolution in aquatic tetrapods: insights from an exceptional fossil mosasaur. *PLoS One* 5: e11998

Llewellyn SR, Britton GJ, Contijoch EJ, Vennaro OH, Mortha A, Colombel JF, Grinspan A, Clemente JC, Merad M, Faith JJ (2018) Interactions between diet and the intestinal microbiota alter intestinal permeability and colitis severity in mice. *Gastroenterology* 154: 1037−1046

Love RA, Webon C, Glue DE, Harris S (2010) Changes in the food of British Barn Owls (*Tyto alba*) between 1974 and 1997. *Mam Rev* 30: 107−129

Mazmanian SK, Liu CH, Tzianabos AO, Kasper DL (2005) An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* 122: 107−118

McGuckin MA, Lindén SK, Sutton P, Florin TH (2011) Mucin dynamics and enteric pathogens. *Nat Rev Microbiol* 9: 265−278

de Mendoza A, Sebe-Pedros A, Ruiz-Trillo I (2014) The evolution of the GPCR signaling system in eukaryotes: modularity, conservation, and the transition to metazoan multicellularity. *Genome Biol Evol* 6: 606−619

Natarajan N, Pluznick JL (2014) From microbe to man: the role of microbial short chain fatty acid metabolites in host cell biology. *Am J Physiol Cell Physiol* 307: C979−C985

Nguyen LP, Pan J, Dinh TT, Hadeiba H, O'Hara E, Ebtikar A, Hertweck A, Gökmen MR, Lord GM, Jenner RG *et al* (2014) Role and species-specific expression of colon T cell homing receptor GPR15 in colitis. *Nat Immunol* 16: 207−213

Nomiyama H, Yoshie O (2015) Functional roles of evolutionary conserved motifs and residues in vertebrate chemokine receptors. *J Leukoc Biol* 97: 39−47

Ocon B, Pan J, Dinh TT, Chen W, Ballet R, Bscheider M, Habtezion A, Tu H, Zabel BA, Butcher EC (2017) A Mucosal and cutaneous chemokine ligand for the lymphocyte chemoattractant receptor GPR15. *Front Immunol* 8: 1111

Perrigoue J, Das A, Mora JR (2014) Interplay of nutrients and microbial metabolites in intestinal immune homeostasis: distinct and common mechanisms of immune regulation in the small bowel and colon. *Nestle Nutr Inst Workshop Ser* 79: 57−71

Pierce KL, Premont RT, Lefkowitz RJ (2002) Seven-transmembrane receptors. *Nat Rev Mol Cell Biol* 3: 639−650

Remely M, Simone D, Hippe B, Kesselring J, Aumueller E, Helmut B, Haslberger A (2013) Abundance and diversity of microbiota in type 2 diabetes and obesity. *J Diabetes Metab* 4: 253

Rios-Covian D, Ruas-Madiedo P, Margolles A, Gueimonde M, de Los Reyes-Gavilan CG, Salazar N (2016) Intestinal short chain fatty acids and their link with diet and human health. *Front Microbiol* 7: 185

Rioux G, Ridha Z, Simard M, Turgeon F, Guerin SL, Pouliot R (2020) Transcriptome profiling analyses in psoriasis: a dynamic contribution of keratinocytes to the pathogenesis. *Genes* 11: 1155

Rohe I, Gobel TW, Goodarzi Boroojeni F, Zentek J (2017) Effect of feeding soybean meal and differently processed peas on the gut mucosal immune system of broilers. *Poult Sci* 96: 2064−2073

Schrage R, Schmitz A-L, Gaffal E, Annala S, Kehraus S, Wenzel D, Büllesbach KM, Bald T, Inoue A, Shinjo Y *et al* (2015) The experimental power of FR900359 to study Gq-regulated biological processes. *Nat Commun* 6: 10156

Selvaraj RK (2013) Avian CD4(+)CD25(+) regulatory T cells: properties and therapeutic applications. *Dev Comp Immunol* 41: 397−402

Sharma V, Lehmann T, Stuckas H, Funke L, Hiller M (2018) Loss of RXFP2 and INSL3 genes in Afrotheria shows that testicular descent is the ancestral condition in placental mammals. *PLoS Biol* 16: e2005293

Sievers F, Higgins DG (2018) Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci* 27: 135−145

Slosky LM, Bai Y, Toth K, Ray C, Rochelle LK, Badea A, Chandrasekhar R, Pogorelov VM, Abraham DM, Atluri N *et al* (2020) β-arrestin-biased allosteric modulator of ntsr1 selectively attenuates addictive behaviors. *Cell* 181: 1364−1379

Smith PM, Howitt MR, Panikov N, Michaud M, Gallini CA, Bohlooly YM, Glickman JN, Garrett WS (2013) The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science* 341: 569−573

Strotmann R, Schröck K, Böselt I, Stäubert C, Russ A, Schöneberg T (2011) Evolution of GPCR: change and continuity. *Mol Cell Endocrinol* 331: 170−178

Su Y, Deng Y, Cheng C, Ma H, Feng J (2019) Molecular characterization and expression analysis of the CCR9 gene from cobia (*Rachycentron canadum*) following bacterial and poly I: C challenge. *J Appl Anim Res* 47: 454−462

Sudhir K, Glen S, Koichiro T (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33: 1870−1874

Sues HD, Reisz RR (1998) Origins and early evolution of herbivory in tetrapods. *Trends Ecol Evol* 13: 141−145

Sumida SS, Martin KLM (1997) *Amniote origins: completing the transition to land*, San Diego: Academic Press

Suply T, Hannedouche S, Carte N, Li J, Grosshans B, Schaefer M, Raad L, Beck V, Vidal S, Hiou-Feige A *et al* (2017) A natural ligand for the orphan receptor GPR15 modulates lymphocyte recruitment to epithelia. *Sci Signal* 10: eaal0180

Tang H, Shu C, Chen H, Zhang X, Zang Z, Deng C (2019) Constitutively active BRS3 is a genuinely orphan GPCR in placental mammals. *PLoS Biol* 17: e3000175

Tian L, Xiao H, Li M, Wu X, Xie Y, Zhou J, Zhang X, Wang B (2020) A novel Sprouty4-ERK1/2-Wnt/β-catenin regulatory loop in marrow stromal progenitor cells controls osteogenic and adipogenic differentiation. *Metabolism* 105. 154189

Vinet J, van Zwam M, Dijkstra IM, Brouwer N, van Weering HR, Watts A, Meijer M, Fokkens MR, Kannan V, Verzijl D *et al* (2013) Inhibition of CXCR3-mediated chemotaxis by the human chemokine receptor-like protein CCX-CKR. *Br J Pharmacol* 168: 1375−1387

Vogel KJ, Brown MR, Strand MR (2013) Phylogenetic investigation of peptide hormone and growth factor receptors in five dipteran genomes. *Front Endocrinol* 4: 193

Wang Z, Du J, Lam SH, Mathavan S, Matsudaira P, Gong Z (2010) Morphological and molecular evidence for functional organization along the rostrocaudal axis of the adult zebrafish intestine. *BMC Genom* 11: 392

Xiao Y-T, Yan W-H, Cao Y, Yan J-K, Cai W (2016) Neutralization of IL-6 and TNF-α ameliorates intestinal permeability in DSS-induced colitis. *Cytokine* 83: 189−192

Yang L, Mo W, Yu X, Yao N, Zhou Z, Fan X, Zhang L, Piao M, Li S, Yang D *et al* (2018) Reconstituting *Arabidopsis* CRY2 signaling pathway in mammalian cells reveals regulation of transcription by direct binding of CRY2 to DNA. *Cell Rep* 24: 585−593

Yang L, Smyth GK, Wei S (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30: 923−930

Yang M, Tang M, Ma X, Yang L, He J, Peng X, Guo G, Zhou L, Luo N, Yuan Z *et al* (2015) AP-57/C10orf99 is a new type of multifunctional antimicrobial peptide. *Biochem Biophys Res Comm* 457: 347−352

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

Yu W, Yang C, Bi Y, Long F, Li Y, Wang J, Huang F (2016) Characterization of hepatitis E virus infection in tree shrew (*Tupaia belangeri chinensis*). *BMC Infect Dis* 16: 80

Zabel BA, Rott A, Butcher EC (2015) Leukocyte chemoattractant receptors in human disease pathogenesis. *Annu Rev Pathol* 10: 51 – 81

Zhang L, Song J, Zang Z, Tang H, Li W, Lai S, Deng C (2020) Adaptive evolution of GPR39 in diverse directions in vertebrates. *Gen Comp Endocrinol* 299: 113610

Zlotnik A, Yoshie O (2012) The chemokine superfamily revisited. *Immunity* 36: 705 – 716