Comparative evolutionary histories of kisspeptins and kisspeptin receptors in vertebrates reveal both parallel and divergent features

Jérémy Pasquier1, Anne-Gaëlle Lafont1, Hervé Tostivint2, Hubert Vaudry3, Karine Rousseau1 and Sylvie Dufour1*

1 Research Unit BOREA, Biology of Aquatic Organisms and Ecosystems, Centre National de la Recherche Scientifique 7208, Institut de Recherche pour le Développement 207, Université Pierre et Marie Curie, Muséum National d’Histoire Naturelle, Paris, France
2 UMR 7221 CNRS/MNHN Evolution des Régulations Endocriniennes, Muséum National d’Histoire Naturelle, Paris, France
3 Laboratory of Cellular and Molecular Neuroendocrinology, INSERM U982, European Institute for Peptide Research (IFRMP 23), University of Rouen, Mont-Saint-Aignan, France

INTRODUCTION

As increasing vertebrate genomes have been explored, the understanding of their structure and evolution has progressed in parallel. Indeed, the comparison of their gene organization shed light on various large-scale genomic events that occurred along vertebrate radiation. Among those events, in the early stages of their history, vertebrates experienced two rounds of whole genome duplication (1R and 2R), resulting in fourfold-replicated genomes (Dehal and Boore, 2005; Van de Peer et al., 2010). These two events can be traced through the study of gene families currently presenting up to four paralogs. In addition, the comparison of teleost genomes with other vertebrate genomes revealed a teleost-specific third round of whole genome duplication (3R), resulting in up to eight paralogs in the same gene family in this lineage (Amores et al., 1998; Meyer and Van de Peer, 2005; Kasahara et al., 2007).

During the past decade, the kisspeptin system has been identified in various vertebrates, leading to the discovery of multiple genes encoding both peptides (Kiss) and receptors (Kissr). The investigation of recently published genomes from species of phylogenetic interest, such as a chondrichthyan, the elephant shark, an early sarcopterygian, the coelacanth, a non-teleost actinopterygian, the spotted gar, and an early teleost, the European eel, allowed us to get new insights into the molecular diversity and evolution of both Kiss and Kissr families. We identified four Kissr in the spotted gar and coelacanth genomes, providing the first evidence of four Kissr genes in vertebrates. We also found three Kiss in the coelacanth and elephant shark genomes revealing two new species, in addition to Xenopus, presenting three Kiss genes. Considering the increasing diversity of kisspeptin system, phylogenetic, and synteny analyses enabled us to clarify both Kiss and Kissr classifications. We also could trace back the evolution of both gene families from the early steps of vertebrate history. Four Kissr and four Kiss paralogs may have arisen via the two whole genome duplication rounds (1R and 2R) in early vertebrates. This would have been followed by multiple independent Kiss and Kissr gene losses in the sarcopterygian and actinopterygian lineages. In particular, no impact of the teleost-specific 3R could be recorded on the numbers of teleost Kissr or Kiss paralogs. The origin of their diversity via 1R and 2R, as well as the subsequent occurrence of multiple gene losses, represent common features of the evolutionary histories of Kiss and Kissr families in vertebrates. In contrast, comparisons also revealed un-matching numbers of Kiss and Kissr genes in some species, as well as a large variability of Kiss/Kissr couples according to species. These discrepancies support independent features of the Kiss and Kissr evolutionary histories across vertebrate radiation.

Keywords: kisspeptin, kisspeptin receptor, phylogeny, synteny, evolutionary history, spotted gar, coelacanth, European eel
has been identified in various vertebrate species, leading to the discovery of multiple genes encoding Kiss as well as multiple genes encoding Kissr (Biran et al., 2008; Felip et al., 2009; Kitahashi et al., 2009; Lee et al., 2009).

Concerning Kiss and Kissr diversity, contrasting situations are found in the different vertebrate phyla. Indeed, in eutherian species, one single gene, named KissIr, encodes the kisspeptin receptor and one single gene, named Kiss1, encodes kisspeptin. In protochondrichthyan, such as platypus (Ornithorhynchus anatinus), two Kiss, and two Kissr are present (Lee et al., 2009). To date, in teleosts, two situations have been reported. One Kiss gene and one Kissr gene are present in some species such as fugu (Takifugu niphobles), tetraodon (Tetraodon nigroviridis), and stickleback (Gasterosteus aculeatus). In contrast, a second Kiss as well as a second Kissr genes have been characterized in some species including zebrafish (Danio rerio; Biran et al., 2008), goldfish (Carassius auratus; Li et al., 2009), medaka (Oryzias latipes; Lee et al., 2009), and striped bass (Morone saxatilis; Zmora et al., 2012). Until recently the maximum number of Kiss and Kissr genes was found in an amphibian species, the Xenopus (Xenopus tropicalis), with three paralogs of each gene. On the opposite, in birds (chicken, Gallus gallus, zebra finch, Taeniopygia guttata, and turkey, Meleagris gallopavo) neither Kiss nor Kissr have been found. So far, in all these cases, a matching number of Kiss and Kissr genes had been reported, leading to the suggestion of the occurrence of “paired Kiss/Kissr” systems in vertebrates (Kim et al., 2012).

The recent publications of genomes from representative species of key phylogenetic positions makes it possible to revisit the diversity, the origin and the evolution of Kiss and Kissr in vertebrates. These genomes include a chondrichthyan, the elephant shark (Callorhinchus milii; Venkatesh et al., 2007), a representative of early sarcopterygian, the coelacanth (Latimeria chalumnae, coelacanth genome project, Broad Institute), a non-teleost actinopterygian, the spotted gar (Lepisosteus oculatus; Amores et al., 2011), and an early teleost (elopomorph), the European eel (Anguilla anguilla; Henkel et al., 2012). Gene characterization, phylogenetic, and syntenic analyses allowed us to provide new insights on the respective evolutionary histories of Kiss and Kissr families. Furthermore, the comparison of proposed Kiss and Kissr phylogenetic histories highlighted common processes as well as divergent events leading to the discussion of the existence of conserved Kiss/Kissr couples among the various vertebrate lineages.

MATERIALS AND METHODS

GENOMIC DATABASES

The following genomic databases were investigated:

- the chicken genome
- the coelacanth genome
- the elephant shark genome
- the European eel genome
- the human genome
- the lizard genome
- the platypus genome
- the sea lamprey genome
- the spotted gar genome
- the stickleback genome
- the Xenopus genome
- the zebrafish genome

TBLASTN SEARCH

The TBLASTN algorithm of the CLC DNA Workbench software (CLC bio, Aarhus, Denmark) was used on the European eel genome database and the elephant shark genome database. The TBLASTN algorithm (search sensitivity: near exact matches) of the eGENOMIC website was used on the coelacanth and spotted gar genomic databases.

GENE PREDICTIONS

Considering that Kissr gene structure as well as coding sequences (CDS) are well conserved among vertebrate species, it was possible to predict the exon and intron sequences for new Kissr genes (Pasquier et al., 2012). The splicing junctions were predicted using the empirical nucleotidic splicing signatures, i.e., intron begins with “GT” and ends with “AG.” Concerning Kiss structures, the fact that their CDS are split on two exons (Figure 1), appear to be conserved across vertebrates (for review: Tena-Sempere et al., 2012). However, they are highly variable among species, except for the sequence encoding the Kp(10) localized on the final exon (Figure 1). Therefore, only this Kp(10) conserved sequence can be predicted when investigating new genomes. This small sequence was used to identify the open reading frame (ORF) encompassing a part of the putative Kiss final exon and a part of the putative intron sequence (Figure 1). We therefore focused on this ORF encompassing the sequence encoding Kp(10), in the various genomes. The ORFs of the European eel, the coelacanth, the spotted gar, and the elephant shark, were determined using ORF finder tool of the CLC DNA Workbench software.

SYNTENIC ANALYSES

The synteny analyses of the eel genomic regions were manually performed using CLC DNA Workbench 6 software and the European eel genome database. The analyses of the spotted gar genomic regions were performed using the preliminary gene annotation of the genome assembly LepOcu1 generated by Ensembl release 67. Synteny maps of the conserved genomic regions in human, platypus, lizard (Anolis carolinensis), Xenopus, zebrafish, medaka, stickleback, tetraodon, and coelacanth, as well as of the corresponding region in chicken, G. gallus, were performed using the PhyloView of Genomics v67.01 web site (Muffato et al., 2010).
RESULTS AND DISCUSSION

As one of the aims of this study was to compare the Kiss and Kissr histories, we first propose to make a short overview of our recent findings concerning the diversity, classification, and origin of Kissr gene family. Then, we will expose and discuss our new findings concerning Kiss family. Finally, we will compare and discuss the Kiss and Kissr evolutionary histories in order to get a better understanding of the kisspeptin system evolution.

DIVERSITY AND EVOLUTIONARY HISTORY OF Kissr IN VERTEBRATES

Diversity and classification of Kissr New advances in Kissr gene characterization. Recently, we described three Kissr genes in the genome of a basal teleost, the European eel, providing the first evidence of the existence of three Kissr genes in a teleost species (Pasquier et al., 2012). Furthermore, we described four Kissr in the genome of a non-teleost actinopterygian, the spotted gar, as well as in the genome of a more, we described four Kissr described three recently, we New advances in Kissr gene characterization. Diversity and classification of Kissr IN VERTEBRATES

Phylogeny, synteny, and classification of Kissr

Phylogenetic analysis of 51 peptidic Kissr sequences clustered the osteichthyan Kissr into four clades, each one encompassing a coelacanth and spotted gar Kissr (Pasquier et al., 2012). Clade-1 mainly encompassed mammalian Kissr including human Kissr-2, as well as *Xenopus* Kissr-1, spotted gar Kissr-1, and coelacanth Kissr-1. Clade-2 mainly encompassed teleost Kissr including European eel Kissr-2, as well as amphibian, spotted gar, and coelacanth Kissr-2. Clade-3 encompassed a few teleost Kissr including European eel Kissr-3 as well as *Xenopus* Kissr1b, spotted gar, and coelacanth Kissr-3. Clade-4 encompassed two early osteichthyan (spotted gar and coelacanth) Kissr-4 and two tetrapod Kissr (lizard and platypus; Pasquier et al., 2012).

Synteny analysis of Kissr neighboring genes, performed on 11 representative vertebrate species including the European eel, coelacanth, and spotted gar, fully supported the phylogenetic repartition of Kissr in four clades. Based on this classification, we proposed a new nomenclature of the Kissr family (Kissr-1, Kissr-2, Kissr-3, and Kissr-4; Pasquier et al., 2012).

EOLUVATIONARY HISTORY OF Kissr

Origin of Kissr diversity via 1R and 2R

Synteny analysis revealed that the four Kissr neighboring genomic regions were highly conserved, each presenting paralogs from eight gene families, i.e., PALM, PTBP, GRIN3, GADD45, DIRAS, ZCCHC, LPAR, ZNF644/WIZ (Pasquier et al., 2012). The hypothesis of the potential existence of four Kissr paralogs in vertebrates had been previously raised by Lee et al. (2009) and Kim et al. (2012), although only a maximum of three Kissr genes had been discovered at that time. Our finding of four Kissr genes, located on four paralogous genomic regions, in coelacanth and spotted gar, provides direct evidence validating this former hypothesis. These four Kissr paralogs likely resulted from the two successive genomic duplications (1R and 2R) of a single ancestral genomic region (Figure 2).

The currently available data led to a polytomy of the four Kissr clades. This polytomy did not allow to fully solve the homology relationships between the four Kissr resulting from the 2R (Pasquier et al., 2012). A recent study proposed the phylogenetic reconstruction of the PALM family (Hultqvist et al., 2012). The study of these genes, located in the vicinity of Kissr genes, allows us to infer further relationships between the four Kissr. We can

![Figure 1](https://www.frontiersin.org)
FIGURE 2 | Proposed origin of osteichthyan Kissr tetra-paralogons. The paralogous genes of each of the eight identified families delineate four paralogons in the spotted gar, coelacanth, and human genomes and a duplicated tetra-paralogon in the zebrafish genome. This suggests a common origin of the four Kissr before the two whole genome duplication rounds (1R and 2R) which occurred in the early vertebrate history. This also suggests no impact of the teleost-specific 3R on Kissr number in current teleosts. Chr, chromosome; LG, linkage group.

thus hypothesize that Kissr-1 and Kissr-3, on one side, and Kissr-2 and Kissr-4, on the other side, could be sister genes resulting from the 2R.

Recently, one study proposed the reconstruction of 10 proto-chromosomes of the ancestral vertebrate karyotype and their linkage to the corresponding tetra-paralogons in the human genome (Nakatani et al., 2007). Considering our localization of the four Kissr syntenic regions in the human genome, we can hypothesize that the corresponding tetra-paralogons resulted from the duplications of one single region localized on the proto-chromosome-A of the vertebrate ancestor (Pasquier et al., 2012).

A subsequent history of Kissr losses
Since the spotted gar and the coelacanth are the only two vertebrate species in which we discovered four paralogous Kissr, we can hypothesize multiple Kissr loss events to explain the status of this receptor in current vertebrates (Figure 3A). In the sarcopterygian lineage, Kissr-4 may have been lost in amphibians, while Kissr-1 and Kissr-2 would have been lost in early amniotes. Subsequent additional losses may have led to the presence of only Kissr-1 in eutherian mammals and to the absence of any Kissr in birds.

Considering the presence of four Kissr in a non-teleost actinopterygian, the spotted gar, the teleost-specific 3R could have resulted in the potential existence of up to eight Kissr genes. However, we only found three Kissr in the European eel, representing the current maximum number of this gene in teleosts. Furthermore, each eel Kissr is orthologous to a different tetrapod Kissr, supporting the absence of any teleost-specific Kissr. Synteny analysis demonstrated that each of the four Kissr paralogous genomic regions, present in the spotted gar, was duplicated in zebrafish, in agreement with the 3R. This analysis also indicated that all 3R-copies of Kissr were lacking in the corresponding duplicated regions (Figure 2). This suggests an early loss of duplicated Kissr genes, which would have suppressed the impact of the 3R on the number of Kissr in teleosts (Figure 3A). Additional successive deletions may have led to the presence of three Kissr in a basal teleost (the eel), two Kissr in a cypriniform
Diversity and classification of Kiss genes (A) versus Kissr genes (B). (1) Vertebrates, (2) gnathostomes, (3) osteichthyan, (4) actinopterygians, (5) sarcopterygians, (6) teleosts, (7) tetrapods, (8) amniotes, (9) euteleosts, (10) diapsides, (11) mammals. The names of the current representative species of each phylum are given at the end of the final branches, together with the symbol of the Kiss and Kissr genes they possess. These hypotheses assume the presence of four Kiss and four Kissr paralogs in the vertebrate lineage resulting from the two rounds of vertebrate whole genome duplication. Multiple subsequent gene loss events are indicated in the various lineages.

Diversity and evolutionary history of Kiss in vertebrates

In contrast to the receptor proteins which present several conserved domains, the Kiss genes encode precursors which are highly variable among vertebrates, except for the short sequence of the mature decapptide [Kp(10)]. This makes it difficult to obtain Kiss mRNA sequences by classical molecular strategies. Genomic database analyses thus represent the best approach to characterize the Kiss set for a given species. However, the small characteristic sequence of Kiss, encoding Kp(10), could be missing in genomic databases due to sequencing or assembly limitations.

All previously investigated osteichthyan species possessed the same number of Kiss and Kissr genes: one in eutherian mammals, two in prototherians, three in *Xenopus*, none in birds, and one or two in teleosts (Lee et al., 2009). In cyclostomes, two Kiss genes have been reported (Lee et al., 2009), while only one Kissr could be predicted until now (Pasquier et al., 2011).

Two Kiss genes in the European eel genome. The two ORFs containing the sequences encoding Kp(10) are 296 and 327 bp long, respectively (Figure A1 in Appendix). Once translated, each of them leads to a peptidic sequence encompassing a putative Kp(10): YWNNSFGLRY [European eel Kp1(10)] and FNRNPFGLRF [European eel Kp2(10)], respectively (Figure 4). The C-terminal end of the Kp1(10) sequence is followed by a GKR motif, while the Kp2(10) sequence is followed by a GKR motif (Figure A1 in Appendix). The sequences “X-G-Basic-Basic” or “X-G-Basic” are characteristic of the conserved proteolytic cleavage and alpha-amidation sites of neuropeptides (Eipper et al., 1992).

Two Kiss genes in the spotted gar genome. The two ORFs containing the sequences encoding Kp(10) are 348 and 300 bp long, respectively (Figure A2 in Appendix). Once translated, each of them leads to a peptidic sequence presenting a putative Kp(10): YWNNSFGLRY [spotted gar Kp1(10)] and FNRNPFGLRF [spotted gar Kp2(10)], respectively (Figure 4). The C-terminal ends of these two sequences are followed by a GKR motif (Figure A2 in Appendix).

(zebrafish), and only one Kissr in a more recent teleost (stickleback; Figure 3A).
Three Kiss genes in the coelacanth genome. The three ORFs containing the sequences encoding Kp(10) are 363, 396, and 81 bp long, respectively (Figure A3 in Appendix). Once translated, each of them leads to a peptidic sequence encompassing a putative Kp(10): YNWNTFGLRY [coelacanth Kiss1(10)], FNFNPPFGLRF [coelacanth Kiss2(10)], and FNWNSFGLRF [coelacanth Kiss3(10)], respectively (Figure 4). The C-terminal ends of the Kiss1(10) and Kiss2(10) sequences are followed by a GKR motif, while the Kiss3(10) is followed by a GKK motif (Figure A3 in Appendix). Seven amino-acids up-stream the sequence of the Kiss3(10), a stop codon appears (Figure A3 in Appendix) suggesting that coelacanth Kiss3 gene could have a different intro-exon structure compared to what has been described so far or it can suggest that this gene is no longer expressed.

A third Kiss gene in the elephant shark genome. While two Kiss (Kiss1 and Kiss2) were previously identified in the elephant shark genome (Lee et al., 2009), we were able to localize a new ORF of 315 bp containing a third sequence encoding a Kp(10) (Figure A4 in Appendix). Once translated, it leads to a peptidic sequence encompassing a putative Kp(10): YNLNSFGLKF [elephant shark Kiss3(10)] (Figure 4). The C-terminal end of this peptide is followed by a GKR motif (Figure A4 in Appendix).

Kiss sequence alignment and comparisons. The alignment of 56 kisspeptin precursors revealed a high variability of their amino-acid sequences except for the sequences corresponding to Kp(10) and its few surrounding amino-acids which, in contrast, are highly conserved (data not shown). Such a pattern, which is representative of many other neuropeptide precursors, provides poor phylogenetic information within alignment matrix.

This lack of information makes the use of phylogenetic analysis inappropriate to establish homology relationships within this kind of peptide precursor family.

Novel Kp(10). Among the new Kiss genes predicted in the present study, four of them encode novel Kp(10) (Figure 4). The singularity of the elephant shark Kiss3(10) is the presence of a lysine (K) instead of an arginine (R) at the ninth position. The coelacanth Kiss2(10) provides the first case of a threonine (T) at the fifth position. The coelacanth Kiss3(10) is the only one to present both a phenylalanine (F) at the first position and a serine (S) at the fifth position. The European eel Kp2(10) presents at its third position an arginine (R), which possesses different physical and chemical properties from amino-acids commonly present at this position. Up to now, only the musk shrew (Suncus murinus) Kp1(10) presented an arginine at the third position and it was demonstrated that its kisspeptin system was involved in the reproductive function as in other mammals (Inoue et al., 2011). Since the impacted positions by the amino-acid substitutions have not been characterized as highly critical for Kp(10) functional properties (Gutiérrez-Pascual et al., 2009; Curtis et al., 2010), those novel Kp(10) may have conserved their functionality. Since Kp(10) is considered as the smallest required sequence to specifically bind to the receptor (Kotani et al., 2001), it could be of interest to test all those peptides in future pharmacological studies in order to assess their structure/function relationships.

Syntenic analysis and classification of Kiss genes. In order to classify the different Kiss paralogs, we performed a syntenic analysis of the Kiss neighboring genes. We considered the following vertebrate representatives: mammals (human), birds (chicken), squamates (lizard), amphibians (Xenopus), basal sarcopterygian (coelacanth), non-teleost actinopterygians (spotted gar), and teleosts (zebrafish, stickleback, and European eel). Our syntenic analysis demonstrated that the Kiss genes are localized in three different genomic regions.

The human Kiss1, Xenopus Kiss1a, coelacanth Kiss1, spotted gar Kiss1, and zebrafish Kiss1 are positioned in genomic regions containing common loci, including TEAD3, NAV1, PPPIR12B, PPFA4, MYBPH, KCNC4, REN, GOLI1A, PLEKHA6, PPPIR15B, PIK3C2B, and SYT6, thus exhibiting well conserved synteny (Figure 5A). This supports the orthology of these Kiss genes, all considered as Kiss1 genes. Syntenic analysis suggests that the stickleback, lizard, and chicken genomes do not contain any Kiss1 gene, although the above-mentioned neighboring genes are present in the corresponding genomic regions (Figure 5A). The peptidic sequence of ed Kiss1 presents many similarities to the other Kiss1, but the eel Kiss1 gene is located on too small scaffolds to contain any other gene, preventing from any syntenic analysis.

The lizard Kiss2, coelacanth Kiss2, spotted gar Kiss2, zebrafish Kiss2, stickleback Kiss2, and European eel Kiss2 genes are positioned in genomic regions containing common loci including STRAP, PLEKHA5, GOLI1B, C12orf39, GYS2, LDHB, KCNJ8, ABCB9, CMAS, SYT10, NAV3, PPFA2, and KCNC2, thus exhibiting well conserved synteny (Figure 5B). This supports the orthology of these Kiss genes, all considered as Kiss2 genes. Syntenic analysis suggests that human and chicken genomes do not contain...
FIGURE 5 | Conserved genomic synteny of osteichthyan Kiss. Genomic synteny maps comparing the orthologs of Kiss1 (A), Kiss2 (B), Kiss3 (C), Kiss4 (D), and their neighboring genes. Analyses were performed on the genomes of human (Homo sapiens), platypus (Ornithorhynchus anatinus), lizard (Anolis carolinensis), chicken (Gallus gallus), Xenopus (Xenopus tropicalis), coelacanth (Latimeria chalumnae), spotted gar (Lepisosteus oculatus), zebrafish (Danio rerio), stickleback (Gasterosteus aculeatus), and European eel (Anguilla anguilla). This map was established using the PhyloView of Genomicus v67.01 web site, manual annotation of the European eel genome using CLC DNA Workbench 6 software and the gene annotation of the coelacanth and spotted gar genomic databases. Kiss genes are named according to our proposed nomenclature (Kiss1 to Kiss4). The other genes are named after their human orthologs according to the Human Genome Naming Consortium (HGNC). Orthologs of each gene are shown in the same color. The direction of arrows indicates the gene orientation, with the position of the gene (in 10−6 base pairs) indicated below. The full gene names and detailed genomic locations are given in Table S1 in Supplementary Material. Chr, chromosome; LG, linkage group; Gr, group; Sc, scaffold; Cg, contig.

any Kiss2 gene, although the above-mentioned neighboring genes are present in the corresponding genomic region (Figure 5B).

The coelacanth Kiss3 and the Xenopus Kiss1b genes are positioned in genomic regions containing common loci, including TEAD2, PH1I1D1, and ALDH16A1, thus exhibiting well conserved synteny (Figure 5C). This supports the orthology of these two Kiss genes, both considered here as Kiss3 genes. Syntenic analysis suggests that human, lizard, spotted gar, and teleost genomes do not contain any Kiss3 gene, although the above-mentioned neighboring genes are present in the corresponding genomic regions (Figure 5C). Syntenic analysis also suggests that the whole considered region is absent from the chicken genome.

Evolutionary history of Kiss

Origin of Kiss diversity via 1R and 2R. The syntenic analysis also allowed us to investigate the origin of the multiple Kiss genes found in vertebrates. The three conserved genomic regions, presenting Kiss genes, also comprise other paralogs from 11 gene families: TEAD (4 paralogs), NAV (3 paralogs), PPFLA (4 paralogs), KCNC (4 paralogs), GOLT1 (2 paralogs), PLEKHA (4 paralogs), PPP1R15 (2 paralogs), PIK3C2 (3 paralogs), SYT (4 paralogs), GYS (2 paralogs), and PTH (2 paralogs) (Figures 5A–C). The members of those families are present among the three Kiss syntenic regions and they also delineate a fourth conserved region (Figure 5D). The four conserved regions delineated by the 11 gene families can be considered as paralogous (tetra-paralogon).

Considering the reconstruction of the ancestral vertebrate chromosomes, their linkage to the tetra-paralogs in the human genome (Nakatani et al., 2007) and our localization of the four Kiss syntenic regions in the human genome (on Chromosomes 1, 11, 12, and 19), we can hypothesize that the Kiss tetra-paralogs resulted from the duplications of one single genomic region.
localized on the proto-chromosome-D of the vertebrate ancestor. Therefore, we can infer that the current three Kiss genes may have resulted from a single ancestral gene duplicated through 1R and 2R that occurred in early steps of vertebrate evolution (Figure 6).

**A subsequent history of Kiss losses.**

Multiple loss events in sarcopterygians and actinopterygians. The 1R and 2R events should have resulted in four different Kiss genes in vertebrates. Since the fourth Kiss gene (referred to as Kiss4 in this study) has not been observed in any species studied so far, we can hypothesize an early loss of this gene after the 2R. As only a chondrichthyan, the elephant shark, and two sarcopterygians, the coelacanth and Xenopus, still present three Kiss genes, whereas all other species possess less than three Kiss, we can hypothesize multiple additional events of Kiss losses in vertebrates (Figure 3B).

Among the sarcopterygian lineage, in tetrapods, Kiss3 would have been lost in amniotes. Further alternative losses may have occurred in this lineage, with only Kiss1 remaining in eutharian mammals and only Kiss2 in squamates (lizard). Finally, additional losses would have led to the complete absence of Kiss in birds (Figure 3B). Among the actinopterygian lineage, an early loss of Kiss3 would have occurred since it is lacking in the actinopterygian species investigated so far (Figure 3B).

No impact of the teleost-specific 3R on Kiss number in current species. In the actinopterygian lineage, the teleost-specific 3R and the presence of two Kiss in a non-teleost actinopterygian, the spotted gar, implied the potential existence of at least four Kiss genes in the early teleost history. However, our study showed that so far the largest number of Kiss exhibited by current teleosts, including the eel, is two. Furthermore, each teleost Kiss is orthologous to a different tetrapod Kiss, indicating that no teleost-specific Kiss exists. Synteny analysis revealed that each of the four Kiss genomic regions present in the spotted gar is duplicated in teleosts in agreement with the 3R event but that duplicated Kiss genes are lacking (Figures 5 and 6). This suggests an early loss of duplicated Kiss genes suppressing the impact of the 3R on the number of Kiss in teleosts (Figure 3B). Additional deletion may have led to the presence of only Kiss2 in gasterosteiforms (the stickleback; Figure 3B). Kiss evolutionary history was punctuated by numerous loss events through vertebrate radiation (Figure 3B).

**COMPARISON OF THE EVOLUTIONARY HISTORIES OF Kiss AND Kissr IN VERTEBRATES**

These new data concerning Kiss and Kissr diversities enabled us to improve their respective classifications and evolutionary histories. A remaining challenge was to elucidate whether Kiss and Kissr families have experienced parallel histories during vertebrate radiation. The comparative study of the current status of both families in vertebrates allows a better understanding of the whole kisspeptin system.

Features in agreement with parallel histories

Origin of the Kiss and Kissr multiplicity via 1R and 2R. Our syntenic studies suggest that the vertebrate Kiss and Kissr families both resulted from the successive duplications of a single ancestral gene through the 1R and 2R (Figure 3). Thus, Kiss and Kissr experienced, in parallel, the two first genome duplication rounds resulting in four copies of each gene in the early steps of the vertebrate evolutionary history (Figure 3). While Kiss gene homologs were characterized in non-vertebrate species (Strongylocentrotus purpuratus, GenBank accession numbers: XP_793873.1 and XP_796286.1; Saccoglossus kowalevskii, GenBank accession numbers: NP_001161573.1 and NP_001161574.1), tracing back the presence of an ancestral Kissr in early deuterostomes, Kiss genes have not yet been discovered in non-vertebrate species.

Subsequent history of gene losses. Both Kiss and Kissr families were composed of four genes in the early steps of the vertebrate history. However, most of the current vertebrate species investigated so far present less than four copies of Kiss and Kissr genes. The current numbers of both Kiss and Kissr genes suggest that both families underwent numerous independent loss events across vertebrate history (Figure 3).

No impact of the teleost-specific 3R. The teleost lineage, which has experienced a third whole genome duplication round (3R), could have been expected to possess up to eight Kiss and Kissr genes. However, the analyses of the Kiss and Kissr within teleost genomes revealed a maximum of three Kissr and two Kiss genes and did not reveal any 3R-specific copies of Kiss or Kissr genes. This suggests that the teleost-specific 3R did not impact the current number of Kiss or Kissr genes, reflecting massive losses of the 3R-copies of both Kiss and Kissr genes in early teleosts (Figure 3).

Features in opposition to parallel histories: independent loss events

Un-matching number of Kiss and Kissr in some species. In the current gnathostomes, we observed a maximum of four Kissr but only three Kiss paralogs. This difference suggests that this lineage inherited the four Kissr copies resulting from the 1R and 2R, whereas the fourth Kiss may have been lost before or at an early stage of the gnathostome emergence (Figure 3). This situation was observed in an early sarcopterygian, the coelacanth, while an even larger difference in Kissr (four) and Kiss (two) numbers was found in the spotted gar, reflecting an additional independent loss of Kiss in the actinopterygian lineage. An un-matching number of Kissr (three) and Kissr (two) was also observed in an early teleost, the European eel, while additional losses may have led to equal numbers of Kissr and Kissr in more recent teleosts (two or one according to the species). These variations in Kissr/Kissr numbers reflect different timing of Kiss and Kissr loss events. Those hypotheses suggest that Kiss losses occurred independently among the different gnathostome lineages and also independently from the Kissr losses.

Various Kiss/Kissr combinations across vertebrates. The hypothesis of independent Kiss and Kissr evolutionary histories is also strengthened by the comparison of the gene sets present in species with even matching numbers of Kiss and Kissr. For example, lizard, and stickleback both present the Kiss2 gene, whereas they possess different Kissr, i.e., Kissr-4 in the lizard and Kissr-2 in the stickleback (Figure 3). The same observation can be done comparing the kisspeptin systems of platypus and zebrafish, which both present the Kiss1 and Kiss2 genes, whereas their sets of receptors are completely different, i.e., Kissr-1 and Kissr-4 in platypus versus Kissr-2...
and Kissr-3 in zebrafish (Figure 3). These observations strongly suggest a large variety of Kiss and Kissr combinations, resulting from independent loss events. These data shed new lights on the evolution of the kisspeptin system in vertebrates and challenge the former hypothesis of a conservation of Kiss/Kissr couples across vertebrate evolution. This diversity among vertebrates opens new research avenues for comparative physiology and endocrinology of kisspeptin system.

**What could have favored the independent evolutions of Kiss and Kissr?** A few in vitro studies using recombinant receptors have showed cross-reactivities between various kisspeptins and kisspeptin receptors (Biran et al., 2008; Lee et al., 2009; Li et al., 2009). For instance, Lee et al. (2009) tested the specificity of recombinant human GPR54 (Kissr-1 according to our nomenclature), zebrafish GPR54-1, and -2 (Kissr-3 and Kissr-2 according to our nomenclature), and Xenopus GPR54-1a, -1b, and -2 (Kissr-1, Kissr-4, and Kissr-2 according to our nomenclature) toward various kisspeptins. They showed that human, zebrafish, and Xenopus kisspeptins were able to activate all the receptors with differential intra and inter-specific ligand selectivity. Such cross-reactivity could have promoted the independence of the Kiss and Kissr evolutionary histories. This could explain the situation of species presenting un-matching numbers of Kiss and Kissr genes, as well as the high variability of Kiss/Kissr gene combinations across vertebrate species. Another study, using goldfish recombinant GPR54a and GPR54b (Kissr-3 and Kissr-2 according to our nomenclature), revealed that the ligand potency strikingly differed depending on the responsive element used in the reporter gene construction (Li et al., 2009). These data showed the difficulty to define specific Kiss/Kissr couples based only on pharmacological properties.
In the case of the presence of multiple Kiss/Kissr genes in a given species, anatomical relationships between projections of Kiss neurons and target cells expressing Kissr may provide further cues for determining Kiss/Kissr functional couples. Thus, in zebrafish which possess two Kiss genes and two Kissr genes, in situ hybridization and immunocytochemical studies localized the Kiss1 neurons in different nuclei from Kiss2 neurons (Servili et al., 2011). Moreover, Kiss1 neurons are projecting to Kissr1 (Kissr-3 according to our nomenclature) expressing cells, while Kiss2 neurons are projecting to Kissr2 (Kissr-2 according to our nomenclature) expressing cells (Servili et al., 2011). This reveals anatomically separated kisspeptin systems and distinct specific Kiss/Kissr functional couples in zebrafish (Servili et al., 2011). In striped bass, another teleost possessing two Kiss genes and two Kissr genes, in situ hybridization and laser capture microscopy coupled to quantitative PCR showed, in contrast, that Kiss1 and Kiss2 were co-expressed in neurons of the hypothalamus, indicating promiscuous Kiss synthesis sites (Zmora et al., 2012). However, the two Kissr of the striped bass were expressed in different brain cells, indicating that the kisspeptin systems are not fully redundant (Zmora et al., 2012). All these data underline the importance of investigating the gene diversity, the anatomical organization and the functional properties of the kisspeptin system in various species, regarding the potential high variability of this system among vertebrates.

CONCLUSION

Kisspeptin system is known to play a role in many physiological processes such as antimitastasis, energy metabolism homeostasis, pregnancy, and puberty onset. Even though this system has been widely studied in the last few years, its diversity and evolutionary history remained unclear. Thanks to the newly published genomes of osteichthyes of phylogenetical positions, we were able to provide new data on the diversity of Kiss and Kissr genes, to clarify the classification of these genes and to bring new insights on the evolutionary history of these gene families. Four Kiss and four Kissr genes may have arisen via the 1R and 2R in early vertebrates. This would have been followed by multiple independent Kiss and Kissr gene loss events in the sarcopterygian and actinopterygian lineages. In particular, due to massive Kiss and Kissr gene losses, no impact of the teleost-specific 3R can be recorded on the number of Kissr or Kiss paralogs in current teleost species. The comparison of both Kiss and Kissr gene status, in the current vertebrates, supports both parallel and independent evolutionary histories of the Kiss and Kissr families across vertebrate radiation. It also underlines a large diversity of Kiss/Kissr possible combinations that needs to be taken into account in future comparative studies.

ACKNOWLEDGMENTS

We thank Dr. B. Quérat (CNRS, Paris, France) for his helpful advices and discussions concerning phylogeny and synteny. We also thank all the different consortiums at the initiative of sequencing, assembly, annotation, and publication of genomes, in particular the consortiums of the European eel genome, the spotted gar genome, the coelacanth genome, the elephant shark genome, and the sea lamprey genome. Our thanks also include the authors of the Ensembl genome browser web site and all web sites providing free access to their genomic databases. We thank the authors of the Genomicus web site providing a very useful and friendly tool for synteny analysis. Jérémy Pasquier is a recipient of a Ph.D. fellowship from the Ministry of Research and Education. This work was supported by grants from the National Research Agency, PUBEETEL N ANR-08-BLAN-0173 to Karine Rousseau and Sylvie Dufour, and from the European Community, Seventh Framework Program, PRO-EEL N No. 245257 to Anne-Gaëlle Lafont and Sylvie Dufour.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/NeuroendocrineScience/10.3389/fendo.2012.00173/abstract

Table S1 | Names, references, and locations of the genes used in the Kiss syntenic analysis.

| Name | Reference |
|------|-----------|
| Amores, A., Catchen, J., Ferrara, A., Fontenot, Q., and Postlethwait, J. H. (2011). Genome evolution and meiotic maps by massively parallel DNA sequencing: spotted gar, an outgroup for the teleost genome duplication. Genetics 188, 799–888. | |
| Amores, A., Force, A., Yan, Y. L., Joly, L., Amemiya, C., Fritz, A., et al. (1998). Zebrafish box clusters and vertebrate genome evolution. Science 282, 1711–1714. | |
| Biran, J., Ben-Dor, S., and Levavi-Sivan, B. (2008). Molecular identification and functional characterization of the kisspeptin/kisspeptin receptor system in lower vertebrates. Biol. Reprod. 79, 776–786. | |
| Cartwright, J. E., and Williams, P. J. (2012). Altered placental expression of kisspeptin and its receptor in pre-eclampsia. J. Endocrinol. 214, 79–85. | |
| Carvajal, A. E., Cooke, J. H., Baxter, J. E., Parkinson, J. R., Batavia, L., Ghati, M. A., et al. (2010). A kisspeptin-10 analog with greater in vivo bioactivity than kisspeptin-10. Am. J. Physiol. Endocrinol. Metab. 298, E306–E303. | |
| deRoux, N., Genin, E., Carel, J. C., Matsuda, F., Chauassin, J. L., and Milgrom, E. (2003). Hypogonadotropic hypogonadism due to loss of function of the KISS1-derived peptide receptor GPR54. Proc. Natl. Acad. Sci. U.S.A. 100, 10972–10976. | |
| Dinh, P., and Boore, J. L. (2011). Genome evolution of spotted gar genome, the coelacanth genome, the elephant shark genome, and the sea lamprey genome. Our thanks also include the authors of the Genomicus web site providing a very useful and friendly tool for synteny analysis. Jérémy Pasquier is a recipient of a Ph.D. fellowship from the Ministry of Research and Education. This work was supported by grants from the National Research Agency, PUBEETEL N ANR-08-BLAN-0173 to Karine Rousseau and Sylvie Dufour, and from the European Community, Seventh Framework Program, PRO-EEL N No. 245257 to Anne-Gaëlle Lafont and Sylvie Dufour. | |
Evolutionary history of Kisspeptins and Kisspeptin receptors

Pasquier et al.

Inoue, N., Sasagawa, K., Ikai, K., Sasaki, T., Tomikawa, J., Oishi, S., et al. (2011). Kisspeptin neurons mediate reflex ovulation in the musk shrew (Suncus murinus). Proc. Natl. Acad. Sci. U.S.A. 108, 17527–17532.

Kasahara, M., Naruse, K., Sasaki, S., Nakatani, Y., Qu, W., Ahsan, B., et al. (2007). The medaka draft genome and insights into vertebrate genome evolution. Nature 447, 714–719.

Kim, D. K., Cho, E. B., Moon, M. J., Kasahara, M., Naruse, K., Sasaki, S., Inoue, N., Sasagawa, K., Ikai, K., Sasaki, N., Pasquier, J., Lafont, A. G., Leprince, J., Vaudry, H., Rousseau, K., and Dufour, S. (2011). First evidence for a direct inhibitory effect of kisspeptins on LH expression in the eel, Anguilla anguilla. Gen. Comp. Endocrinol. 173, 216–225.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 05 October 2012; accepted: 07 December 2012; published online: 26 December 2012.

Citation: Pasquier J, Lafont A-G, Toistint V, Vaudry H, Rousseau K and Dufour S (2012) Comparative evolutionary histories of kisspeptins and kisspeptin receptors in vertebrates reveal both parallel and divergent features. Front. Endocrin. 3:173. doi: 10.3389/fendo.2012.00173

This article was submitted to Frontiers in Neuroendocrine Science, a specialty of Frontiers in Endocrinology. Copyright © 2012 Pasquier, Lafont, Toistint, Vaudry, Rousseau and Dufour. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.
**APPENDIX**

**A European eel Kiss1 ORF (Contig 384214.1)**

1 - TCG ACC GAG GCC CCG CCC TAC CGG CAC CGC CCC GCC ACC TGG ACC GCC AAT CCA CTG GCC ATG GCA CTG CAC TGG CCC ACC GCA CTG GCC AAT CCA CTG - 66

1 - S T E A P Y P H R P A H W P S T A L A N P L - 22

67 - CCG GGA GGG GCC GGC CTT CCC ACG AGG TCG TGG TGG TGG TGG TAC CCC GAG GTG CCG CCA CAA GCG GCC - 132

23 - P G G A G L P P R S W W W Y P E V P P Q A A - 44

133 - AAG AAG CGG GAA AAC TTC TCC TCC TAC ACG TGG AAC TCT TTC GGT TTA CGT TAT GGG AAA TGA - 195

45 - K K R E N F S S Y N W N S F G L R Y G K * - 65

**B European eel Kiss2 ORF (Scaffold 13.1)**

1 - TTT ATT AAC ATG CTT TGT GTT CAT CGT CTA GGT TCT CTG TAT CCA GAG CGT GCG TCT GCT GGG GTC - 66

1 - F I N M L C V H R L G S L Y P E R A S A G V - 22

67 - TGG GAC CGC TCT AAA GTA CTG AGA GAA GTC GTG CGG AAT TTA TCA GAC GAA GCA GAT CTC TGC - 132

23 - W D R S K V L R E V T G P N Y S D E A D L C - 44

133 - TTC TTC CTG AAA GAC AGT GAA ATC GAG GCC CAC ATT TCC TGC CGA CTC CTA CTC TCT CTA AGC AAA - 198

45 - F F L K D S E I E G H I S C R L R Y S R R S K - 66

199 - TTC ACG CAC AAT CGG TCT GCT CTC GCT TTT AAA GAA CGT GAG TGG AGT TAC CTT CCC AAG AGC AAA - 264

67 - F N R N F P F G L R F G K R E W S Y L P K S K - 88

265 - ACT GCA AAA CCC GGG GCA AGC AAG CTT CTC GCG TAT CTA CGT TAC ATC CAA GAG CGG AAA GCC TGA - 330

89 - T A K P G T S K L L P Y L L Y I Q E R K A * - 110

**FIGURE A1 | Prediction of two Kiss ORFs from the European eel genome.** Nucleotide and deduced amino-acid sequences of the ORF encoding the European eel Kp1(10) (A) and Kp2(10) (B). Nucleotides (top) are numbered from 5' to 3'. The amino-acid residues (bottom) are numbered beginning with the first amino-acid residue encoded by the ORF. The asterisk (*) indicates the stop codon. The predicted Kp(10) are underlined. The amino-acids of the C-terminal α-amidation and cleavage site are shaded in gray.
A  Spotted gar Kiss1 ORF (LG3)

1  -  AAC  CAG  AAC  TCG  GCA  TTA  TGG  TCA  GCC  CTG  CAT  TTG  CCA  AAT  ATG  TTT  CCG  TCC  TTT  CCA  GTG  -  66
1  -  N  Q  N  S  A  L  W  S  A  L  H  L  P  N  H  M  F  P  S  F  P  V  -  22
67  -  TCG  GGG  AAG  AAG  CAG  GAG  CCC  GCC  GTG  CAG  GAC  ATA  CTG  AGA  CAG  ATG  AAG  ACC  CCG  CCA  -  132
23  -  S  G  K  R  L  P  E  P  V  Q  D  I  L  R  R  M  S  T  P  P  P  -  44
133  -  GGA  GCT  GCC  CTG  ATC  CTG  CCT  GCC  GCT  GCC  CTG  GCC  AGC  CTG  CTG  TTC  GGG  -  198
45  -  G  A  R  L  I  L  P  A  A  G  K  I  P  P  A  L  A  S  L  L  F  G  -  66
199  -  TCT  CGG  TCC  CCC  AGG  AAG  GCC  TGG  GCT  CAG  CAG  AGA  CCC  CAG  CCC  CTT  GCG  AAA  CCG  GAG  -  264
67  -  S  R  F  P  P  R  G  W  A  Q  A  R  P  Q  P  P  A  A  K  R  E  -  88
265  -  AAG  AAC  CTT  TCC  GCC  TAC  AAC  TGG  AAT  TCC  TTC  GGG  CTC  CGT  TAC  GGA  AAA  CCG  GCG  TCA  AAC  ACC  -  330
89  -  K  N  L  S  A  V  N  W  N  S  F  G  L  R  V  G  K  R  R  S  N  T  -  110
331  -  CCG  CGG  CCG  CAG  CTC  GCC  TGA
111  -  P  P  P  Q  L  G  *  -  117

B  Spotted gar Kiss2 ORF (LG8)

1  -  GCA  CCT  GTG  CAC  GCA  GCC  AGC  ACA  CCT  GCC  GGT  CTC  GGA  AGC  CTG  CCC  CAC  CTG  GCC  GTC  AGG  GAG  -  66
1  -  A  P  V  H  A  A  S  T  P  A  G  L  G  S  L  P  H  L  A  V  R  E  -  22
67  -  GTT  GAA  GGT  CCA  AAT  TCG  GCA  GAC  CAG  GCG  AGT  CTC  TGC  TAC  TTG  GTC  CAA  GAG  AGC  AAA  GTG  GAG  -  132
23  -  V  E  G  P  N  S  A  D  Q  A  S  L  C  Y  F  V  Q  E  S  E  V  E  -  44
133  -  AGT  CAG  ATC  TCC  TCC  AAA  TCC  CAG  CTG  TTT  AAC  AGT  AAA  TTT  ACC  TCC  AAT  TTG  CCA  GTT  CCC  -  198
45  -  S  Q  I  S  C  K  L  R  F  T  R  S  K  F  N  F  N  P  F  G  L  R  -  66
199  -  TTC  GGA  AAG  CCG  AAC  AGG  AAT  ATC  GTC  GCC  AAC  AGC  AGG  ACC  GCT  ATA  CCG  AGT  GAG  CTG  TGG  CTT  -  264
67  -  F  G  K  R  N  R  N  I  V  A  N  D  R  S  A  I  P  S  E  L  L  -  88
265  -  TAC  CTG  CTG  TAC  CTC  AAA  GAG  ACG  GCC  CTG  GTC  CCT  TGA
89  -  Y  L  L  Y  L  K  E  T  G  L  A  P  *  -  101

FIGURE A2 | Prediction of two Kiss ORFs from the spotted gar genome. Nucleotide and deduced amino-acid sequences of the ORF encoding the spotted gar Kp1(10) (A) and Kp2(10) (B). Nucleotides (top) are numbered from 5′ to 3′. The amino-acid residues (bottom) are numbered beginning with the first amino-acid residue encoded by the ORF. The asterisk (*) indicates the stop codon. The predicted Kp(10) are underlined. The amino-acids of the C-terminal α-amidation and cleavage site are shaded in gray.
FIGURE A3 | Prediction of three Kiss ORFs from the coelacanth genome.

Nucleotide and deduced amino-acid sequences of the ORF encoding the coelacanth Kp1(10) (A), Kp2(10) (B), and Kp3(10) (C). Nucleotides (top) are numbered from 5' to 3'. The amino-acid residues (bottom) are numbered beginning with the first amino-acid residue encoded by the ORF. The asterisk (*) indicates the stop codon. The predicted Kp(10) are underlined. The amino-acids of the C-terminal α-amidation and cleavage site are shaded in gray.
FIGURE A4 | Prediction of a third Kiss ORF from the elephant shark genome. Nucleotide and deduced amino-acid sequences of the ORF encoding the elephant shark Kp3(10). Nucleotides (top) are numbered from 5' to 3'. The amino-acid residues (bottom) are numbered beginning with the first amino-acid residue encoded by the ORF. The asterisk (*) indicates the stop codon. The predicted Kp1(10) is underlined. The amino-acids of the C-terminal α-amidation and cleavage site are shaded in gray.