



OPEN A bioinformatic approach to characterize the vitellogenin receptor and the low density lipoprotein receptor superfamily in the newt *Cynops orientalis*

Chiara Spinsante^{1,2}, Federica Carducci^{1,2}, Elisa Carotti¹, Adriana Canapa¹, Davide Bizzaro¹, Maria Assunta Biscotti¹✉ & Marco Barucca¹

The Low Density Lipoprotein receptors (LDLRs) gene family includes 15 receptors: very low-density lipoprotein receptor (VLDLR), LDLR, Sorting-related receptor with A-type repeats (SORLA), and 12 LDL receptor-related proteins (LRPs): LRP1, LRP1B, LRP2, LRP3, LRP4, LRP5, LRP6, LRP8, LRP10, LRP11, LRP12, LRP13. Most of these are involved in the transduction of key signals during embryonic development and in the regulation of cholesterol homeostasis. In oviparous animals, the VLDL receptor is also known as VTGR since it facilitates the uptake of vitellogenin in ovary. In tetrapods, information concerning genes encoding these proteins is limited to a few taxa. Here, we report the characterization of VTGR in the amphibian *Cynops orientalis*. The secondary structure analyses and the expression profiles obtained from hepatic and gonadal tissues of *C. orientalis* supported the role of VTGR as vitellogenin oocyte membrane receptor in this species. Moreover, to get a holistic view of the evolutionary history of this gene superfamily, we extended our investigation to all 15 genes belonging to the LDLR superfamily analyzing through a phylogenetic analysis a total of 161 sequences belonging to 11 genera of vertebrates. The position of LRP8 in the tree and its expression findings in *C. orientalis* ovary allowed us to suggest that other proteins of the LDLR superfamily could act as receptors during vitellogenesis.

Keywords Amphibians, Vitellogenin receptor, Basal sarcopterygians, Low density lipoprotein receptors gene family, LDLR, VTGR

In non-mammalian vertebrates, vitellogenin (Vtg) is the precursor of the egg yolk, that is the main nutritional reserve for embryo development¹. This multidomain apolipoprotein is synthesized in female liver of oviparous and ovoviviparous species during reproductive season². Low levels of *vtg* have been recorded also in males²⁻⁵. Additional Vtg functions have been attributed as in egg buoyancy^{6,7}, in host defense⁸⁻¹¹, and toxin transport¹².

In vertebrates, in which multiple *vtg* genes are present^{1,5,13-19}, the Vtg production is regulated by hypothalamic-pituitary-gonadal (HPG) axis that leads to 17 β -estradiol release (E2) by follicular cells. This hormone is recognized by estrogen receptors located in hepatocytes stimulating the synthesis of *vtg* genes²⁰. After post-translational modifications, the produced proteins are released in the bloodstream and reach the oocyte in the form of a homodimeric complex where it is incorporated through Vtg receptor (VTGR)-mediated endocytosis. The VTGR, known also as very-low-density lipoprotein receptor (VLDLR)^{21,22}, is member of the low-density lipoprotein receptor (LDLR) superfamily that includes also LDLR, LRP1, LRP1B, LRP2, LRP3, LRP4, LRP5, LRP6, LRP8, LRP10, LRP11, LRP12, LRP13 and the Sortilin-related receptor with A-type repeats (SORLA)^{23,24}. They are type 1 transmembrane receptors having an extracellular domain made up of cysteine-rich ligand-binding repeats (LBRs) responsible for the ligand-binding activity, a transmembrane domain, and an intracellular domain. Differently from the extracellular domain, the cytoplasmic domain is less conserved except for Asn-Pro-X-Tyr (NPxY) motif involved in adapter protein docking and present in VLDLR, LDLR,

¹Department of Life and Environmental Sciences, Polytechnic University of Marche, Via Brecce Bianche, 60131 Ancona, Italy. ²Chiara Spinsante and Federica Carducci contributed equally to this work. ✉email: m.a.biscotti@univpm.it

LRP8, LRP4, LRP1, LRP1B, LRP2, and SORLA. Moreover, VLDLR, LDLR, and LRP8 show an O-glycosylated domain close to the transmembrane domain. These three proteins share a similar domain organization. However, LRP8 and LDLR present seven LDLR class A domains while VLDLR has eight of these domains²⁵ (Fig. 1). LRP8 and LDLR differ for the length of intracellular domain. Although the members of the LDLR family are primarily involved in binding lipoproteins, other functions have been identified such as in neuronal development²⁶, signaling transduction during embryo development²⁷, and cholesterol homeostasis²⁸. Moreover, the genotypization of some *ldlr* gene SNPs highlighted that these might be associated with a higher risk in the onset of Alzheimer disease, especially in human female individuals²⁹.

The binding site of Vtg for VTGR is in the N-terminal of the lipovitellin I domain. In particular, Li and colleagues (2003)³⁰ have demonstrated that Lys185 located in the 84-amino acid fragment at the N-terminus of Vtg is important for electrostatic interaction with VTGR and that the first three LBRs of the receptor were found to be essential for this ionic interaction. Three-dimensional structure information has been provided by Daly and colleagues (1995)³¹ through two-dimensional NMR spectroscopy. In detail, authors have showed that the N-terminal cysteine-rich repeats are characterized by a beta-hairpin structure, followed by a series of beta-turns. In fish, besides VTGR, LRP13 has been designed as vitellogenin oocyte receptors^{23,32–34}. In tetrapods, VTGRs and the LDLR superfamily have been investigated in a restricted number of taxa³⁵. In detail, the process of Vtg/VTGR endocytosis has been investigated in oviparous amphibians³⁶, while the binding of Vtg/VTGR has been described in the African clawed frog³⁷ and in chicken³⁸. In mammals, the VTGR is known as VLDLR and is widely expressed in the heart, in the skeletal muscle, and in the adipose tissue²⁸.

Due to the paucity of information in amphibians, the aim of this work was to characterize VTGR in the newt *Cynops orientalis*. We extended our investigation to the other 14 genes belonging to the LDLR superfamily, screening the available transcriptomic data of *C. orientalis* and the genomic and transcriptomic resources of 10 other vertebrate species. A total of 161 sequences belonging to 11 genera was included in the phylogenetic analysis carried out to unravel the evolutionary history of the LDLR superfamily. Analysis of secondary structures coupled with the hepatic and gonadal expression profiles in *C. orientalis* supported its role as vitellogenin oocyte membrane receptor in this species. In addition, the position of LRP8 in the tree and its expression in *C. orientalis* ovary suggested that other members of this family could act as receptors during vitellogenesis.

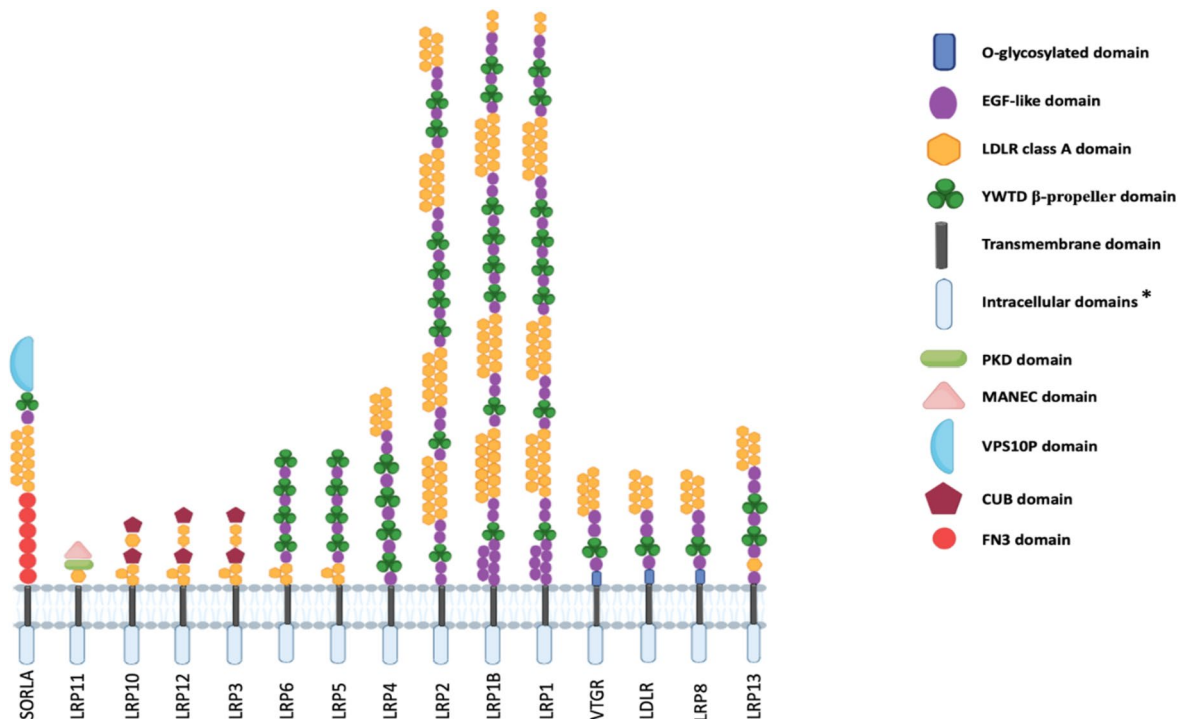


Fig. 1. Domains organization of the LDLR superfamily members. Domains of the low density lipoprotein receptor (LDLR) superfamily members are showed. They are membrane-anchored and contain a transmembrane, a cytoplasmic, and an extracellular domain. The represented domains are: Epidermal Growth Factor (EGF), LDLR class A domain (LDLRA), Tyr-Trp-Thr-Asp (YWTD) β -propeller, Polycystic Kidney Disease (PKD), N-terminus with Eight Cysteines (MANEC), Vacuolar Protein Sorting ten Protein (VPS10P), complement C1r/C1s, Uegf, Bmp1 (CUB), and fibronectin type-III (FN3). Very Low Density Lipoprotein Receptor (VLDLR), LDLR, and Low-Density Lipoprotein Receptor-Related Protein 8 (LRP8) comprise an additional O-glycosylated extracellular domain adjacent to the transmembrane domain. The domain description refers to Principe et al. (2021)²⁴. *Asterisk means that for the intracellular domain are not representative of the real length.

Results

Identification of genes encoding low density lipoprotein receptors in *Cynops orientalis* and phylogenetic analysis

The sequence of VTGR was retrieved from the in-house *de novo* assembled transcriptome of *C. orientalis*. To validate its identity, the investigation was extended to the other 14 genes of the LDLR superfamily. In the newt, 12 sequences were retrieved from our high-quality transcriptome, ten of which were complete. To evaluate the identity and orthology of these sequences 10 other vertebrate species were considered (Supplementary Table S1). Regarding the two basal sarcopterygians, the coelacanth *L. menadoensis* and the lungfish *P. annectens*, in-house *de novo* assembled transcriptomes were screened to identify ortholog sequences. Eight sequences were retrieved for *L. menadoensis* (one of which labelled as incomplete), while the other seven sequences were obtained from the genome of its congeneric *L. chalumnae*. Concerning *P. annectens*, a total of nine complete sequences were obtained, while public repositories were screened for the remaining six sequences. Sequences related to the two actinopterygians and the other six tetrapods were downloaded from public repositories. Details about identification and characterization of sequences analyzed in this study were reported in Supplementary Table S1.

A total of 161 sequences belonging to 11 genera representing the main evolutionary lineages of vertebrates were included in the phylogenetic analysis. The tree topology showed 15 clades corresponding to each member of the LDLR superfamily (Fig. 2). It was noteworthy that for each clade the position of the sequences reflects the evolutionary relationships of the analyzed species. In particular, within each clade, two monophyletic groups can be distinguished: that including actinopterygians (ray-finned fish, *L. oculatus* and *T. rubripes*) and that comprising sarcopterygians (lobe-finned fish and tetrapods). In the case of sarcopterygians, sequences related

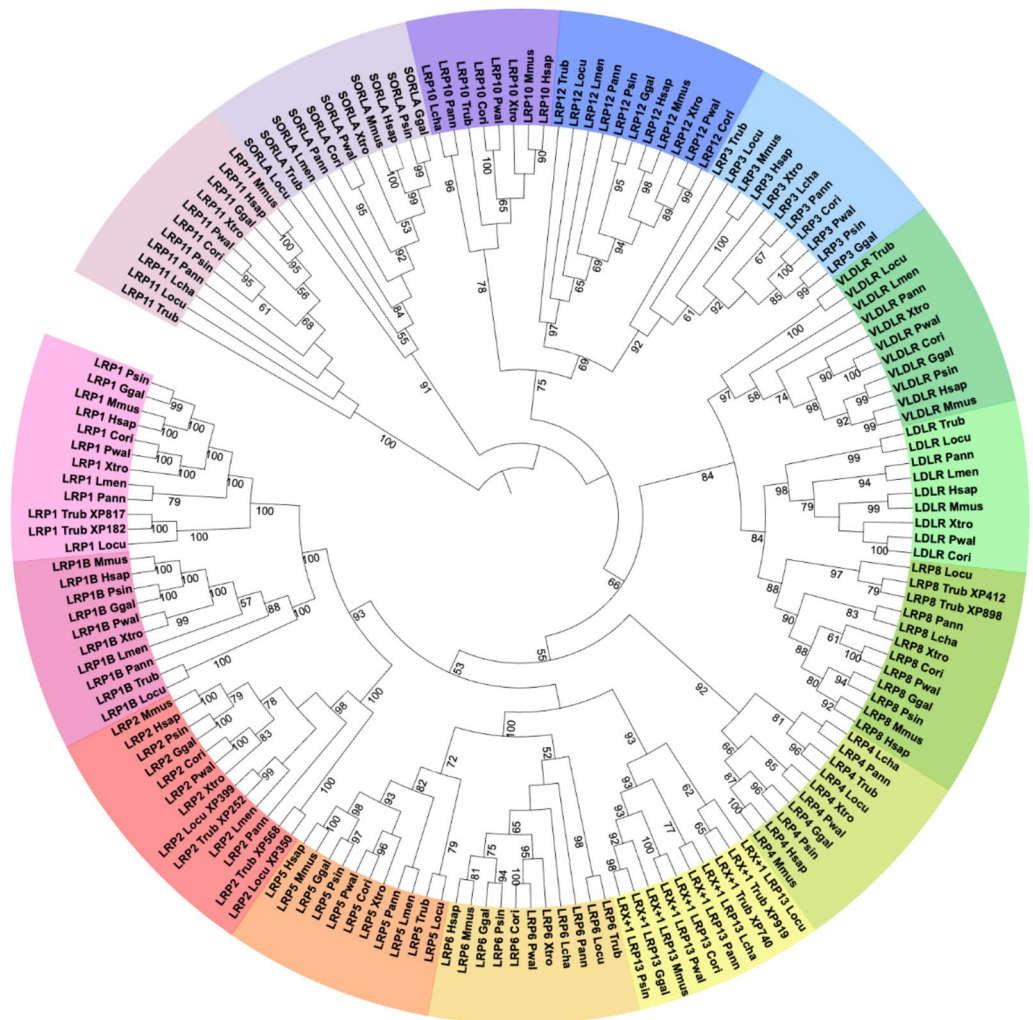


Fig. 2. Maximum Likelihood phylogenetic reconstruction of receptors belonging to the LDLR superfamily. Circular representation of the Maximum Likelihood phylogenetic reconstruction performed on the other members belonging to the Low Density Lipoprotein Receptor (LDLR) superfamily, with bootstrap values. The different colors aimed to highlight the different receptors belonging to the superfamily. Locu: *Lepisosteus oculatus*, Trub: *Takifugu rubripes*, Lmen: *Latimeria menadoensis*, Lcha: *Latimeria chalumnae*, Pann: *Protopterus annectens*, Cori: *Cynops orientalis*, Pwal: *Pleurodeles waltl*, Xtro: *Xenopus tropicalis*, Psin: *Pelodiscus sinensis*, Ggal: *Gallus gallus*, Mmus: *Mus musculus*, Hsap: *Homo sapiens*.

to *Latimeria* and *Protopterus* branched external to those of tetrapods. An exception was represented by the duplicate LRP2 sequences (XP_029693252.1 and XP_015214399.1) of *T. rubripes* and *L. oculatus*, that were placed within the group characterized by sarcopterygians, unlike the sequences XP_029693568.1 of *T. rubripes* and XP_015217350.1 of *L. oculatus* that were placed externally. This might be explained considering the so-called “long branch attraction phenomenon”. Indeed, the sequences of *L. menadoensis* and *P. annectens* were attracted by those of fish due to their high divergence compared to those of the other sarcopterygians. Besides *lrp2*, in *T. rubripes*, the presence of duplicate genes emerged for *lrp1*, *lrp13*, and *lrp8*. In particular, the microsyntenic analyses revealed that the two copies of *lrp1*, *lrp8*, and *lrp13* in *T. rubripes* were located on separate chromosomes while *lrp2* copies were present on the same chromosome in the case of *T. rubripes* and on two distinct linkage groups in the spotted gar (see Supplementary File S1).

In general, the tree topology presented in the external position the sequences related to LRP11 and SORLA. Closer paralog relationships supported by high bootstrap values were highlighted for LRP10/LRP12/LRP3, VLDLR/LDLR/LRP8, and LRP2/LRP1B/LRP1. On the contrary, the clade composed of LRX + 1/LRP5/LRP6 did not show significant bootstrap value indicating that these relationships were not reliable. However, this clade grouped with LRP2/LRP1B/LRP1 and LRP4 and all together shared a common ancestor with VLDLR/LDLR/LRP8. None of the identified sequences belonging to *C. orientalis* was located in the LRP1B and LRP4 clades. Since the gene sequences encoding these two proteins were found in *P. waltl* genome, the other salamander considered, the lack of identification in *C. orientalis* might be due to the absence of transcriptional activity for these two genes in the transcriptomes here analyzed.

Protein structure prediction and comparison

Based on the phylogenetic analysis results, we made the decision to determine the structure of basal sarcopterygian VTGR (*L. menadoensis*, *P. annectens*, and *P. waltl*) and those of LDLR and LRP8 sequences of *C. orientalis*. Models were used to carry out multiple sequence alignments to evaluate the conservation level of secondary structures (Supplementary File S2). Our findings showed that the VTGR in the four basal sarcopterygians was conserved and eight β -hairpin motifs were found at the N-terminal (Supplementary File S2). This structural region appeared to be also conserved in LDLR and LRP8 receptors of *C. orientalis*. However, in the case of LDLR, the lack of the first β -sheet, which constitutes the first β -hairpin structure, was highlighted (Supplementary File S3). Moreover, due to the higher similarity between LRP8 and VLDLR, the protein structural organization of *C. orientalis* LRP8 was compared with orthologs of species considered in phylogenetic analysis highlighting a conservation in terms of protein domain organization (Supplementary File S4).

Gene expression analyses

The expression analysis of the 15 genes belonging to the low-density protein receptors superfamily was conducted in the transcriptomes of *C. orientalis* obtained from female liver and gonadal tissues of both sexes. Overall, in the female gonad the higher transcriptional level was reported for *vldlr*, *lrp8*, *lrp13*, and *ldlr* (Fig. 3). Considering the comparison between female gonad and liver, the highest transcriptional activity was reported for *vldlr* in the gonad, with statistically significant values. It was noteworthy the absence of *lrp8* and *lrp13* in the hepatic tissue (Fig. 4A). Moreover, regarding the gene expression profiles obtained in female and male gonads, statistically significant higher values were obtained for *vldlr*, *lrp8*, and *lrp13* in ovary (Fig. 4B) while for *ldlr* in testis. Concerning the expression values of *vldlr* gene in the three female gonad replicates, the highest transcriptional activity was recorded in the third female individual (Fig. 3).

Discussion

The Vtg receptor, belonging to the LDLR superfamily^{21,39}, is fundamental for the vitellogenin endocytosis within the oocytes⁴⁰. In particular, the receptors belonging to this superfamily are cell surface proteins, important for the transport of various ligands across cell membranes. Most of these are not only involved in endocytosis, but also in the transduction of important signals during embryonic development and in the regulation of cholesterol homeostasis²⁷. In tetrapods, information on the LDLR superfamily and Vtg receptors are limited to a few taxa³⁵. Therefore, gene characterization studies of these receptors in the newt *C. orientalis* and in other basal sarcopterygians might deepen our knowledge on the evolution of these proteins and their roles. Our analysis allowed us to identify 13 complete sequences in *C. orientalis*, eight sequences in *L. menadoensis*, and nine sequences in *P. annectens*. The phylogenetic analysis confirmed the orthology of the candidate sequences to the Vtg and LDLR superfamily receptors. Moreover, the position of these sequences in the phylogenetic tree helped us to unravel the evolutionary history of these genes. Given the presence of all 15 genes belonging to the LDLR superfamily both in actinopterygians and sarcopterygians, this gene set was already present in their common ancestor originated in the Ordovician era (510 mya) and that already experienced two rounds of WGD (1R and 2R)⁴¹. Furthermore, duplicate genes were detected in actinopterygians (*lrp1*, *lrp13*, *lrp8*, and *lrp2*). In particular, microsyntenic analyses revealed that the two copies of *lrp1*, *lrp8*, and *lrp13* in *T. rubripes* were located on separate chromosomes known to be originated by teleost specific WGD 3R estimated to have occurred between 226 and 350 mya⁴¹. Differently, the duplication event of *lrp2* cannot be attributable to the 3R WGD event, given the presence of two copies also in the non-teleost *L. oculatus*. Therefore, these two copies were already present in the common ancestor of spotted gar and pufferfish, in the Devonian era (409 mya)⁴¹.

Interestingly, the domain organization of the receptors considered in the present study corroborated results obtained through phylogenetic analysis (Fig. 1). Indeed, LRP1, 1B, and 2, which grouped together, showed a very similar protein domain organization; the same result was also observed for the groups of receptors LRP3/LRP10/LRP12 and VLDLR/LDLR/LRP8. The monophyletic relationship observed for these two clades has also been highlighted in previous phylogenetic analyses^{23,32,34}. Therefore, the common origin of the considered sequences allowed us to exclude structural similarity due to convergent evolution. Moreover, VLDLR, LDLR, and LRP8

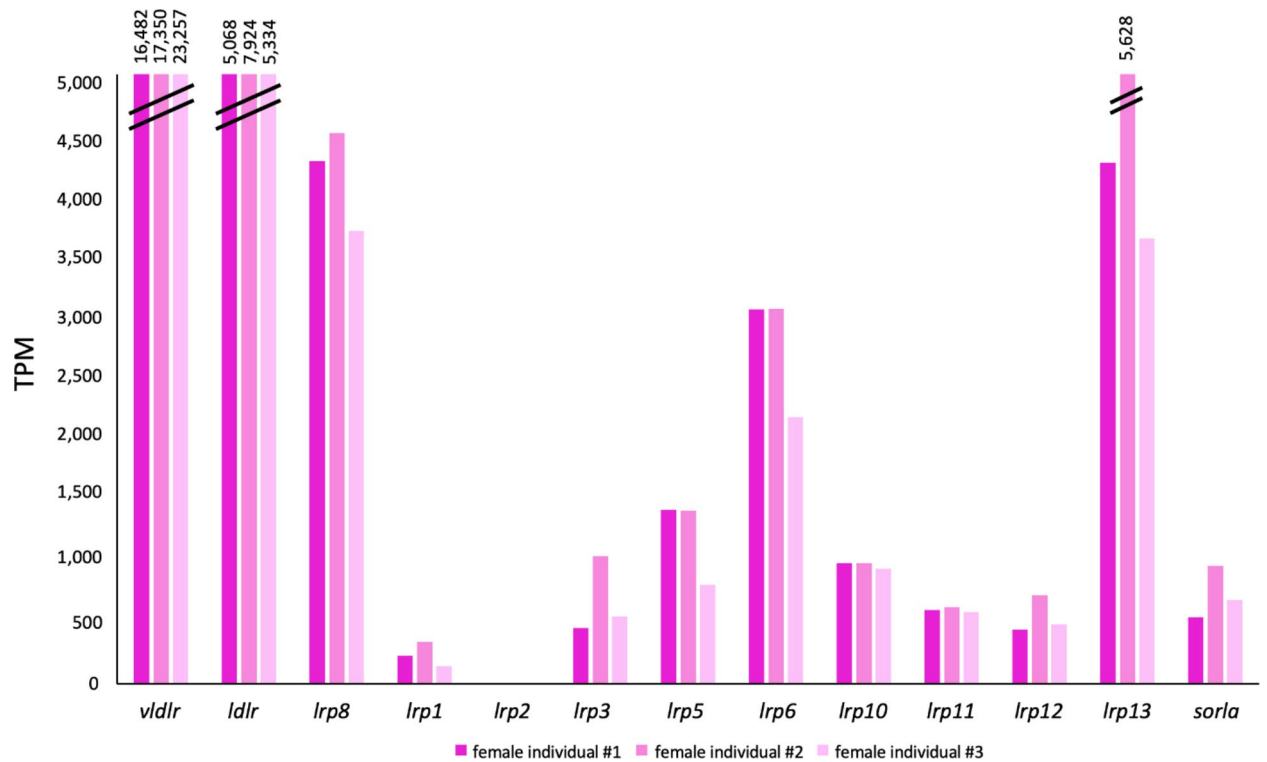


Fig. 3. Transcriptional activity of members belonging to the LDLR superfamily in the three replicates of *Cynops orientalis* female gonad. Expression values of the low density lipoprotein receptor (LDLR) superfamily members obtained for the three replicates of *Cynops orientalis* female gonad. Values are reported as transcripts per million kilobases (TPM).

structural models showed a conserved alignment of secondary structure elements in *C. orientalis*. This suggested that the function of the three proteins could be similar.

LDLR has been reported to be involved in the lipid transfer within maturing oocytes in the salmonid *Oncorhynchus clarki*³³. However, the lack of the first beta sheet in the N-terminus secondary structure element belonging to LBR and known to be fundamental for the interaction between vitellogenin receptor and its ligand^{30,31}, suggested that it might not function as vitellogenin receptor. Moreover, the transcriptional activity of the *ldlr* gene showed comparable levels in the two female tissues and was higher in testis than in ovary in *C. orientalis*. This latter evidence agreed with the recently published papers in which LDLR has been associated with the regulation of lipid homeostasis in normal male reproductive function^{42,43}.

At the gonadal level, LRP8 is involved in the formation of cholesterol storage, also useful for the synthesis of steroids during the folliculogenesis⁴⁴ and in the achieving of the right weight of the yolk and albumen⁴⁵. Recently, Akhavan and colleagues (2020)⁴⁶ have proposed that LRP8 might mediate ovarian yolk and lipid uptake in the beluga sturgeon *Huso huso* L. Therefore, a possible interaction of LRP8 with Vtg or other lipoprotein particles might be hypothesized also in *C. orientalis*.

Indeed, analyzing transcriptomic data, *lrp8* as *vldlr* and *lrp13*, involved in vitellogenesis, showed a greater expression in the female gonad compared to the female liver and the male gonad in the newt. Moreover, the high degree of conservation observed aligning secondary structures of LRP8 between considered species suggested a similar function in oviparous and ovoviviparous species. The different transcriptional activity reported for the three genes *vldlr*, *lrp8*, and *lrp13* might be related to the variable expression levels of the diverse forms of *vtgs*, evidenced by Carducci and collaborators (2021)¹³ for the same specimens here analyzed. Moreover, the different transcriptional activity of the *vldlr* gene in the three female individuals can be explained considering gonadal histological observations⁴⁷ and *vtg* expression levels¹³. Indeed, gonads of the three females were characterized by a clear asynchronous maturation stage⁴⁷. Of these, the third one was reported as less mature than others, in agreement with the lower transcriptional level reported for the *vtg* and the highest value obtained for the *vtg* receptor. This scenario suggested that the gonad is producing receptor to receive Vtg, whose production has just begun in the liver. In *O. mykiss*, the expression of *vtgr* mRNA starts in the pre-vitellogenic stage of oocytes and diminishes at the beginning of the vitellogenesis⁴⁸. On the contrary, the highest transcriptional value detected for the *vtg* and the lower value for the *vtg* receptor were obtained for the first *C. orientalis* female specimen, having mature gonads.

Conclusion

In this work, we report the characterization of the VTGR in the newt *C. orientalis* and an extensive phylogenetic analysis of the LDLR superfamily receptors that helped unravel the evolutionary history of these genes. Our

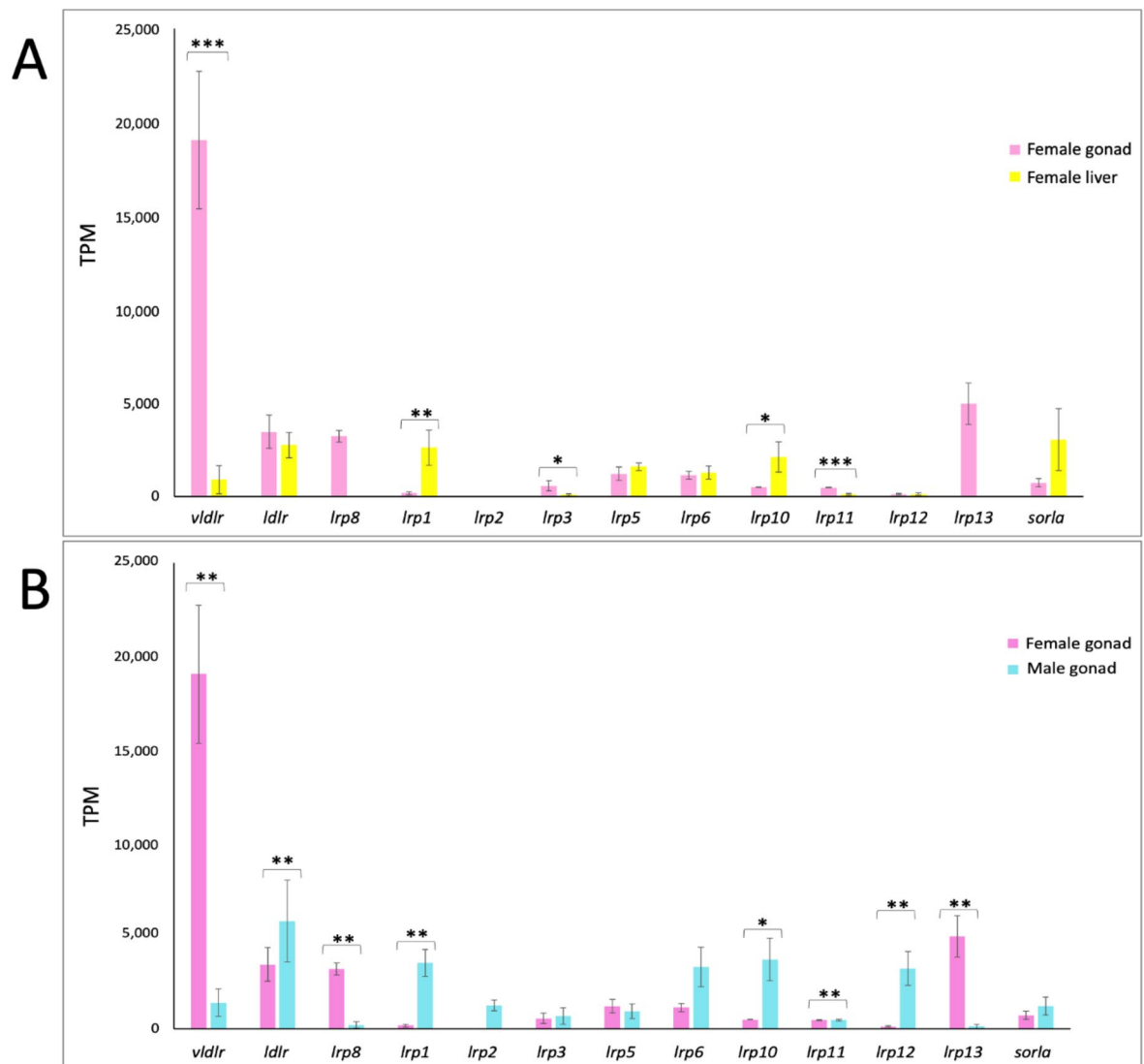


Fig. 4. Transcriptional activity of members belonging to the low density lipoprotein receptor (LDLR) superfamily in *Cynops orientalis* between gonad and liver of female specimens (A) and between female and male gonads (B). (A) Expression values of the 13 members of the Low Density Lipoprotein Receptor (LDLR) superfamily in *Cynops orientalis*. Means and standard deviations were obtained from the triplicates of the female gonad (pink) and female liver (yellow) *Cynops orientalis* transcriptomes. Values are reported as transcripts per million kilobases (TPM). (B) Expression values of the 13 members of the Low Density Lipoprotein Receptor (LDLR) superfamily in *Cynops orientalis*. Means and standard deviations were obtained from the triplicates of the female (pink) and male gonad (light blue) *Cynops orientalis* transcriptomes. Values are reported as transcripts per million kilobases (TPM). Statistically significant differences are presented as * for $p > 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$.

results revealed that the gene set of this superfamily was present in the common ancestor of sarcopterygians and actinopterygians and remained conserved throughout these evolutionary lineages. Moreover, combining phylogenetic analysis, secondary structure predictions, and gene expression analyses, our findings supported the role of VTGR as vitellogenin oocyte membrane receptor in the newt. The close relationship of LRP8 with VLDLR in the tree and expression of these genes in *C. orientalis* ovary allowed us to propose that also LRP8 could be involved in vitellogenesis.

Materials and methods

Vitellogenin receptor sequences identification and characterization

In this study, 11 genera belonging to 10 orders were considered: Coelacanthiformes (*L. menadoensis* and *L. chalumnae*), Ceratodontiformes (*P. annectens*), Lepisosteiformes (*Lepisosteus oculatus*), Tetraodontiformes (*Takifugu rubripes*), Caudata (*C. orientalis* and *P. waltl*), Anura (*Xenopus tropicalis*), Testudines (*Pelodiscus sinensis*), Galliformes (*Gallus gallus*), Rodentia (*Mus musculus*), and Primates (*Homo sapiens*).

Analyses on *L. menadoensis*, *P. annectens*, and *C. orientalis* were carried out using transcriptomes obtained by the Molecular Phylogenetic laboratory of the Polytechnic University of Marche^{47,49,50} (accession numbers: PRJNA175365, PRJNA164839, PRJNA574599, respectively); while, for the remaining organisms, the search was performed using the transcriptomes available in NCBI (<https://www.ncbi.nlm.nih.gov>) and the genomic information present in ENSEMBL (<https://www.ensembl.org/index.html>). In detail, for *C. orientalis* a search using BLAST + v2.11.0⁵¹ was performed, setting protein sequences of the evolutionary related species *X. tropicalis* as queries. Nucleotide sequences corresponding to subjects with the highest score were then translated and the correctness checked through a BLASTP search. The same procedure was used for *L. menadoensis* and *P. annectens*. The coding sequence of each transcript identified as potential candidate of this gene superfamily was then characterized using ExPasy translate tool (SIB, www.expasy.org). To complete the sequences of *L. menadoensis* not retrieved at the transcriptome-level, the genome of *L. chalumnae*, available in NCBI, was used (GCF_000225785.1). Moreover, the analysis of this superfamily was extended to other organisms: *L. chalumnae*, *L. oculatus*, *T. rubripes*, *P. waltl*, *X. tropicalis*, *P. sinensis*, *G. gallus*, *M. musculus*, and *H. sapiens* (Supplementary Table S1). Sequences obtained in this study were deposited at GenBank under the accession numbers PP516878:PP516907. Selected sequences were then used to generate preliminary phylogenetic trees, to confirm their sequence orthology.

Phylogenetic analysis

In order to identify and evaluate evolutionary distances of VTGRs of organisms considered in this study, a phylogenetic analysis was conducted. Sequences obtained from the transcriptomes of *L. menadoensis*, *P. annectens*, and *C. orientalis* were aligned with another 131 sequences of the other organisms included in the present study, retrieved in NCBI and ENSEMBL public repositories (the complete list of accession numbers of sequences used was reported in Supplementary Table S1). The alignment was performed with ClustalW⁵² with default parameters. The phylogenetic analysis was carried out with RAxML-NG⁵³ with 1000 bootstraps. ITOL was used for tree visualization⁵⁴. The JTT + G4 model was defined as the most appropriate model using ModelTest-NG⁵⁵.

Microsyntenic analysis in actinopterygians and basal sarcopterygians

The microsyntenic arrangement of *lrp1*, *lrp2*, *lrp8*, and *lrp13* genes was obtained combining information retrieved from NCBI Gene, Genomicus, and Ensembl to determine upstream and downstream flanking genes of those of interest.

Protein structure prediction and comparison

The three-dimensional (3D) model prediction for VTGR, LDLR, and LRP8, was built using Swiss-Model⁵⁶ web server, available online (<http://swissmodel.expasy.org>). Global Model Quality Estimate (GMQE) and Quality Model Energy ANalysis (QMEAN) parameters were used to have an estimate of the quality of the model globally and per residue. The 3D predictions were used to determine the secondary structures present along the aminoacidic sequences of the proteins analyzed. The secondary protein structures obtained for VTGR, LDLR, and LRP8 were compared using 2DSS server⁵⁷.

Gene expression analyses

C. orientalis trimmed reads of RNA-seq data of the three female livers and the three male and female gonads were mapped against *ldlr* superfamily genes in CLC Genomics Workbench v.12 environment (Qiagen, Hilden, Germany) using the RNA-seq mapping tool, setting parameters as highly stringent (length fraction 0.75 and similarity fraction 0.8). Gene expression levels were computed as Transcripts Per Million (TPM), as this metric allows to efficiently compare gene expression levels both within and between samples. We followed the same strategy described in Biscotti et al. (2016)⁵⁰. In summary, trimmed sequencing reads were aligned to the transcriptome using RNA-seq mapping tools in CLC Genomics Workbench v.24.0.1 (Qiagen, Hilden, Germany). The mapping parameters were set at length fraction and a similarity fraction of 0.80. Gene expression levels, reported as transcripts per million (TPM), were calculated using a scaling factor derived from the cumulative expression of 3,354 unequivocal single-copy orthologs shared between analyzed tissues. Statistically significant differences were determined using one-way ANOVA, with significance denoted by * for $p > 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$.

Data availability

The datasets generated during the current study are available in the Genbank repository under the accessions PP516878:PP516907. Details about all data analysed during the current study are provided in Supplementary Table S1.

Received: 16 July 2024; Accepted: 23 January 2025

Published online: 27 January 2025

References

1. Finn, R. N. & Kristoffersen, B. A. Vertebrate vitellogenin gene duplication in relation to the 3R hypothesis: correlation to the pelagic egg and the oceanic radiation of teleosts. *PLoS One* **2**, e169 (2007).
2. Verderame, M. & Scudiero, R. Estrogen-dependent, extrahepatic synthesis of vitellogenin in male vertebrates: a mini-review. *C. R. Biol.* **340**, 139–144 (2017).
3. Barucca, M. et al. Characterization and phylogenetic analysis of vitellogenin coding sequences in the Antarctic fish *Trematomus bernacchii*. *J. Exp. Zool. B Mol. Dev. Evol.* **314**, 645–652 (2010).

4. Canapa, A. et al. Vitellogenin gene expression in males of the Antarctic fish *Trematomus bernacchii* from Terra Nova Bay (Ross Sea): a role for environmental cadmium? *Chemosphere* **66**, 1270–1277 (2007).
5. Canapa, A. et al. Composition and phylogenetic analysis of vitellogenin coding sequences in the Indonesian coelacanth *Latimeria menadoensis*. *J. Exp. Zool. B Mol. Dev. Evol.* **318**, 404–416 (2012).
6. Finn, R. N., Østby, G. C., Norberg, B. & Fyhn, H. J. In vivo oocyte hydration in Atlantic halibut (*Hippoglossus hippoglossus*); proteolytic liberation of free amino acids, and ion transport, are driving forces for osmotic water influx. *J. Exp. Biol.* **205**, 211–224 (2002).
7. Kolarevic, J., Nerland, A., Nilsen, F. & Finn, R. N. Goldsinny Wrasse (*Ctenolabrus rupestris*) is an extreme vtgaa-type pelagophil teleost. *Mol. Reprod. Dev.* **75**, 1011–1020 (2008).
8. Zhang, S., Wang, S., Li, H. & Li, L. Vitellogenin, a multivalent sensor and an antimicrobial effector. *Int. J. Biochem. Cell. Biol.* **43**, 303–305 (2011).
9. Zhang, S., Dong, Y. & Cui, P. Vitellogenin is an immunocompetent molecule for mother and offspring in fish. *Fish. Shellfish Immunol.* **46**, 710–715 (2015).
10. Sun, C. & Zhang, S. Immune-relevant and antioxidant activities of vitellogenin and yolk proteins in fish. *Nutrients* **7**, 8818–8829 (2015).
11. Sattar Khan, M. A. et al. Bactericidal action of egg yolk phosvitin against *Escherichia coli* under thermal stress. *J. Agric. Food Chem.* **48**, 1503–1506 (2000).
12. Yin, X. et al. A novel function of vitellogenin subdomain, vWF type D, as a toxin-binding protein in the pufferfish *Takifugu pardalis* ovary. *Toxicol.* **136**, 56–66 (2017).
13. Carducci, F., Biscotti, M. A., Canapa, A. & Barucca, M. The vitellogenin genes in *Cynops orientalis*: new insights on the evolution of the vtg gene family in amphibians. *J. Exp. Zool. B Mol. Dev. Evol.* **336**, 554–561 (2021).
14. Biscotti, M. A., Barucca, M., Carducci, F. & Canapa, A. New perspectives on the evolutionary history of vitellogenin gene family in vertebrates. *Genome Biol. Evol.* **10**, 2709–2715 (2018).
15. Andersen, Ø. et al. Resolving the complexity of vitellogenins and their receptors in the tetraploid Atlantic salmon (*Salmo salar*): ancient origin of the phosvitin-less VtgC in chondrichthyan fishes. *Mol. Reprod. Dev.* **84**, 1191–1202 (2017).
16. Yamane, K. et al. Characterization of vitellogenin and its derived yolk proteins in cloudy catshark (*Scyliorhinus torazame*). *Fish. Physiol. Biochem.* **39**, 373–390 (2013).
17. Babin, P. J. Conservation of a vitellogenin gene cluster in oviparous vertebrates and identification of its traces in the platypus genome. *Gene* **413**, 76–82 (2008).
18. Brawand, D., Wahli, W. & Kaessmann, H. Loss of egg yolk genes in mammals and the origin of lactation and placentation. *PLoS Biol.* **6**, e63 (2008).
19. Reading, B. J. et al. Conserved and variant molecular and functional features of multiple egg yolk precursor proteins (vitellogenins) in white perch (*Morone americana*) and other teleosts. *Mar. Biotechnol. (NY)*. **11**, 169–187 (2009).
20. Mushirobira, Y., Niida, M., Hotta, T., Fujinami, Y. & Soyano, K. Hepatic expression profiles of three subtypes of vitellogenin and estrogen receptor during vitellogenesis in cultured female yellowtail. *Gen. Comp. Endocrinol.* **299**, 113612 (2020).
21. Bujo, H. et al. Chicken oocyte growth is mediated by an eight ligand binding repeat member of the LDL receptor family. *EMBO J.* **13**, 5165–5175 (1994).
22. Hiramatsu, N. et al. Multiple ovarian lipoprotein receptors in teleosts. *Fish. Physiol. Biochem.* **39**, 29–32 (2013).
23. Reading, B. J. et al. Lrp13 is a novel vertebrate lipoprotein receptor that binds vitellogenins in teleost fishes. *J. Lipid Res.* **55**, 2287–2295 (2014).
24. Príncipe, C., Dionísio de Sousa, I. J., Prazeres, H., Soares, P. & Lima, R. T. LRP1B: a Giant Lost in Cancer Translation. *Pharmaceuticals (Basel Switzerland)*. **14**, 836 (2021).
25. Dlugosz, P. & Nimpf, J. The Reelin receptors Apolipoprotein E receptor 2 (ApoER2) and VLDL receptor. *Int. J. Mol. Sci.* **19**, 3090 (2018).
26. May, P. & Herz, J. LDL receptor-related proteins in neurodevelopment. *Traffic* **4**, 291–301 (2003).
27. Li, Y., Cam, J. & Bu, G. Low-density lipoprotein receptor family: endocytosis and signal transduction. *Mol. Neurobiol.* **23**, 53–67 (2001).
28. Go, G. W. & Mani, A. Low-density lipoprotein receptor (LDLR) family orchestrates cholesterol homeostasis. *Yale J. Biol. Med.* **85**, 19–28 (2012).
29. Lämsä, R. et al. Genetic study evaluating LDLR polymorphisms and Alzheimer's disease. *Neurobiol. Aging*. **29**, 848–855 (2008).
30. Li, A., Sadasivam, M. & Ding, J. L. Receptor-ligand interaction between vitellogenin receptor (VtGR) and vitellogenin (vtg), implications on low density lipoprotein receptor and apolipoprotein B/E. The first three ligand-binding repeats of VtGR interact with the amino-terminal region of Vtg. *J. Biol. Chem.* **278**, 2799–2806 (2003).
31. Daly, N. L., Scanlon, M. J., Djordjevic, J. T., Kroon, P. A. & Smith, R. Three-dimensional structure of a cysteine-rich repeat from the low-density lipoprotein receptor. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 6334–6338 (1995).
32. Mushirobira, Y. et al. Molecular cloning and partial characterization of a low-density lipoprotein receptor-related protein 13 (Lrp13) involved in vitellogenin uptake in the cutthroat trout (*Oncorhynchus clarki*). *Mol. Reprod. Dev.* **82**, 986–1000 (2015).
33. Luo, W. et al. Molecular cloning and partial characterization of an ovarian receptor with seven ligand binding repeats, an orthologue of low-density lipoprotein receptor, in the cutthroat trout (*Oncorhynchus clarki*). *Comp. Biochem. Physiol. Mol. Integr. Physiol.* **166**, 263–271 (2013).
34. Babio, L., Damsteegt, E. L. & Lokman, P. M. Lipoprotein receptors in ovary of eel, *Anguilla australis*: molecular characterisation of putative vitellogenin receptors. *Fish. Physiol. Biochem.* **49**, 117–137 (2023).
35. Ruan, Y. et al. Vitellogenin receptor (VgR) mediates oocyte maturation and ovarian development in the Pacific white shrimp (*Litopenaeus vannamei*). *Front. Physiol.* **11**, 485 (2020).
36. Wall, D. A. & Patel, S. The intracellular fate of vitellogenin in *Xenopus* oocytes is determined by its extracellular concentration during endocytosis. *J. Biol. Chem.* **262**, 14779–14789 (1987).
37. Okabayashi, K. et al. cDNA cloning and expression of the *Xenopus laevis* vitellogenin receptor. *Biochem. Biophys. Res. Commun.* **224**, 406–413 (1996).
38. Stifani, S., Nimpf, J. & Schneider, W. J. Vitellogenesis in *Xenopus laevis* and chicken: cognate ligands and oocyte receptors. The binding site for vitellogenin is located on lipovitellin I. *J. Biol. Chem.* **265**, 882–888 (1990).
39. Davail, B. et al. Evolution of oogenesis: the receptor for vitellogenin from the rainbow trout. *J. Lipid Res.* **39**, 1929–1937 (1998).
40. Opresko, L. K. & Wiley, H. S. Receptor-mediated endocytosis in *Xenopus* oocytes. I. characterization of the vitellogenin receptor system. *J. Biol. Chem.* **262**, 4109–4115 (1987).
41. Van de Peer, Y., Maere, S. & Meyer, A. The evolutionary significance of ancient genome duplications. *Nat. Rev. Genet.* **10**, 725–732 (2009).
42. Casado, M. E. et al. Hormone-sensitive lipase deficiency affects the expression of SR-BI, LDLr, and ABCA1 receptors/transporters involved in cellular cholesterol uptake and efflux and disturbs fertility in mouse testis. *Biochim. Biophys. Acta Mol. Cell. Biol. Lipids.* **1866**, 159043 (2021).
43. Bai, Y. et al. LDLR heterozygous deletion reduces hamster testicular cholesterol toxicity via AMPK/Sirt1/PGC-1 α pathway. *Toxicol. Lett.* **384**, 30–43 (2023).
44. Argov, N., Moallem, U. & Sklan, D. Lipid transport in the developing bovine follicle: messenger RNA expression increases for selective uptake receptors and decreases for endocytosis receptors. *Biol. Reprod.* **71**, 479–485 (2004).

45. Wang, C. et al. Molecular cloning, expression and association study with reproductive traits of the duck LRP8 gene. *Br. Poult. Sci.* **54**, 567–574 (2013).
46. Akhavan, S. R., Falahatkar, B., McCormick, S. P. A. & Lokman, P. M. Changes in lipid biology during ovarian development in farmed beluga sturgeon, *Huso huso* L. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **319**, 376–386.
47. Biscotti, M. A. et al. The transcriptome of the newt *Cynops orientalis* provides new insights into evolution and function of sexual gene networks in sarcopterygians. *Sci. Rep.* **10**, 5445 (2020).
48. Perazzolo, L. M. et al. Expression and localization of messenger ribonucleic acid for the vitellogenin receptor in ovarian follicles throughout oogenesis in the rainbow trout, *Oncorhynchus mykiss*. *Biol. Reprod.* **60**, 1057–1068 (1999).
49. Pallavicini, A. et al. Analysis of the transcriptome of the Indonesian coelacanth *Latimeria menadoensis*. *BMC Genom.* **14**, 538 (2013).
50. Biscotti, M. A. et al. The lungfish transcriptome: a glimpse into molecular evolution events at the transition from water to land. *Sci. Rep.* **21571** (2016).
51. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
52. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
53. Kozlov, A. M., Darriba, D., Flouri, T., Morel, B. & Stamatakis, A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**, 4453–4455 (2019).
54. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **4**, 242–245 (2016).
55. Darriba, D. et al. ModelTest-NG: a new and scalable tool for the selection of dna and protein evolutionary models. *Mol. Biol. Evol.* **37**, 291–294 (2020).
56. Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, 296–303 (2018).
57. Lotun, D. P., Cochard, C., Vieira, F. R. J. & Bernardes, J. S. 2DSS: a web server for protein secondary structure visualization. *BioRxiv* **649426** (2019).

Author contributions

C.S.: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation. F.C.: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation. E.C.: Writing – review & editing, Formal Analysis, Data curation. A.C.: Writing – review & editing, Supervision. D.B.: Writing – review & editing, Supervision. M.A.B.: Writing – original draft, Resources, Investigation, Conceptualization. M.B.: Writing – original draft, Supervision, Investigation, Formal analysis, Data curation, Conceptualization. All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Funding

This research was funded by Polytechnic University of Marche in the framework of “Progetto Scientifico di Ateneo”.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-88011-6>.

Correspondence and requests for materials should be addressed to M.A.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025