General and Comparative Endocrinology 296 (2020) 113532

Contents lists available at ScienceDirect

General and Comparative Endocrinology

journal homepage: www.elsevier.com/locate/ygcen

Research paper

Identification and structural characterization of the factors involved in vitellogenesis and its regulation in the African Osteoglossiforme of aquacultural interest *Heterotis niloticus* (Cuvier, 1829)

N'Zi Daniel Koua[a,b,c], Jésus Núñez-Rodriguez[d], Julie Orjuela[e], Céline Zatylny-Gaudin[a,c], Marie-Pierre Dubos[a,c], Benoît Bernay[f], Julien Pontin[f], Erwan Corre[g], Joël Henry[a,c,*]

[a] NORMANDIE UNIV, UNICAEN, CNRS, BOREA, 14000 Caen, France
[b] INP-HB, Département FOREN, BP 1313 Yamoussoukro, Cote d'Ivoire
[c] Laboratoire de Biologie des Organismes et Ecosystèmes Aquatiques (BOREA), Université de Caen-Normandie, MNHN, SU, UA, CNRS, IRD, Esplanade de la paix, 14032 Caen Cedex, France
[d] IRD UMR207, UAGRM, Santa Cruz, Bolivia
[e] IRD UMR BOREA, France
[f] NORMANDIE UNIV, UNICAEN, SF ICORE, Proteogen Platform, Esplanade de la paix, 14032 Caen, France
[g] Sorbonne Université, CNRS, FR2424, ABiMS, Station Biologique, F-29680 Roscoff, France

ARTICLE INFO

Keywords:
Heterotis niloticus
Reproduction
Vitellogenin
Neurohormones
Pituitary gland
Estradiol receptors
Transcriptome
Proteome

ABSTRACT

The African bonytongue (*Heterotis niloticus*) is an excellent candidate for fish farming because it has outstanding biological characteristics and zootechnical performances. However, the absence of sexual dimorphism does not favor its reproduction in captivity or the understanding of its reproductive behavior. Moreover, no molecular data related to its reproduction is yet available. This study therefore focuses on the structural identification of the different molecular actors of vitellogenesis expressed in the pituitary gland, the liver and the ovary of *H. niloticus*. A transcriptomic approach based on *de novo* RNA sequencing of the pituitary gland, ovary and liver of females in vitellogenesis led to the creation of three transcriptomes. *In silico* analysis of these transcriptomes identified the sequences of pituitary hormones such as prolactin (PRL), luteinizing hormone (LH) and follicle-stimulating hormone (FSH) and their ovarian receptors (PRLR, FSHR, LHR). In the liver and ovary, estrogen receptors (ER) beta and gamma, liver vitellogenins (VtgB and VtgC) and their ovarian receptors (VLDLR) were identified. Finally, the partial transcript of an ovarian Vtg weakly expressed compared to hepatic Vtg was identified based on structural criteria. Moreover, a proteomic approach carried out from mucus revealed the presence of one Vtg exclusively in females in vitellogenesis. In this teleost fish that does not exhibit sexual dimorphism, mucus Vtg could be used as a sexing biomarker based on a non-invasive technique compatible with the implementation of experimental protocols *in vivo*.

1. Introduction

In teleost fish as in all vertebrates, reproductive functions, from gametogenesis to sexual behavior during pair formation, are mostly controlled by the hypothalamic-pituitary–gonadal axis (Maruska and Fernald, 2011; Takahashi et al., 2016). Under the effect of sensory stimuli such as environmental factors (Plant, 2015; Yang et al., 2017) and social factors (Maruska and Fernald, 2011), gonadotropin-releasing hormones (GnRH) produced by the brain stimulate the synthesis of neurohormones such as prolactin (PRL), luteinizing hormone (LH) and follicle stimulating hormone (FSH) by the pituitary gland (Zohar et al., 2010). PRL participates in the regulation of several physiological processes including growth, reproduction and osmoregulation (Bu et al., 2015; Freeman et al., 2000; Manzon, 2002; Whittington and Wilson, 2013). LH and FSH stimulate the production of steroids that play a role in sexual differentiation by regulating gonad development in both sexes, the reproductive cycle (gametogenesis, secondary sexual characteristics, sexual behavior) and growth (Devlin and Nagahama, 2002; Taranger et al., 2010; Zohar et al., 2010). These gonadotropins (GTHs) are used in aquaculture as inducers to improve the reproduction of certain fish species (Mehdi and Ehsan, 2013; Mylonas et al., 2010). 17β-Estradiol (E2) is synthesized by the follicular cells of oocytes under the

* Corresponding author.
E-mail address: joel.henry@unicaen.fr (J. Henry).

https://doi.org/10.1016/j.ygcen.2020.113532
Received 6 March 2020; Received in revised form 29 May 2020; Accepted 3 June 2020
Available online 12 June 2020

control of FSH (Hara et al., 2016; Norris and Lopez, 2011); it is the main sex hormone that stimulates oogenesis and vitellogenin (Vtg) synthesis by the liver in females (Hara et al., 2016; Nelson and Habibi, 2013). This hormone is perceived *via* intracellular estrogen receptors (ERs) mainly expressed in the liver (Nelson and Habibi, 2013; Unal et al., 2014).

Vitellogenin (Vtg) is a phospholipoglycoprotein precursor of vitelline proteins that accumulate in female oocytes during vitellogenesis in the majority of oviparous species such as fish teleostans, amphibians, reptiles, birds, most invertebrates, and platypus (Wallace et al., 1990; Zhang et al., 2015). Vtg is mainly expressed in the liver, but also in non-hepatic tissues (Zhong et al., 2014) where its level of expression is much lower (Wang et al., 2005). In general, teleost fish Vtg comprises a heavy chain named lipovitelline (LvH or LvI), a phosvitin domain rich in phosphorylated serines (Pv), a light chain of lipovitelline (LvL or LvII), a β'-C component, and a C-terminal coding region (CT) (Hara et al., 2016; Hiramatsu et al., 2006; Zhang et al., 2015) Vtg plays a vital role in embryo development and larval survival. Methods based on the detection of plasma Vtg in sexually mature adult individuals can be used to sex fish that do not display sexual dimorphism.

For example, in *Arapaima gigas* that belongs to the order Osteoglossiformes and does not exhibit sexual dimorphism (Chu-Koo et al., 2009; Dugué et al., 2008), the use of Vtg as a sexing marker has made it possible to breed pairs and optimize reproduction in captivity (Carreiro et al., 2011; Chu-Koo et al., 2009; Nuñez, 2008). Blood-borne Vtg is incorporated into oocytes by endocytosis mediated by Vtg receptors (VtgR) integrated in the oocyte membrane (Dominguez et al., 2014; Le Menn, 1979; Mañanós et al., 2007; Núñez-Rodríguez et al., 1996; Stifani et al., 1990). VLDLRs are recycled following endocytosis (Hara et al., 2016; Hiramatsu et al., 2006).

Heterotis niloticus, the species addressed in this study, is of economic interest because it is important for inland fisheries and aquaculture in Africa. It shows exceptional growth performances (Adite et al., 2006; Ezekiel and Abowei, 2013; Moreau, 1982; Odo et al., 2009). Several technical barriers currently make it difficult to set up a zootechnical path compatible with mass production. Biological particularities (Monentcham et al., 2009) such as the absence of sexual dimorphism (Carreiro et al., 2011; Oladosu et al., 2007) observed in this Osteoglossiforme species phylogenetically close to *A. gigas* (Betancur-R et al., 2017; Guo-Qing and Wilson, 1996; Hilton, 2001, 2003; Lavoué, 2016; Lavoué and Sullivan, 2004; Nelson, 1968, 1969) makes it difficult to breed them in captivity or understand their reproductive behavior. Information on Vtg (Chu-Koo et al., 2009; Dugué et al., 2008) is available about *A. gigas*, and the structures of the α (Faria et al., 2013) and β (Sevilhano et al., 2017) subunits of LH and FSH are known (Borella et al., 2009; Marcos and Adalberto, 2015), but no molecular data is yet available on vitellogenesis and its regulation in *H. niloticus*.

The aim of this study was therefore to characterize Vtg(s) and their ovarian receptor(s), pituitary neurohormones and their receptors, as well as E2 receptors in *H. niloticus* using a combination of several "omic" approaches. A sexing method based on the presence of vitellogenin was further tested from the mucus and blood of mature individuals using a proteomic approach.

2. Materials and methods

2.1. Sampling and tissue samples

The selected tissues were the pituitary gland, the ovary, the liver, the blood and the mucus of *H. niloticus* from a fish farm located in central Côte d'Ivoire (5°08,713′W; 6°50,929′N) nearby the political capital Yamoussoukro. The species was introduced for the first time in this country in 1957, from Cameroon (Lazard, 1990). Tissue samples were taken from fish anesthetized in water containing 96° ethanol (1/10 dilution) (Bhanu and Philip, 2011). The samples were taken from male and female individuals (whose sex was only known after dissection) during the vitellogenesis period (May and June).

For RNA sequencing of the liver, ovary and pituitary gland, pooled samples from three females were stored in RNAlater stabilization solution (Sigma) at 4 °C. For proteomic analyses of blood and mucus, the extraction medium was composed of 200 mM PBS, 0.1 mM EDTA, 1 mM DTT and 400 mM PMSF, and the samples were stored at −80 °C until extraction.

2.2. Extraction of total RNAs

The total RNAs of each tissue pool were extracted separately. The organs were ground in liquid nitrogen, and 50 to 100 mg of tissue powder were mixed with 1 ml of TRizol (Ambion, Life Technologies, Carlsbad, Calif., USA). Then extraction was performed according to the manufacturer's recommendations. Total RNA was quantified using a NanoDrop Spectrophotometer ND-1000 (NanoDrop Technologies, Inc.) and its integrity was assessed on a 2100 Bioanalyzer (Agilent Technologies). Libraries were generated from 250 ng of total RNA as following: mRNA enrichment was performed using the NEBNext Poly (A) Magnetic Isolation Module (New England BioLabs). cDNA synthesis was achieved with the NEBNext RNA First Strand Synthesis and NEBNext Ultra Directional RNA Second Strand Synthesis Modules (New England BioLabs). The remaining steps of library preparation were done using and the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England BioLabs). Adapters and PCR primers were purchased from New England BioLabs. Libraries were quantified using the Quant-iT™ PicoGreen® dsDNA Assay Kit (Life Technologies) and the Kapa Illumina GA with Revised Primers-SYBR Fast Universal kit (Kapa Biosystems). Average size fragment was determined using a LabChip GX (PerkinElmer) instrument.

2.3. RNA sequencing and in silico approach

2.3.1. Illumina sequencing, assembly and annotation

The RNA samples were sequenced by the Génome Québec platform (http://www.genomequebec.com/ressources-et-platformes-technologiques.html) on an Illumina HiSeq4000 sequencer in 2 × 100 base pairs paired mode.

The raw readings were cleaned and filtered, and the traces of adapters were removed using Trimmomatic v.0.33 (http://www.usadellab.org/cms/index.php?page = trimmomatic). The readings were filtered using a quality threshold of 30 and a minimum size of 50 base pairs. SortMeRNA v2.1 was used to filter ribosomal RNAs from reads (Kopylova et al., 2012). Finally, the cleanup process was verified using fastQC v.0.11.5 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

The cleaned readings from the different libraries were subsequently assembled together using the *de novo* transcriptome assembler Trinity v.2.5.1 (Grabherr et al., 2011). Relative abundances were estimated using kallisto v.0.43.1 (Bray et al., 2016). To obtain the expression values (FPKM: fragments per kilobase of exon per million fragments mapped) and identify low coverage contigs (FPKM < 1) and rare isoforms (< 1%) so as to exclude them from the analysis, the two software programs were launched *via* the Trinity package scripts. Peptide prediction was performed using Transdecoder v.3.0.0 (Haas et al., 2013). A similarity search (blastx of assembled transcripts and blastp of predicted peptides) was carried out against the uniprot-swissprot database (version 09-2013). Signal peptides were predicted using signalP v4.0 (Petersen et al., 2011). Transmembrane domains were detected using TMHMM v2.0c (Krogh et al., 2001). Finally, the functional annotation of the transcriptome was carried out using the Trinotate v.3.0.1 pipeline (http://trinotate.github.io) described by Bryant et al. (2017).

2.3.2. In silico analysis of the transcripts

In silico analyses of the *H. niloticus* liver, ovary and pituitary

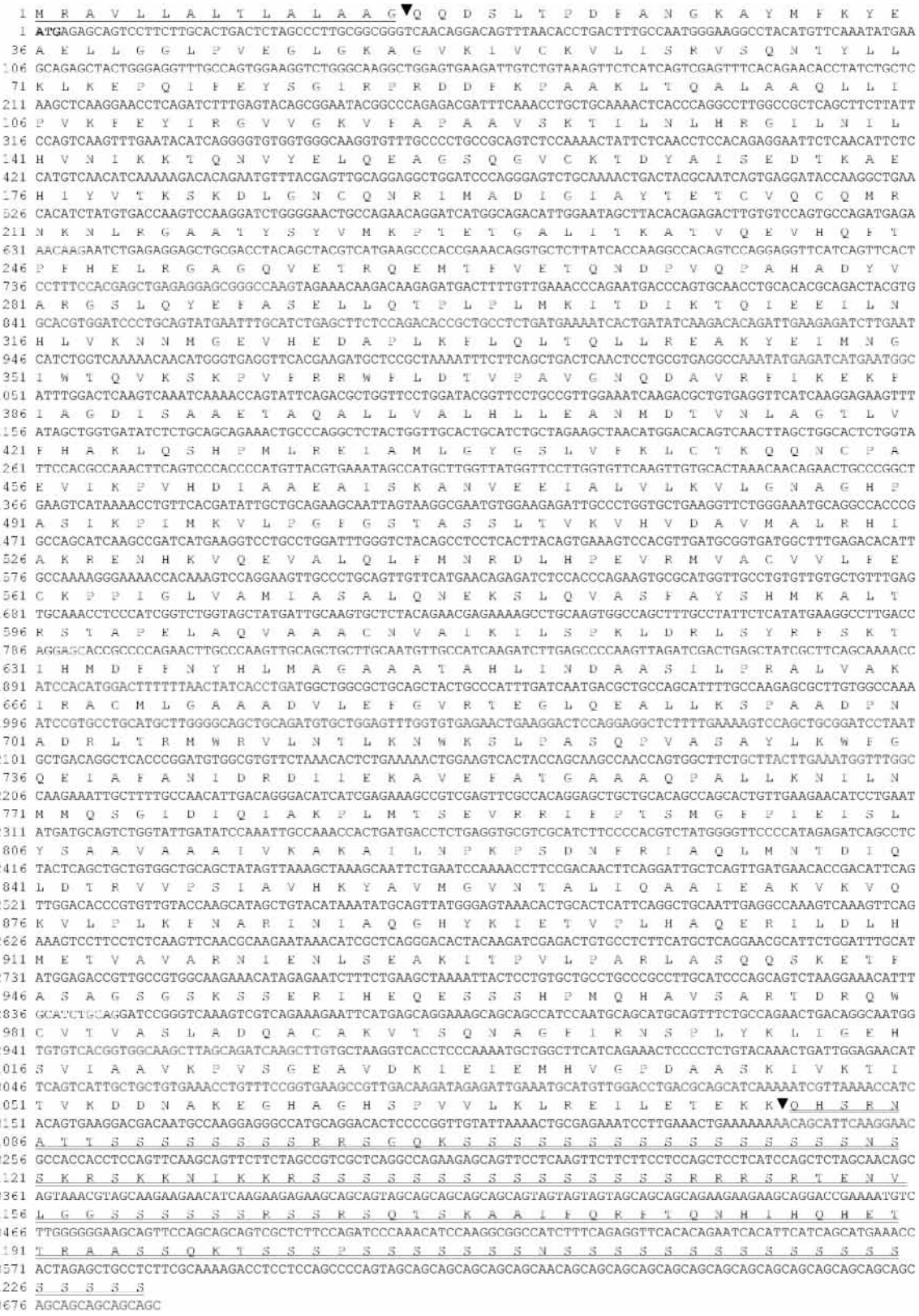

Fig. 1a. Nucleotide sequence and deduced protein sequence of the N-terminal end of Vtg1 identified in *Heterotis niloticus* liver (N-ter1 HnVtg1). The start codon is in bold type. The signal peptide is underlined, lipovitellin I (LvI) is delineated by two black triangles, and partial phosvitin (Pv) is double-underlined.

transcripts resulting from the assembly were mainly performed with Peptraq software developed internally (Zatylny-Gaudin et al., 2016). Peptraq is a software program dedicated to the analysis of sequence files in fasta or txt formats. It translates and filters transcripts and protein precursors on the basis of annotation using keywords or structural criteria related to the presence of particular subsequences. Peptraq can also assign the signal peptide, cleave it from preproprotein to proprotein, carry out the cleavages of the convertases and identify the

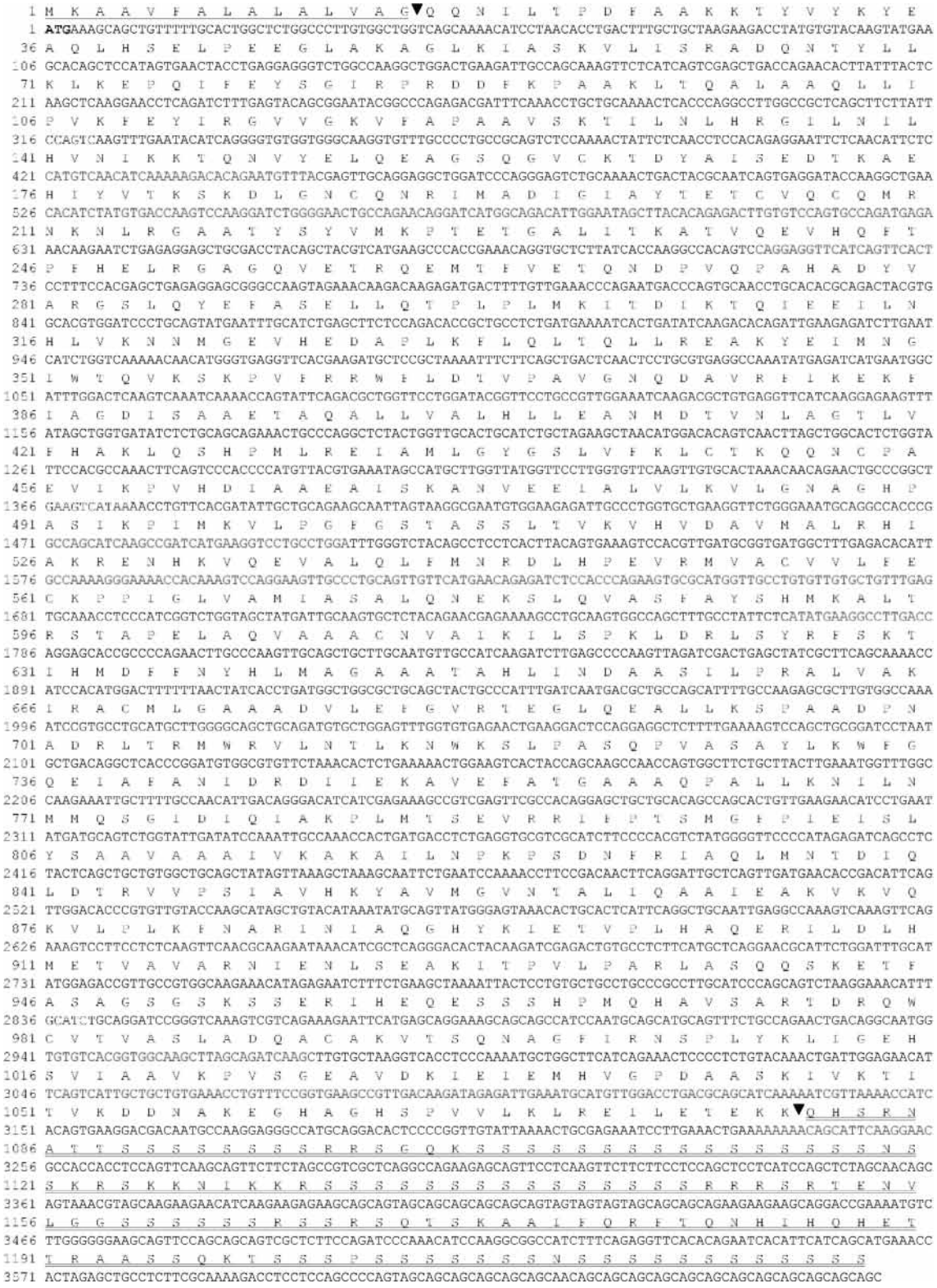

Fig. 1b. Nucleotide sequence and deduced protein sequence of the N-terminal end of Vtg2 identified in *Heterotis niloticus* liver (N-ter2 HnVtg2). The start codon is in bold type. The signal peptide is underlined, lipovitellin I (LvI) is delineated by two black triangles, and partial phosvitin (Pv) is double-underlined.

characteristic sequence repeats of the neuropeptides.

The transmembrane domains were predicted by TMHMM v2.0 (www.cbs.dtu.dk/services/TMHMM). Secondary structures were obtained by alignment with proteins of species whose domains were known or by searching for conserved domains on the NCBI site (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb. IRM). Tertiary structure was modeled under I-TASSER (https://zhanglab.ccmb.med.umich.edu/I- TASSER/). The N-glycosylation sites were identified under NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/). The nucleic and protein sequences were aligned with Show translation (https://www.bioinformatics.org/sms/show_trans.html). Protein sequence alignments of different species were performed under CLC Sequence

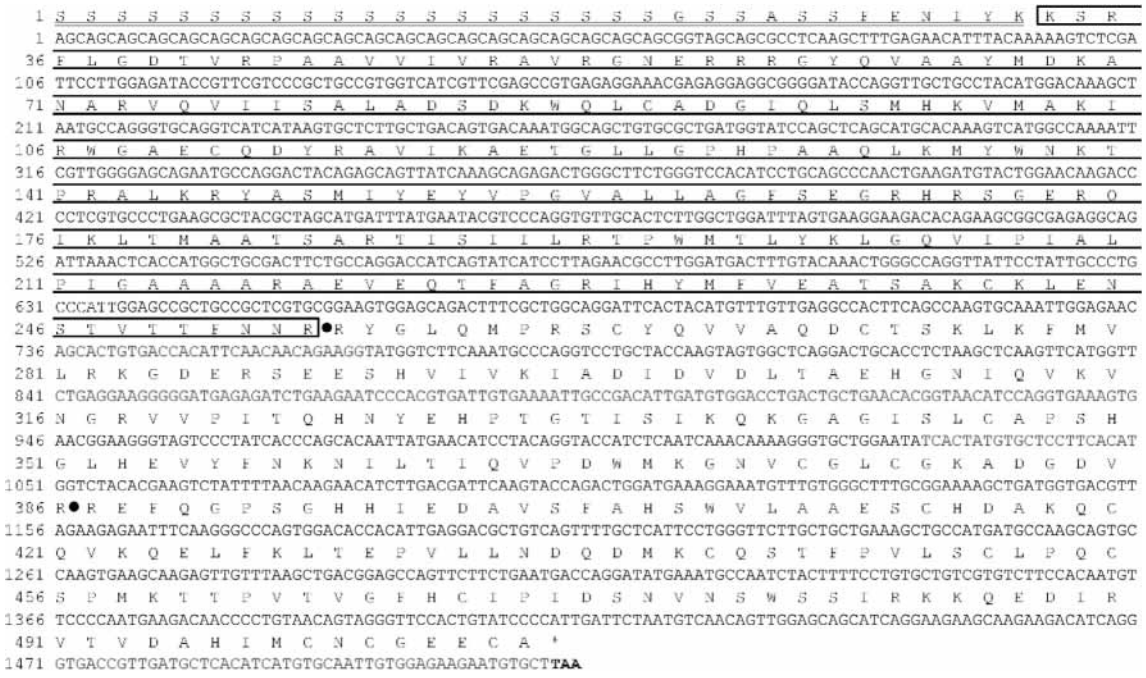

Fig. 1c. Nucleotide sequence and deduced protein sequence of the C-terminal end (C-ter) of Vtg1 and Vtg2 identified in *Heterotis niloticus* liver. The stop codon is in bold type. The phosvitin (Pv) is double-underlined, lipovitellin II (LvII) is in boxes, the β'-C domain is delineated by two black circles, and followed by the C-terminal (CT) coding region.

Viewer 7.6.1 (http://www.clcbio.com). The amino acid identities across protein sequences were determined by Protein Blast (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

2.4. Proteomic approach focused on the search for Vtg

2.4.1. Sample preparation for MS analysis

Proteomic analysis samples were first thawed in ice water and centrifuged for several minutes. Proteins were precipitated at a rate of 20 µl *per* 120 µl of refrigerated acetone. The samples were stored at −20 °C overnight, and then centrifuged at 20,000 g for 20 min. The protein pellet was resuspended in ammonium bicarbonate buffer (50 mM, pH 7). Proteins were digested overnight at 37 °C using 0.25 µg of porcine trypsin (Promega, Madison, USA). The digestates were desalted and concentrated on an Omix ZipTip µC18 (Agilent) before analysis.

2.4.2. NanoLC-MS/MS

The chromatography step was performed on an ultra-high-pressure nano-chromatography system (NanoElute, Bruker Daltonics). The peptides were concentrated on a C18 pepmap 100 (5 mm × 300 µm) precolumn (Thermo Scientific) and separated at 50 °C on a Reprosil reverse phase column (25 cm × 75 µm, 1.6 µm, C18) (Ionopticks). The mobile phases consisted of 0.1% formic acid, 99.9% water (v/v) (A) and 0.1% formic acid in 99.9% ACN (v/v) (B). The flow rate was set at 400 nl/min, and the gradient profile was as follows: from 2 to 15% B in 60 min, followed by a 25% increase in B over 30 min and then a 37% increase in 10 min, a 95% B wash step and a 2% B rebalance.

MS analyses were performed on a TIMS-TOF mass spectrometer (Bruker Daltonics) with a nano-electrospray ion source (CaptiveSpray, Bruker Daltonics). The system was calibrated weekly, and the accuracy of the mass measurements was greater than 1 ppm. A capillary voltage of 1,400 V was used for ionization. The MS spectra were positive in the 100 to 1,700 *m/z* mass range. The mass spectrometer was used in PASEF (parallel accumulation-serial fragmentation) mode (Meier et al., 2015), excluding monocharged peptides. Ten PASEF MS/MS scans were performed in 1.25 s from the 2–5 load range.

2.4.3. Peptide sequencing and vitellogenin detection

The fragmentation spectrum was used to determine peptide sequences. The database was searched using Mascot 2.6.1 program (Matrix Science). Two local *H. niloticus* databases were used: an ovarian protein database (437,476 entries) and a liver protein base (244,234 entries), both constructed from the corresponding transcriptomes using Peptraq software.

The allowed variable modifications were as follows: C-carbamidomethyl, K-acetylation, oxidation and dioxidation of methionine. "Trypsin" was selected with a tolerant mode including two missing cleavage sites (Šlechtová et al., 2015). Mass accuracy was set at 20 ppm and 0.05 Da for the MS and MS/MS modes, respectively. The Mascot data was then transferred to Proline validation software (http://www.profiproteomics.fr/proline/) for data filtering with a significance level of less than 0.05 and to eliminate redundant proteins.

3. Results

3.1. Rnaseq and in silico approach

A total of 1,003,989,936 raw reads paired with read lengths of 100 base pairs were generated. After cleaning the poor quality adapters and sequences, 667,119,586 high quality paired readings were used to generate a first overall assembly of 257,484 transcripts (corresponding to 156,636 Trinity "genes"). The length of the transcripts varies between 201 and 24,080 base pairs, an average length of 1,008.5 base pairs and a median length of 418 base pairs. 90% of the cleaned reads were successfully pseudo-aligned with the overall complete transcriptome, indicating strong support for the transcriptome assembled by the reads. Weakly expressed transcripts (FPKM < 1) and rare isoforms (< 1%) were excluded from the initial assembly, which resulted in a filtered assembly of 62,852 transcripts (corresponding to 40,302 "genes" of Trinity), with lengths between 201 and 18,966 base pairs, an average length of 1,382.9 base pairs and a median length of 807 base pairs.

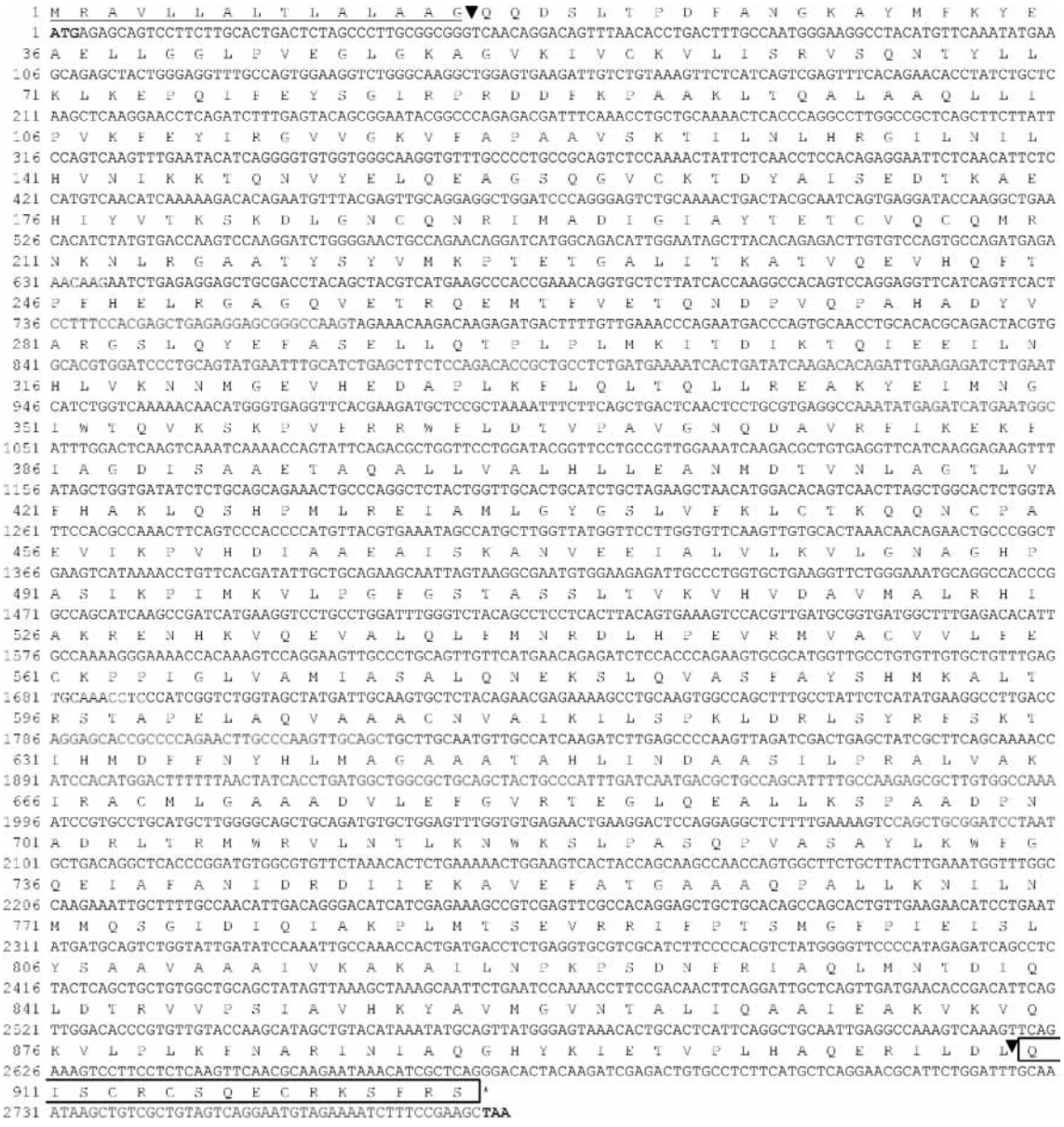

Fig. 1d. Nucleotide sequence and deduced protein sequence of the Vtg3 identified in *Heterotis niloticus* liver. The start codon and the stop codon are in bold type. The signal peptide is underlined, lipovitellin I (LvI) is delineated by two black triangles, and LvII is in boxes.

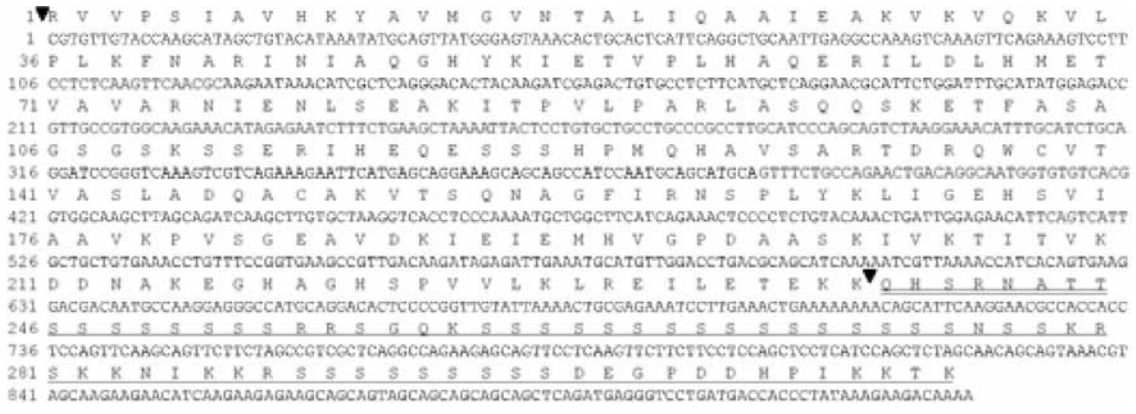

Fig. 1e. Nucleotide sequence and deduced protein sequence of the partial ovarian vitellogenin identified in *Heterotis niloticus* (HnVtg4). The lipovitellin I (LvI) is delineated by two black triangles, and phosvitin (Pv) is double-underlined.

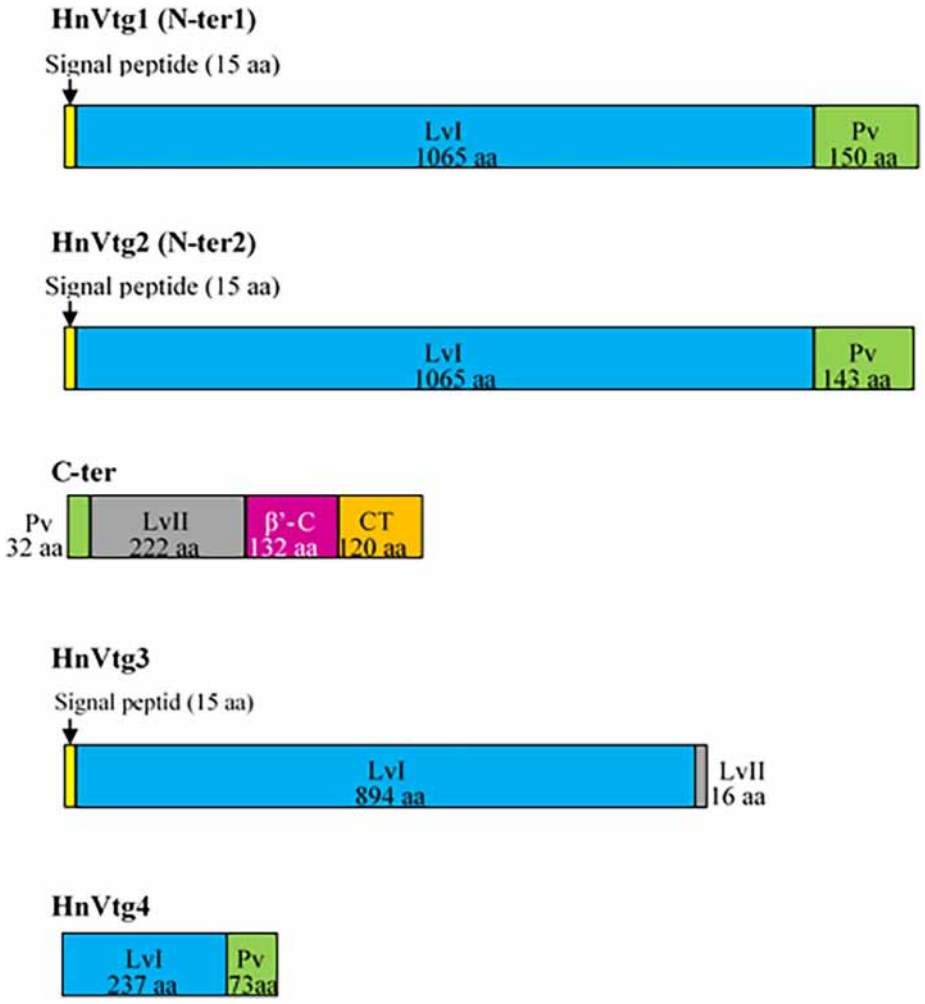

Fig. 2. Linear constructions of the different *Heterotis niloticus* Vtg domains. The numbers of amino acids (aa) are indicated.

Fig. 3a. Alignments of protein sequences deduced from the N-terminal end (N-ter1 and N-ter2) of the hepatic Vtg of *Heterotis niloticus* (HnVtg1 and HnVtg2) with little conserved domains in gray.

3.1.1. Hepatic and ovarian vitellogenin (Vtg)

Three hepatic Vtgs and one ovarian Vtg have been identified in *H. niloticus*. Two of the Vtgs expressed in the liver (HnVtg1 and HnVtg2) are represented by two transcripts for the N-terminal end of 3,690 base pairs (N-ter1, Fig. 1a) and 3,669 base pairs (N-ter2, Fig. 1b) and by a transcript in the C-terminal position of 1,518 base pairs (C-ter, Fig. 1c). The deduced protein sequences show that N-ter1 (1,230 amino acids) and N-ter2 (1,223 amino acids) respectively have 19 and 12 serines at

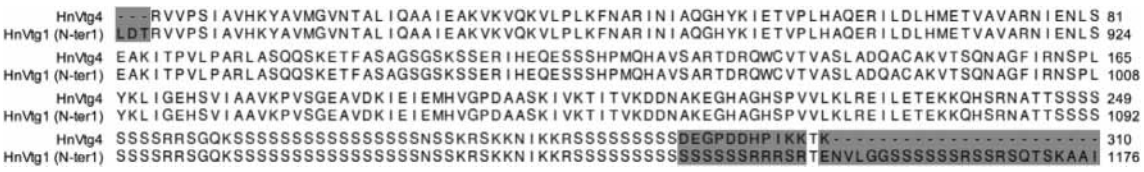

Fig. 3b. Alignments of protein sequences deduced from the N-terminal end (N-ter1) of hepatic Vtg (HnVtg1) and partial ovarian Vtg (HnVtg4) of *Heterotis niloticus* with little conserved domains in gray.

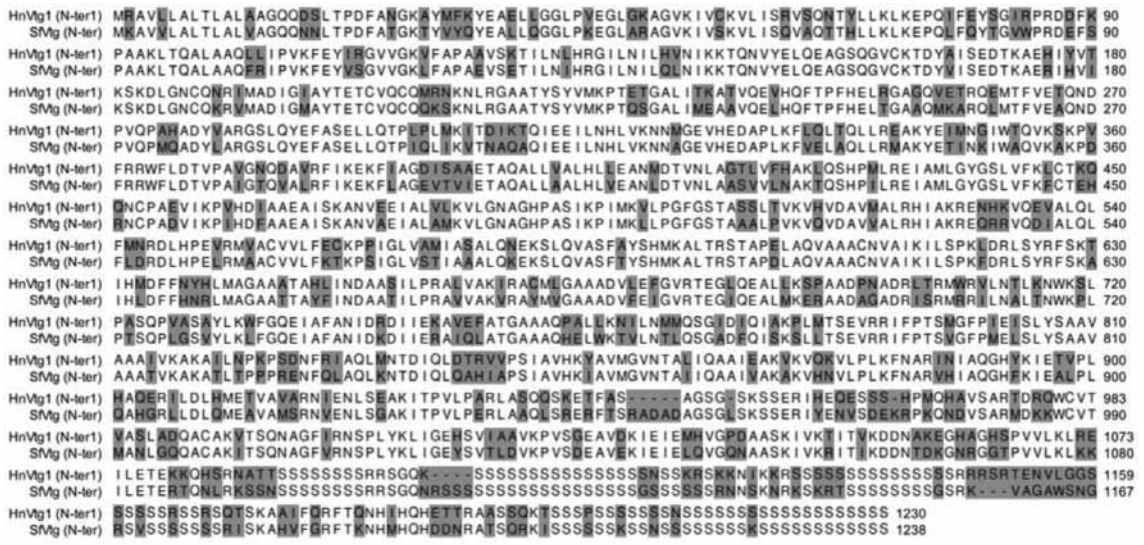

Fig. 3c. Alignment of the protein sequences deduced from the N-terminal end (N-ter1) of the Vtg1 of *Heterotis niloticus* (HnVtg1) and the N-terminal end -N-ter) of the Vtg of *S. formosus* (SfVtg) with little conserved domains in gray.

Fig. 3d. Alignment of the protein sequences deduced from the C-terminal end (C-ter) of the Vtg1 of *Heterotis niloticus* (HnVtg1) and of the Vtg of *S. formosus* (SfVtg) with little conserved domains in.

Fig. 3e. Highlighting of orphan serine located in a box between N-ter (in bold) and C-ter of *S. formosus* Vtg. The two domains are defined following alignment with N-ter and C-ter of HnVtg1.

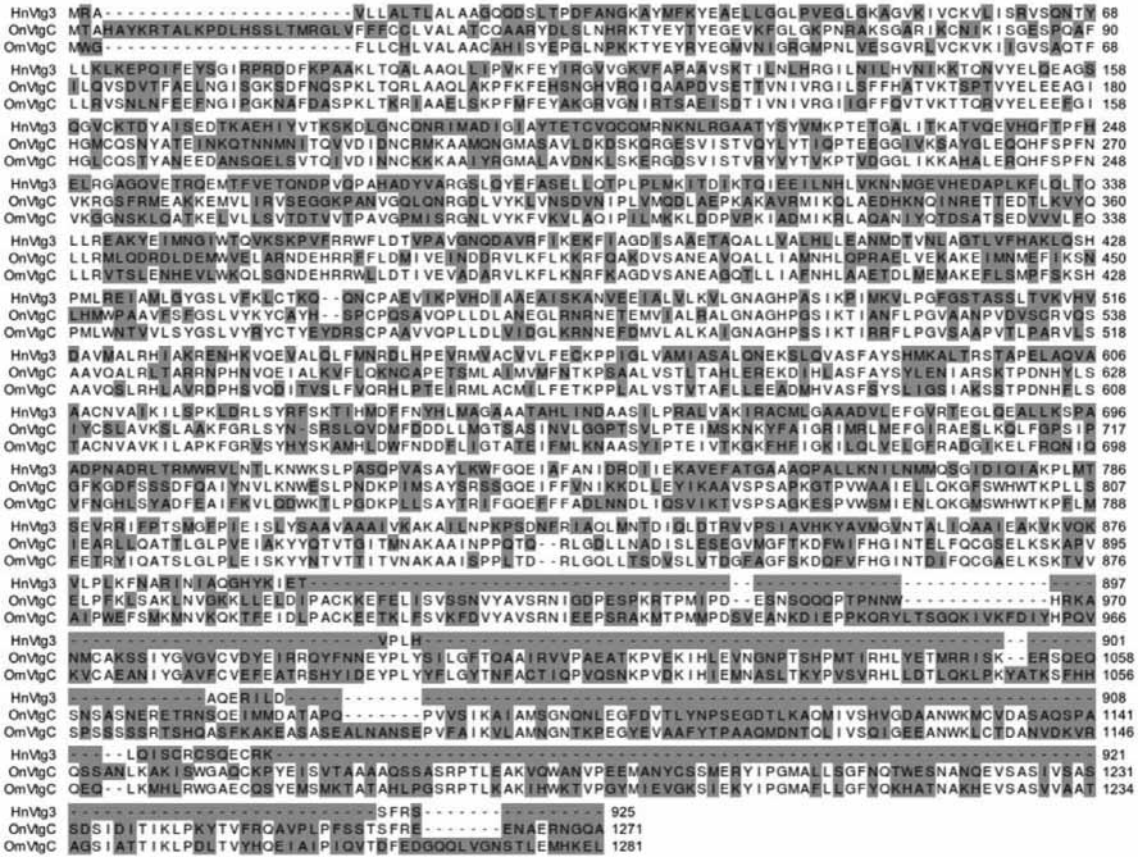

Fig. 3f. Alignment between the VtgC of *Heterotis niloticus* (HnVtg3) and the VtgC of *O. mykiss* (OmVtgC) and *O. niloticus* (OnVtgC) with little conserved domains in gray.

the end of the sequences all represented by the same AGC codon and corresponding to the N-terminal end of the phosvitin domain (Figs. 1a and 1b). The C-ter transcript of 506 amino acids presents at the start of the sequence 20 serines also represented by the AGC codon (Fig. 1c). The assembly of transcripts at the phosvitin domain is therefore impracticable by Trinity since the codons corresponding to the polyserine site are identical.

One of the hepatic Vtgs (HnVtg3) has a complete 2,775 base pairs sequence encoding a 925 amino acids protein (Fig. 1d). As for ovary Vtg (HnVtg4), it came from a partial 930 base pairs transcript encoding 310 amino acids (Fig. 1e). The secondary structure of *H. niloticus* Vtg showed that hepatic Vtgs had a signal peptide of fifteen amino acids followed by only lipovitellin domains (LvI and LvII) in HnVtg3 and by domains LvI, Pv, LvII, β'-C and CT in HnVtg1 and HnVtg2 (Fig. 2). The incomplete ovary Vtg included the LvI and Pv domains.

Alignments showed that the divergence between N-ter1 and N-ter2 was at the beginning of the sequence, with a more complete Pv domain in HnVtg1 (Fig. 3a). The Met^{1} to Lys^{909} domains of HnVtg1 and HnVtg3 (corresponding to the signal peptide and LvI of HnVtg3) were identical. The Arg^{1} domain at Ser^{297} of the incomplete ovarian Vtg (HnVtg4) was identical to the Arg^{844} domain at Ser^{1140} of N-ter1 and N-ter2 of hepatic Vtg; divergence occurred at the end of the sequence in the Pv domain (Fig. 3b).

We successively aligned *Scleropages formosus* Vtg (1,745 amino acids, AWG47880) with *H. niloticus* N-ter1 (Fig. 3c) and C-ter (Fig. 3d), and found only one orphan serine (Fig. 3e). The homology between *H. niloticus* VtgC (HnVtg3) and VtgC from other teleost fish species was analyzed; it showed 34.08% amino acid identity with *O. mykiss* (BBA57869) and 31.41% with *O. niloticus* (XP_005459969) (Fig. 3f).

3.1.2. Vitellogenin receptor

Teleost vitellogenin receptors (VtgR) are classified as very low-density lipoprotein receptors (VLDLR) belonging to the low-density lipoprotein superfamily (LDLR) (Li et al., 2003; Mizuta et al., 2013; Pousis et al., 2012).

Two transcripts of 2,583 and 2,523 base pairs encoding two VLDLRs of 861 and 841 amino acids were identified in the ovary (Fig. 4a) and the pituitary gland, respectively. HnVLDLR comprised a signal peptide (Met^{1}-Arg^{26}) followed by seven and eight domains for the pituitary and ovarian forms, respectively (Fig. 4b). HnVLDLR expressed in the ovary had an O-glycosylated domain located between Ala^{748} and Ser^{767}; this domain was not found in the pituitary form. Apart from the O-glycosyl domain, the other domains of the two forms of HnVLDLR had identical amino acid sequences.

The ligand-binding domain (LBD) was composed of eight ligand-binding repeats (LBRs) of 39 to 44 amino acids each, among which six cysteines and the SDE (LBR1-2, 4–7) and EDE (LBR3) motifs. The precursor domains of the epidermal growth factor (EGF) was found in three EGF1 (Cys^{355}-Cys^{389}), EGF2 (Cys^{395}-Cys^{429}) and EGF3 (Cys^{702}-Cys^{745}) repeats; each of them comprised six cysteines.

Between EGF2 and EGF3 stood a domain comprising five 4-aa sequences of the YWS(V)D and FWA(T)D types. A single 23-aa transmembrane domain (Ala^{786} to Trp^{808} for the ovarian form) was present, and a cytoplasmic domain containing the FDNPVY sequence was found in C-terminal position.

The analysis of amino acid sequence homology showed that apart from the O-glycosylated domain that was absent from the VLDLRs of certain species, the particular sequences (SDE, EDE, YWS(V)D, FWA(T)D and FDNPVY) were highly conserved. Moreover, HnVLDLR shared a strong identity with VtgR or VLDLR of *S. formosus* (XP_018607358) (94%), *O. bicirrhosum* (AXN72824) (92%), *O. mykiss* (CAD10640)

```
   1 M  V  T  S  I  L  G  F  L  I  L  P  I  C  L  Q  Q  C  A  Y  V  R  G  S  Q  A  E  C  D  P  S  Q  F  Q  C
   1 ATGGTCACGTCTATTCTGGGCTTTCTTATCCTGCCAATATGTCTTCAGCAGTGTGCCTATGTGCGGGGATCCCAGGCGGAATGTGATCCCTCTCAGTTCCAGTGT
  36 G  N  G  R  C  I  P  S  V  W  Q  C  D  G  D  E  D  C  S  D  G  S  D  E  N  T  C  V  R  K  T  C  A  E  V
 106 GGAAATGGCCGCTGTATTCCCTCAGTGTGGCAGTGTGACGGTGATGAAGACTGTTCTGATGGCAGTGATGAAAACACCTGTGTGAGGAAGACCTGCGCAGAAGTG
  71 D  F  V  C  R  N  G  Q  C  V  P  K  R  W  H  C  D  G  E  P  D  C  E  D  G  S  D  E  S  S  Q  V  C  H  M
 211 GACTTTGTCTGCCGCAATGGCCAGTGTGTCCCCAAGAGGTGGCACTGTGATGGCGAGCCCGATTGTGAAGACGGGGTCCGACGAAAGCTCCCAAGTCTGCCACATG
 106 R  T  C  R  M  N  E  F  S  C  G  A  G  T  T  Q  C  I  P  V  F  W  K  C  D  G  E  K  D  C  D  N  G  E  D
 316 CGGACCTGTCGTATGAACGAGTTCAGCTGTGGGGCAGGTACAACCCAGTGCATCCCAGTTTTCTGGAAATGCGATGGAGAGAAGGACTGCGACAACGGGGAGGAT
 141 E  V  N  C  G  N  I  T  C  A  P  L  E  F  T  C  A  S  G  R  C  V  S  Q  N  F  V  C  N  G  E  D  D  C  G
 421 GAGGTTAACTGTGGCAACATCACCTGCCGCCCCCTTGGAGTTCACCTGTGCAAGCGGACGCTGTGTGTCTCAGAATTTGTGTGCAATGGCGAGGACGACTGCGGT
 176 D  G  S  D  E  L  D  C  E  P  S  S  C  G  P  S  Q  F  Q  C  G  N  A  T  C  I  P  D  S  W  V  C  D  D  D
 526 GATGGCAGCGACGAGCTAGATTGTGAGCCCTCATCCTGCGGACCCAGCCAGTTCCAGTGTGGCAATGCCACCTGTATCCCTGACAGCTGGGTATGTGATGATGAT
 211 V  D  C  Q  D  Q  S  D  E  S  P  Q  R  C  G  R  H  P  T  P  P  A  K  C  P  S  S  E  M  R  C  G  S  G  E
 631 GTGGACTGCCAGGACCAGTCAGATGAGTCACCACAGCGCTGTGGCCGCCACCCTACACCTCCCGCCAAGTGCCCTTCCAGCGAGATGCGCTGTGGCTCGGGCGAG
 246 C  I  H  R  K  W  R  C  D  G  D  A  D  C  K  D  G  S  D  E  D  N  C  P  V  R  T  C  R  P  D  Q  F  K  C
 736 TGCATCCACCGCAAGTGGCGCTGCGATGGCGATGCGGACTGTAAGGATGGCAGTGACGAGGACAACTGCCCTGTCCGAACCTGCAGACCAGACCAGTTCAAGTGT
 281 N  D  G  S  C  I  H  G  S  R  Q  C  N  G  M  R  D  C  S  D  G  S  D  E  L  N  C  K  N  V  S  Q  C  S  G
 841 AATGATGGTAGCTGCATCCATGGCAGCCGGCAGTGCAATGGGATGCGTGACTGCTCAGATGGCAGTGATGACTCAACTGCAAAAATGTGTCCCAATGCAGTGGC
 316 P  D  K  F  K  C  R  S  G  E  C  I  E  M  S  K  V  C  N  K  V  R  D  C  S  D  W  S  D  E  P  L  K  E  C
 946 CCAGATAAATTCAAGTGCCGTAGCGGCGAGTGCATCGAGATGAGCAAAGTGTGCAACAAGGTCCGGGACTGCTCAGACTGGAGCGATGAGCCCCTCAAAGAGTGC
 351 N  L  N  E  C  L  Q  N  N  G  G  C  S  H  I  C  R  D  N  I  I  G  Y  E  C  D  C  T  P  G  L  Q  L  I  D
1051 AATCTGAATGAATGCCTTCAGAACAATGGTGGCTGCTCCCACATCTGTCGGGACAACATCATTGGTTACGAATGCGACTGCACTCCAGGCCTCCAGCTCATTGAC
 386 R  K  T  C  G  D  I  N  E  C  L  N  P  G  I  C  S  Q  I  C  I  N  L  K  G  G  Y  K  C  E  C  H  N  G  Y
1156 CGCAAGACCTGTGGTGATATTAATGAGTGCCTAAACCCTGGGATCTGTAGCCAGATCTGCATCAACTTGAAGGGGGGGTACAAGTGTGAGTGTCACAATGGGTAC
 421 Q  M  D  P  T  T  G  V  C  K  A  V  G  K  E  P  C  L  I  F  T  N  R  R  D  I  R  K  L  T  L  E  R  R  E
1261 CAGATGGATCCTACCACAGGCGTGTGCAAAGCAGTGGGCAAGGAGCCATGCCTGATCTTCACCAACCGGCGTGACATCCGCAAGCTGACCCTGGAGCGGCGGGAG
 456 Y  T  Q  I  V  E  Q  L  R  N  T  V  A  L  D  A  D  F  T  Q  Q  M  I  F  W  A  D  L  G  Q  K  A  I  F  S
1366 TACACCCAGATTGTGGAGCAGTTCGCCAACACTGTGGCCCTAGATGCAGACTTCACCCAACAGATGATCTTCTGGGCTGATCTAGGTCAAAAGGCAATCTTCAGC
 491 M  S  L  D  K  R  D  E  G  S  I  T  K  V  I  E  S  V  Q  T  P  V  G  I  A  V  D  W  I  Y  K  N  I  Y  W
1471 ATGTCACTTGACAAGCGGGACGAAGGCAGTATCACAAAAGTGATTGAGAGTGTGCAGACTCCTGTCGGGATTGCAGTAGACTGGATTTACAAGAATATATACTGG
 526 S  D  L  G  T  K  T  I  S  V  A  N  F  N  G  T  K  T  K  I  L  F  S  S  G  L  K  E  P  A  S  I  A  V  D
1576 TCAGATTTGGGTACCAAGACGATATCTGTAGCAAACTTCAATGGCACGAAGACCAAAATTTTGTTCAGCAGTGGGCTAAAGGAACCGGCCTCCATTGCTGTGGAT
 561 P  L  S  G  F  L  Y  W  S  D  W  G  E  P  A  R  I  E  K  S  G  M  N  G  V  D  R  Q  V  L  V  E  T  D  I
1681 CCTCTCTCTGGGTTTCTGTATTGGTCAGACTGGGGTGAGCCAGCCAGGATTGAGAAGTCTGGCATGAATGGTGTAGATCGGCAGGTTCTGGTTGAGACGGACATT
 596 Q  W  P  N  G  I  T  L  D  L  I  K  S  R  L  Y  W  V  D  S  K  L  H  M  L  C  S  V  D  L  N  G  D  N  R
1786 CAGTGGCCTAATGGGATCACTCTAGACTTGATTAAAAGCAGGCTCTACTGGGTGGACTCCAAACTACACATGCTCTGCAGCGTGGATCTGAACGGGGATAACCGG
 631 R  K  V  L  Q  S  P  D  Y  L  A  H  P  F  A  L  T  V  F  E  D  R  V  F  W  T  D  G  E  N  E  A  I  Y  G
1891 CGGAAAGTTCTCCAGTCTCCTGACTATCTCGCGCATCCTTTTGCTCTCACGGTTTTTGAGGATCGAGTGTTCTGGACAGATGGGGAAAACGAGGCCATCTATGGT
 666 A  N  K  F  T  G  S  D  V  I  M  L  A  S  N  L  N  E  P  Q  D  I  I  V  Y  H  E  L  I  Q  L  S  G  T  N
1996 GCCAACAAATTCACTGGGTCAGATGTGATTATGCTTGCCAGCAACCTGAATGAGCCACAGGACATCATTGTTTATCATGAGCTCATCCAGCTATCCGGAACCAAT
 701 W  C  T  D  K  G  D  N  G  G  C  A  F  L  C  L  P  A  P  Q  I  N  K  H  S  P  K  Y  T  C  I  C  S  Q  G
2101 TGGTGCACTGACAAAGGGGACAATGGAGGCTGTGCCTTCCTGTGCTTGCCAGCACCACAGATCAACAAGCACTCTCCAAAGTACACCTGCATTTGCTCACAGGGG
 736 M  E  L  A  A  D  G  L  N  C  R  P  A  T  I  S  A  H  P  G  D  G  K  T  L  I  H  P  T  H  A  S  E  G  N
2206 ATGGAGCTTGCTGCCGATGGCTTGAACTGCAGACCAGCCACCATTTCTGCACACCCAGGAGATGGAAAGACACTAATACACCCCACTCATGCTTCAGAGGGGAAT
 771 I  S  T  S  I  H  E  V  N  S  S  A  K  G  S  A  A  A  W  A  I  L  P  V  L  L  L  A  M  A  V  A  G  G  Y
2311 ATTAGTACCTCCATCCATGAGGTGAACTCCTCAGCCAAAGGGTCAGCTGCTGCATGGGCCATCCTCCCTGTCTTGCTGCTGGCCATGGCTGTGGCAGGAGGCTAT
 806 L  M  W  R  N  W  Q  L  K  N  K  K  S  M  N  P  D  N  P  V  Y  L  K  T  T  E  E  D  L  N  I  D  I  S  R
2416 CTAATGTGGCGCAACTGGCAGCTGAAGAACAAAAGAGCATGAACTTTGACAATCCTGTGTATCTGAAGACCACCGAGGAGGACCTCAACATCGACATCAGCCGG
 841 H  G  A  S  V  G  H  T  Y  P  A  I  S  V  V  N  T  E  D  D  L  *
2521 CACGGTGCATCTGTGGGCCACACGTACCCTGCTATATCAGTTGTGAATACAGAAGATGACTTGTAG
```

Fig. 4a. Nucleotide sequence and deduced protein sequence of the *Heterotis niloticus* vitellogenin receptor or very-low-density lipoprotein receptor (HnVLDLR). The start codon and the stop codon are in bold type, the signal peptide is underlined, the eight repeats that bind to the ligand (LBR) are delineated by two consecutive black triangles, the three EGF1, EGF2 and EGF3 domains are located in the boxes, the highly conserved SDE, EDE, YWS(V)D, FWA(T)D and FDNPVY sequences are underlined, the O-glycosylation domain is double-underlined, and the transmembrane domain is delineated by two black circles and followed by the cytoplasmic domain in C-terminal position.

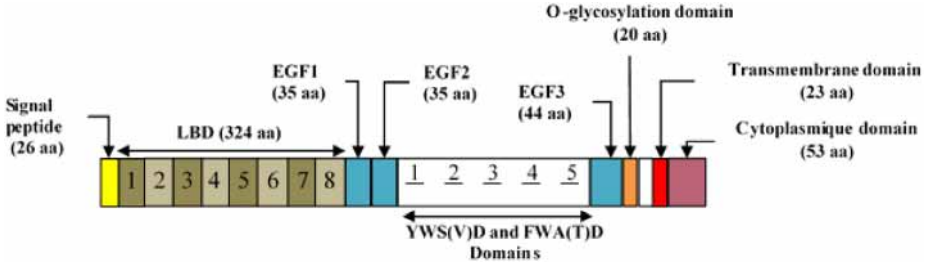

Fig. 4b. Linear constructions of the different *Heterotis niloticus* Vtg receptor domains. The numbers of amino acids (aa) are indicated.

(88%), *C. carpio* (XP_018935318) (87%) and *O. aureus* (AAO27569) (85%) (Fig. 4c).

3.1.3. 17β-estradiol receptors (ERs)

The nomenclature chosen for the 17β-estradiol receptors included the alpha subtype (ERα), the beta or beta2 subtype (ERβ) and the gamma or beta subtype 1 (ERγ).

Using *in silico* analysis, we identified the transcripts of four ERβ and one ERγ: HnERβ_X1 (1,674 base pairs) in the liver and the ovary (Fig. 5a), HnERβ_X2 (1,653 base pairs) and HnERβ_X3 (1,281 base pairs) only in the liver (Fig. 5b), and HnERβ_X4 (1,524 base pairs) and HnERγ (1,443 base pairs) only in the ovary (Fig. 5c). No ERα was identified. Three receptors linked to the estrogen receptors (ERRs) were also identified: HnERRα (1,296 base pairs) in the liver, and HnERRγ_X1

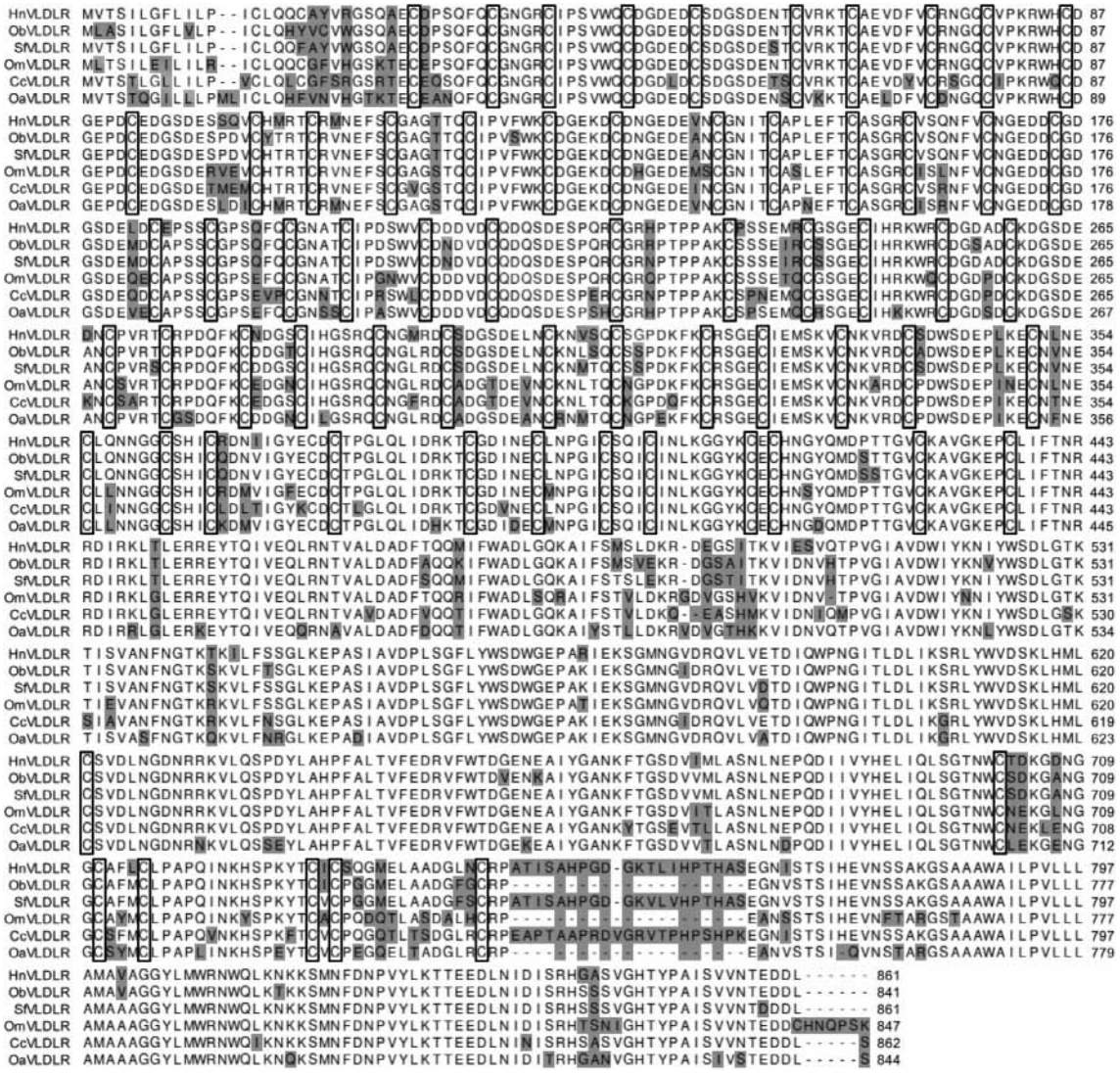

Fig. 4c. Alignment of deduced protein sequences of *Heterotis niloticus* vitellogenin receptor (HnVLDLR) with *Scleropages formosus* VLDLR (SfVLDLR), *Osteoglossum bicirrhosum* VLDLR (ObVLDLR), *Oncorhynchus mykiss* VLDLR (OmVLDLR), *Oreochromis aureus* VLDLR (OaVLDLR) and *Cyprinus carpio* VLDLR (CcVLDLR). The cysteines conserved across species are in boxes, the highly conserved SDE, EDE, YWS(V)D, FWA(T)D and FDNPVY sequences are underlined, the O-glycosylation domain is double-underlined, the little conserved domains are in gray.

(1,068 base pairs) and HnERRγ_X2 (1,101 base pairs) in the ovary (Fig. 6).

The deduced protein sequences of HnERβ_X1, HnERβ_X2, HnERβ_X3, HnERβ_X4, HnERγ, HnERRα, HnERRγ_X1 and HnERRγ_X2 comprised 558, 551, 427, 508, 481, 432, 356 and 367 amino acids, respectively.

The secondary structure of HnERs (Fig. 7) and HnERRs (Fig. 8) was composed of five domains, i.e., i) the domain at the N-terminus (A/B) that contained the first AF-1 activation function, ii) the DNA-binding domain (DBD), iii) flanking sequence D which promotes the rotation of the DBD (Aranda and Pascual, 2001) and functions as a site of interaction with corepressor proteins (Kumar et al., 2011), iv) the ligand-binding domain (LBD) where 17β-estradiol (E2) binds to amino acids to induce the second AF-2 activation function (Aranda and Pascual, 2001; Kato et al., 1995), and v) a C-terminal region F whose specific function is as yet unknown in fish (Aranda and Pascual, 2001; Nelson and Habibi, 2013).

The four forms of HnERβ had identical DBD and D domains (Fig. 9a). The A/B domain of HnERβ_X3 had a reduced number (32) of little conserved amino acids at its N-terminal end, whereas the other HnERβs had an identical 163-aa A/B domain. HnERβ_X1 and HnERβ_X3 had identical LBDs, whereas the LBDs of HnERβ_X2 and HnERβ_X4 did not contain the sequence NMCVNSPE and the sequence located between Ser^{494} and Ala^{523} of HnERβ_X1, respectively. The F domain of HnERβ_X4 contained 14 amino acids and differed from the other HnERβs, which had an identical 35-aa F domain. HnERβ and HnERγ shared 58 to 68% amino acid identity and had no common domain. HnERs shared less than 48% identity with HnERRs. The different HnERRγ shared 78% identity with one another and less than 59% identity with ERRα. However, the cores of the DBDs of the different receptors shared greater identity (Fig. 9b): 89% between ERβ and ERγ, 68 to 70% between ERs and ERRs, 97% between ERRγ_X1 and ERRγ_X2, and 94% between ERRγ and ERRα. The DBD core of all *H. niloticus* receptors (ERs and ERRs) was composed of 66 amino acids including six cysteines followed by a COOH-terminal extension (CTE) composed of 10 (HnERγ) to 25 (all ERRs) amino acids. The LBD was much more variable than the DBD core: HnERβ and HnERγ shared 76 to 77% identity, while HnERs shared less than 33% identity with HnERRs.

A comparison of *H. niloticus* ERβ with those of other teleost species showed that the core of the DBD of HnERβ was identical to that of ERβ of *S. formosus* (XP_018583081.1) and *O. bicirrhosum* (BAT68972. 1) and shared 98% identity with the EBβ DBD core of *O. mykiss* (CAC06714),

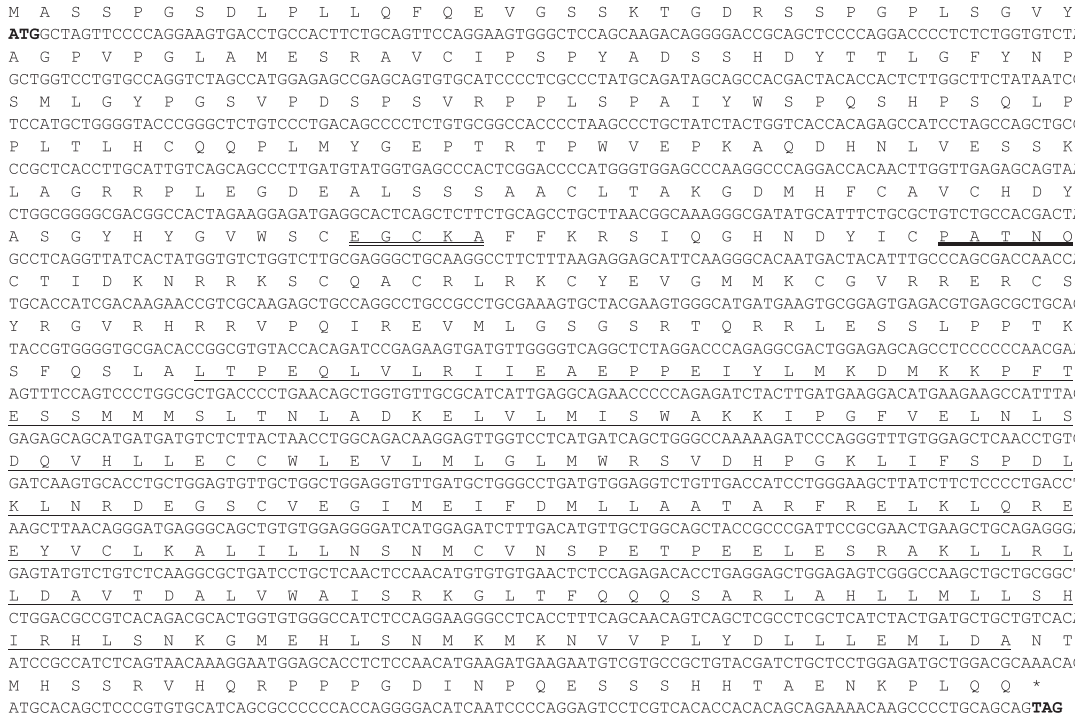

Fig. 5a. Nucleotide sequence and deduced protein sequence of *Heterotis niloticus* (Hn) estrogen receptors beta (ERβ_X1) identified in the liver and ovary of *Heterotis niloticus*. The start codon and the stop codon are in bold type. The A/B region in N-terminal position is followed by the DNA-binding domain (DBD) located in the boxes and containing a double-underlined P-box and a single-underlined D-box. The ligand-binding domain (LBD) is underlined and followed by an F-domain in C-terminal position. The amino acids implied in the binding with estradiol are circled.

O. aureus (ACF75103) and *C. carpio* (BAF99814) (Fig. 9c). The LBD of HnERβ_X1 shared 95% identity with *S. formosus* LBD, 93% with *O. bicirrhosum* LBD, and 82% with *O. mykiss*, *O. aureus* and *C. carpio* LBDs. A comparison of HnERγ DBD/LBD sequences with the ERγs of other species showed 89/74% identity with *O. mykiss* (NP_001118042), 89/72% with *O. aureus* (ACF75102.1), and 89/70% with *C. carpio* (BAF99813.1) ERγ DBDs/LBDs (Fig. 9d). Amino acid identities across ERγs were lower than across ERβs.

The P box involved in DNA recognition at the DBD level (Aranda and Pascual, 2001) was represented by two peptides: EGCKA at the level of HnERs and EACKA at the level of HnERRs. The D box involved in receptor dimerization (Aranda and Pascual, 2001) was represented by four peptides: PATNQ (HnERβ), PANNQ (HnERγ), PASNE (HnERRα), and PATNE (HnERRγ).

The six amino acids involved in the binding with E2 (Sabo-Attwood et al., 2004) were present in HnERβ_X1 (Glu^{330}, Arg^{371}, Gly^{498}, His^{501}, Leu^{502} and Met^{505}), HnERβ_X2 (Glu^{330}, Arg^{371}, Gly^{498}, His^{501}, Leu^{502} and Met^{505}), HnERβ_X3 (Glu^{330}, Arg^{371}, Gly^{498}, His^{501}, Leu^{502} and Met^{505}) and HnERγ (Glu^{273}, Arg^{314}, Gly^{441}, His^{444}, Leu^{445} and Met^{448}) (Fig. 5 and Fig. 9). HnERβ_X2 had only two conserved amino acids involved in the binding with E2 (Glu^{330}, Arg^{371}), which were absent in HnERRs. The LBDs of HnERβ_X1, HnERβ_X2, HnERβ_X3 and HnERγ (Fig. 7) had a tertiary structure composed of 12 α-helices, with helix 12 located between Leu^{513} and Leu^{521}, Leu^{506} and Leu^{514}, Leu^{382} and Leu^{390}, and Tyr^{457} and Asp^{465}, respectively. As for the LBD of HnERβ_X2, it did not have helix 12, and its helix 11 was much smaller than the helices 11 of the other HnERβs.

3.1.4. Prolactin (PRL)

Prolactin (PRL) is a multifunctional polypeptide neurohormone whose reproductive function in fish includes the reproductive cycle and development, spawning, and parental care (Whittington and Wilson, 2013). Expression of *H. niloticus* prolactin (HnPRL) in the pituitary gland was identified from a 627 base pairs transcript encoding 209 amino acids (Fig. 10a).

The protein precursor of HnPRL had a 25-aa signal peptide (Met^{1}-Cys^{25}), and four major α-helices representing 56% of the total protein, i.e., helix 1 (Phe^{28}-His^{56}), helix 2 (Glu^{69}-Ser^{115}), helix 3 (Gln^{122}-Lys^{149}) and helix 4 (Asp^{169}-Val^{201}), as well as two short α-helices between helix 1 and helix 2: helix 1′ (Val^{60}-Lys^{62}) and helix 1″ (Lys^{81}-Ser^{86}) (Fig. 10b). The tertiary structure of HnPRL showed that the four major helices formed two antiparallel pairs, i.e., helix 1/helix 4 and helix 2/helix 3 (Fig. 10c). The mature HnPRL contained four cysteines, one between helix 1′ and helix 1″ (Cys^{70}), two in helix 4 (Cys^{182} and Cys^{199}) and one at the end of the sequence (Cys^{209}). The alignments showed high or medium amino acid identity between HnPRL and the PRLs of *S. formosus* (XP_018592416) (74%), *C. carpio* (CAA37063) (58%), *O. mykiss* (AAA49611) (53%) and *O. niloticus* (AAA53281) (49%), but the four cysteines were conserved (Fig. 10d).

3.1.5. Prolactin receptor (PRLR)

The HnPRL receptor (HnPRLR) was expressed in the ovary of females in vitellogenesis and was represented by a 1,767 base pairs transcript encoding a 589-aa protein (Fig. 11a). HnPRLR consisted of a 25-aa signal peptide (Met^{1}-Leu^{25}) followed by three domains (Fig. 11b), one 207-aa extracellular domain (EC) (Ser^{26}-Arg^{232}), one 23-aa transmembrane domain (TM) (Ser^{233}-Ile^{255}), and one 334-aa intracellular domain (IC) (Asn^{256}-Tyr^{589}). Alignment of HnPRLR sequences with the PRLRs of other teleost species showed 85% amino acid identity with *S. formosus* (XP_018604730), 57% with *O. mykiss* (NP_001118071), 55% with *C. carpio* (AAK95833), and 51% with *O. niloticus* (NP_001266477) (Fig. 11c).

3.1.6. Alpha subunit of the glycoprotein hormone (GPα)

An α-subunit of the glycoprotein hormone (GPα) common to the luteinizing hormone (LH) and the follicle-stimulating hormone (FSH) (Acharjee et al., 2015) was characterized in the pituitary gland of *H. niloticus*. The nucleic sequence of HnGPα comprised 345 base pairs encoding a 115-aa protein (Fig. 12a). The secondary structure showed the presence of a 24-aa signal peptide and three α-loop domains of 25,

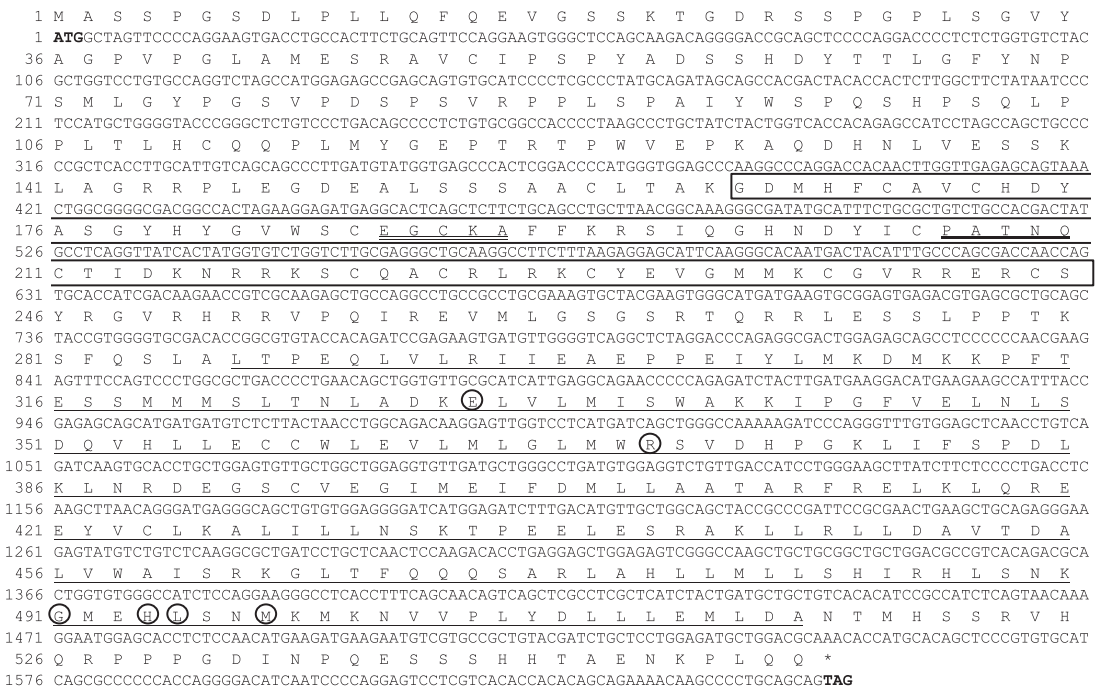

Fig. 5b. Nucleotide sequences and deduced protein sequences of *Heterotis niloticus* (Hn) estrogen receptors beta (ERβ_X2 and ERβ_X3) identified in the liver of *Heterotis niloticus*. The start codon and the stop codon are in bold type. The A/B region in N-terminal position is followed by the DNA-binding domain (DBD) located in the boxes and containing a double-underlined P-box and a single-underlined D-box. The ligand-binding domain (LBD) is underlined and followed by an F-domain in C-terminal position. The amino acids implied in the binding with estradiol are circled.

28 and 23 amino acids, respectively (Fig. 12b). HnGPα had ten cysteines (Cys^1 to Cys^{10}) and α-loop 1 was located between Cys^1 and Cys^4, α-loop 2 between Cys^5 and Cys^6, and α-loop3 between Cys^7 and Cys^8 (Fig. 12a). Two N-glycosylation sites were predicted, one in α-loop 2 (NITS) and the other in α-loop 3 (NHTD).

Sequence comparison showed that HnGPα was identical to *A. gigas* GPα (AIG51239) and that some domains were common with *S. formosus* (XP_018588619.1) (82%), *C. carpio* (CAA39852) (74%), *O. niloticus* (AAP49577) (65%) and *O. mykiss* (BAB17685) GPα (61%) (Fig. 12c).

3.1.7. β-subunit of the follicle-stimulating hormone (FSHβ)

The β-subunit transcript of *H. niloticus* follicle-stimulating hormone (HnFSHβ) had a 381 base pairs sequence encoding a 127-aa protein (Fig. 13a).

The secondary structure of HnFSHβ was composed of a 19-aa signal peptide, three β-loop domains 1, 2 and 3 of 26, 20 and 17 amino acids, respectively, and one 18-aa "seat-belt" loop (Fig. 13b). The mature HnFSHβ had 12 cysteines (Cys^1-Cys^{12}). β-loop 1 was located between Cys^1 and Cys^4, β-loop 2 between Cys^5 and Cys^6, β-loop 3 between Cys^7

(ERβ_X4)

```
   1 M  A  S  S  P  G  S  D  L  P  L  L  Q  F  Q  E  V  G  S  S  K  T  G  D  R  S  S  P  G  P  L  S  G  V  Y
   1 ATGGCTAGTTCCCCAGGAAGTGACCTGCCACTTCTGCAGTTCCAGGAAGTGGGCTCCAGCAAGACAGGGGACCGCAGCTCCCCAGGACCCCTCTCTGGTGTCTAC
  36 A  G  P  V  P  G  L  A  M  E  S  R  A  V  C  I  P  S  P  Y  A  D  S  S  H  D  Y  T  T  L  G  F  Y  N  P
 106 GCTGGTCCTGTGCCAGGTCTAGCCATGGAGAGCCGAGCAGTGTGCATCCCCTCGCCCTATGCAGATAGCAGCCACGACTACACCACTCTTGGCTTCTATAATCCC
  71 S  M  L  G  Y  P  G  S  V  P  D  S  P  S  V  R  P  P  L  S  P  A  I  Y  W  S  P  Q  S  H  P  S  Q  L  P
 211 TCCATGCTGGGGTACCCGGGCTCTGTCCCTGACAGCCCCTCTGTGCGGCCACCCCTAAGCCCTGCTATCTACTGGTCACCACAGAGCCATCCTAGCCAGCTGCCC
 106 P  L  T  L  H  C  Q  Q  P  L  M  Y  G  E  P  T  R  T  P  W  V  E  P  K  A  Q  D  H  N  L  V  E  S  S  K
 316 CCGCTCACCTTGCATTGTCAGCAGCCCTTGATGTATGGTGAGCCCACTCGGACCCCATGGGTGGAGCCCAAGGCCCAGGACCACAACTTGGTTGAGAGCAGTAAA
 141 L  A  G  R  R  P  L  E  G  D  E  A  L  S  S  S  A  A  C  L  T  A  K  G  D  M  H  F  C  A  V  C  H  D  Y
 421 CTGGCGGGGCGACGGCCACTAGAAGGAGATGAGGCACTCAGCTCTTCTGCAGCCTGCTTAACGGCAAAGGGCGATATGCATTTCTGCGCTGTCTGCCACGACTAT
 176 A  S  G  Y  H  Y  G  V  W  S  C  E  G  C  K  A  F  F  K  R  S  I  Q  G  H  N  D  Y  I  C  P  A  T  N  Q
 526 GCCTCAGGTTATCACTATGGTGTCTGGTCTTGCGAGGGCTGCAAGGCCTTCTTTAAGAGGAGCATTCAAGGGCACAATGACTACATTTGCCCAGCGACCAACCAG
 211 C  T  I  D  K  N  R  R  K  S  C  Q  A  C  R  L  R  K  C  Y  E  V  G  M  M  K  C  G  V  R  R  E  R  C  S
 631 TGCACCATCGACAAGAACCGTCGCAAGAGCTGCCAGGCCTGCCGCCTGCGAAAGTGCTACGAAGTGGGCATGATGAAGTGCGGAGTGAGACGTGAGCGCTGCAGC
 246 Y  R  G  V  R  H  R  R  V  P  Q  I  R  E  V  M  L  G  S  G  S  R  T  Q  R  R  L  E  S  S  L  P  P  T  K
 736 TACCGTGGGGTGCGACACCGGCGTGTACCACAGATCCGAGAAGTGATGTTGGGGTCAGGCTCTAGGACCCAGAGGCGACTGGAGAGCAGCCTCCCCCCAACGAAG
 281 S  F  Q  S  L  A  L  T  P  E  Q  L  V  L  R  I  I  E  A  E  P  P  E  I  Y  L  M  K  D  M  K  K  P  F  T
 841 AGTTTCCAGTCCCTGGCGCTGACCCCTGAACAGCTGGTGTTGCGCATCATTGAGGCAGAACCCCCAGAGATCTACTTGATGAAGGACATGAAGAAGCCATTTACC
 316 E  S  S  M  M  M  S  L  T  N  L  A  D  K  E  L  V  L  M  I  S  W  A  K  K  I  P  G  F  V  E  L  N  L  S
 946 GAGAGCAGCATGATGATGTCTCTTACTAACCTGGCAGACAAGGAGTTGGTCCTCATGATCAGCTGGGCCAAAAAGATCCCAGGGTTTGTGGAGCTCAACCTGTCA
 351 D  Q  V  H  L  L  E  C  C  W  L  E  V  L  M  L  G  L  M  W  R  S  V  D  H  P  G  K  L  I  F  S  P  D  L
1051 GATCAAGTGCACCTGCTGGAGTGTTGCTGGCTGGAGGTGTTGATGCTGGGCCTGATGTGGAGGTCTGTTGACCATCCTGGGAAGCTTATCTTCTCCCCTGACCTC
 386 K  L  N  R  D  E  G  S  C  V  E  G  I  M  E  I  F  D  M  L  L  A  A  T  A  R  F  R  E  L  K  L  Q  R  E
1156 AAGCTTAACAGGGATGAGGGCAGCTGTGTGGAGGGGATCATGGAGATCTTTGACATGTTGCTGGCAGCTACCGCCCGATTCCGCGAACTGAAGCTGCAGAGGGAA
 421 E  Y  V  C  L  K  A  L  I  L  L  N  S  N  M  C  V  N  S  P  E  T  P  E  E  L  E  S  R  A  K  L  L  R  L
1261 GAGTATGTCTGTCTCAAGGCGCTGATCCTGCTCAACTCCAACATGTGTGTGAACTCTCCAGAGACACCTGAGGAGCTGGAGAGTCGGGCCAAGCTGCTGCGGCTG
 456 L  D  A  V  T  D  A  L  V  W  A  I  S  R  K  G  L  T  F  Q  Q  Q  S  A  R  L  A  H  L  L  M  L  L  S  H
1366 CTGGACGCCGTCACAGACGCACTGGTGTGGGCCATCTCCAGGAAGGGCCTCACCTTTCAGCAACAGTCAGCTCGCCTCGCTCATCTACTGATGCTGCTGTCACAC
 491 I  R  H  L  R  V  P  V  V  W  E  F  P  V  F  P  G  H  *
1471 ATCCGCCATCTCAGGGTTCCGGTGGTCTGGGAATTTCCTGTATTCCCTGGGCACTAA
```

(ERγ)

```
   1 M  L  L  S  L  V  Y  S  S  A  T  P  I  N  I  P  S  P  L  M  D  N  C  Q  D  Y  S  V  Q  Q  Y  A  C  P  G
   1 ATGTTGCTGTCCCTGGTCTACAGCTCTGCTACGCCTATCAACATCCCATCTCCATTAATGGACAACTGCCAAGATTACTCAGTACAGCAATATGCTTGCCCTGGT
  36 L  S  D  S  S  F  L  D  P  P  F  D  W  K  P  D  S  N  F  L  S  I  V  K  K  I  T  P  P  S  S  P  L  A  E
 106 CTCAGTGATTCCAGCTTTTTGGACCCACCGTTTGATTGGAAACCCGACAGCAATTTCTTGTCAATAGTCAAGAAAATTACACCACCCAGCAGTCCTCTGGCTGAG
  71 P  E  P  Q  S  L  F  S  T  D  K  Q  Q  Q  R  R  K  R  K  T  N  G  M  C  R  V  K  S  L  Q  P  F  P  G  D
 211 CCTGAGCCCCAAAGTCTTTTCTCTACTGATAAACAGCAGCAGAGAAGAAAGAGGAAAACGAATGGCATGTGCAGAGTGAAATCACTGCAGCCTTTCCCAGGGGAT
 106 E  T  E  K  R  L  C  F  V  C  K  D  Y  A  S  G  Y  H  Y  S  V  W  S  C  E  G  C  K  A  F  F  K  R  S  I
 316 GAGACAGAAAAACGCTTGTGTTTTGTGTGCAAGGACTATGCTTCAGGGTACCACTACAGTGTGTGGTCCTGTGAGGGCTGCAAGGCCTTTTTTAAGAGGAGCATC
 141 Q  G  P  T  D  Y  I  C  P  A  N  N  Q  C  T  I  D  K  S  R  R  K  S  C  Q  A  C  R  L  R  K  C  Y  E  V
 421 CAAGGGCCTACTGATTATATCTGTCCAGCAAACAACCAGTGCACCATTGACAAGAGCCGCCGCAAAAGCTGCCAGGCTTGTCGTCTTCGCAAGTGCTATGAAGTG
 176 G  M  V  K  N  G  V  R  R  E  R  R  P  S  Q  E  V  Q  G  I  R  M  C  Q  H  L  S  V  M  G  E  E  A  S  L
 526 GGCATGGTCAAGAATGGTGTCCGAAGAGAACGTCGTCCCTCCCAGGAGGTGCAAGGCATACGTATGTGCCAGCACCTTTCAGTAATGGGGGAAGAAGCCAGTCTT
 211 L  S  Q  E  K  S  H  L  S  A  S  L  E  E  T  Q  Q  L  L  F  S  P  K  Q  L  I  S  R  F  L  E  V  E  P  P
 631 CTCTCACAAGAGAAGTCCCACCTTTCTGCATCTTTGGAGGAGACCCAACAACTGTTGTTCAGCCCAAAGCAGCTCATTTCTCGCTTTCTGGAGGTAGAACCTCCA
 246 D  I  Y  L  M  W  D  L  K  K  P  F  T  E  V  S  M  M  T  S  L  T  H  L  A  D  K  E  L  F  Y  M  I  S  W
 736 GACATTTACTTGATGTGGGACCTGAAAAAACCCTTTACAGAGGTCAGCATGATGACATCACTCACTCACCTGGCAGACAAAGAGCTCTTTTACATGATCAGCTGG
 281 A  K  K  I  P  G  F  T  E  L  G  L  S  D  Q  V  H  L  L  D  C  C  W  L  E  V  L  M  L  G  L  I  W  R  S
 841 GCTAAAAAGATCCCAGGCTTCACTGAGCTCGGCCTATCAGACCAGGTACACTTACTAGACTGTTGCTGGCTGGAGGTGCTAATGCTGGGCTTGATCTGGAGGTCT
 316 V  D  H  P  G  K  L  I  F  S  P  D  L  I  L  S  R  D  E  G  S  C  I  E  G  I  L  E  I  F  D  M  L  L  A
 946 GTGGACCACCCTGGAAAGCTGATCTTCTCCCCTGACCTCATCCTTAGCAGGGATGAGGGCAGCTGCATTGAGGGGATTCTGGAGATCTTTGACATGCTGCTTGCA
 351 A  T  S  R  F  R  E  L  R  L  Q  K  E  E  Y  V  C  L  K  A  L  V  L  L  N  S  S  I  Y  L  Y  T  P  N  G
1051 GCAACTTCTCGCTTCCGGGAGCTGCGGCTACAAAAGGAGGAATATGTCTGCCTGAAGGCCCTGGTTCTCCTCAACTCTAGCATATACCTCTATACCCCAAATGGT
 386 G  Y  E  H  Q  S  R  A  K  L  Q  Q  L  L  N  A  V  T  D  A  L  I  W  V  I  A  K  L  G  L  S  F  Q  Q  Q
1156 GGGTATGAGCATCAGAGCAGGGCCAAGCTACAACAGCTGCTGAATGCTGTGACTGATGCACTTATATGGGTTATCGCTAAGCTGGGACTGAGCTTCCAGCAGCAG
 421 S  A  R  L  A  H  L  I  M  L  L  S  H  I  R  H  V  S  N  K  G  M  D  H  L  H  C  M  K  M  K  K  V  V  P
1261 TCCGCCCGACTGGCCCATCTCATCATGCTGTTGTCACACATCCGCCATGTCAGTAACAAAGGAATGGACCACCTGCACTGCATGAAGATGAAGAAAGTAGTGCCA
 456 F  Y  D  L  L  L  E  M  L  D  A  H  V  M  H  N  S  R  I  L  S  P  E  P  S  G  *
1366 TTCTATGATCTCCTCTTGGAAATGCTGGACGCCCATGTTATGCACAACAGCCGCATTCTCAGTCCTGAGCCTTCAGGATGA
```

Fig. 5c. Nucleotide sequences and deduced protein sequences of *Heterotis niloticus* (Hn) estrogen receptors beta (ERβ_X4) and gamma (ERγ) identified in the ovary of *Heterotis niloticus*. The start codon and the stop codon are in bold type. The A/B region in N-terminal position is followed by the DNA-binding domain (DBD) located in the boxes and containing a double-underlined P-box and a single-underlined D-box. The ligand-binding domain (LBD) is underlined and followed by an F-domain in C-terminal position. The amino acids implied in the binding with estradiol are circled.

and Cys^{8}, and the "seat-belt" loop between Cys^{10} and Cys^{12} (Fig. 13a). A probable N-glycosylation site (NVSI) was found in β-loop 1 between Cys^{1} and Cys^{2}.

A comparison of the FSHβ sequences showed that HnFSHβ shared a high identity with FSHβ of *A. gigas* (AIA09918) (86%), but much less with *S. formosus* (KPP62167) (69%), *C. carpio* (O13050) (51%), *O. mykiss* (BAB17686) (44%) and *O. niloticus* (AAP49575) (37%) FSHβs (Fig. 13c).

3.1.8. Follicle-stimulating hormone receptor (FSHR)

The HnFSH receptor (HnFSHR) was expressed in the ovary of females in vitellogenesis and corresponded to a 2,037 base pairs transcript encoding a 679-aa protein (Fig. 14a). HnFSHR had a 37-aa signal peptide (Met^{1} - Ala^{37}) followed by four extracellular domains (EC) of 322 (EC1), 21 (EC2), 19 (EC3) and 9 (EC4) amino acids, seven transmembrane domains of 23 amino acids each, and four intracellular domains (IC) of 8 (IC1), 20 (IC2), 24 (IC3) and 58 (IC4) amino acids

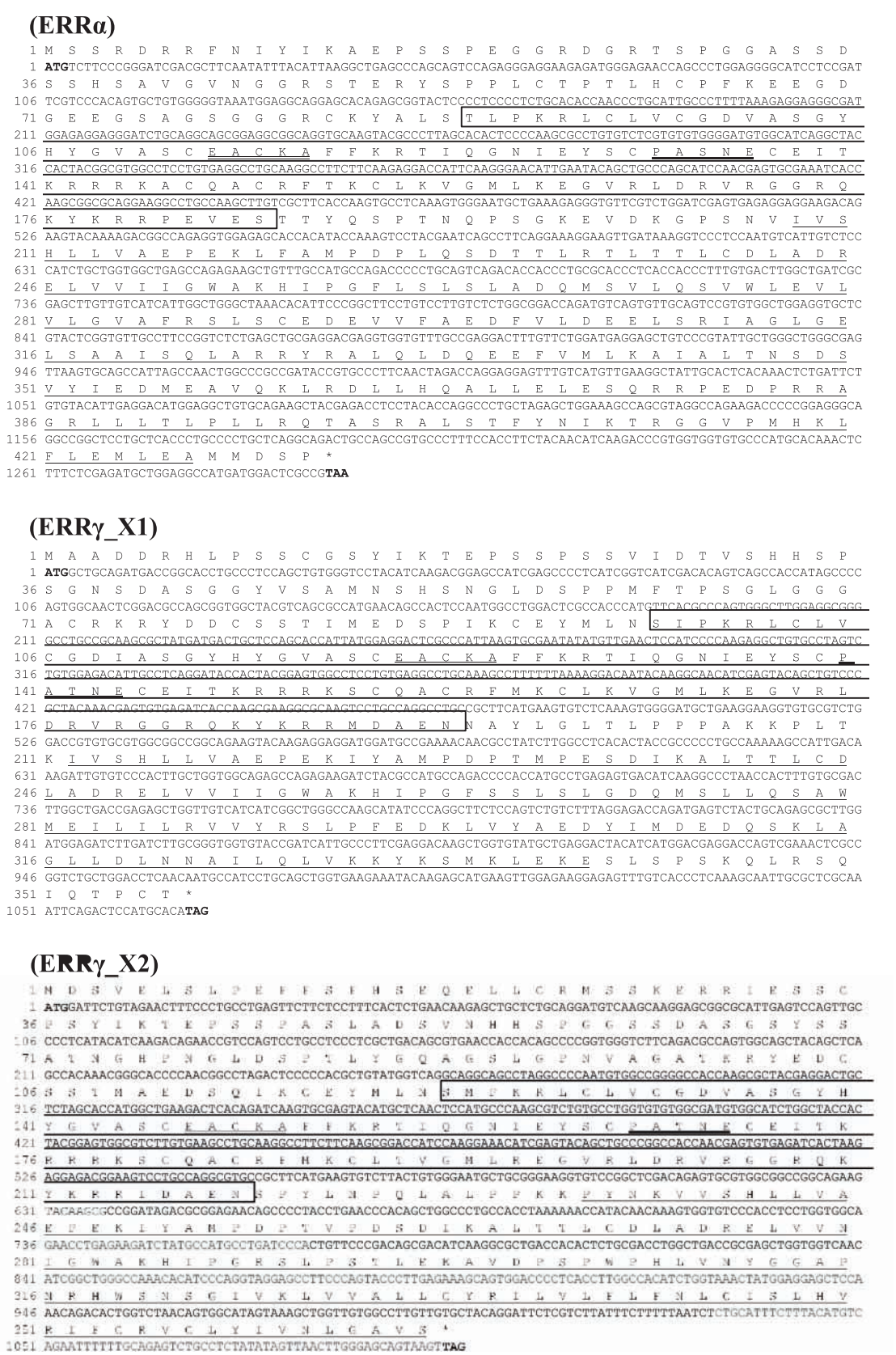

Fig. 6. Nucleotide sequences and deduced protein sequences of the receptors related to *Heterotis niloticus* estrogen receptors alpha (HnERRα) and gamma (ERRγ) (HnERRγ_X1 and HnERγ_X2). The start codon and the stop codon are in bold type. The A/B region in N-terminal position is followed by the DNA-binding domain (DBD) located in the boxes and containing a double-underlined P-box and a single-underlined D-box. The ligand-binding domain (LBD) is underlined and followed by an F-domain in C-terminal position.

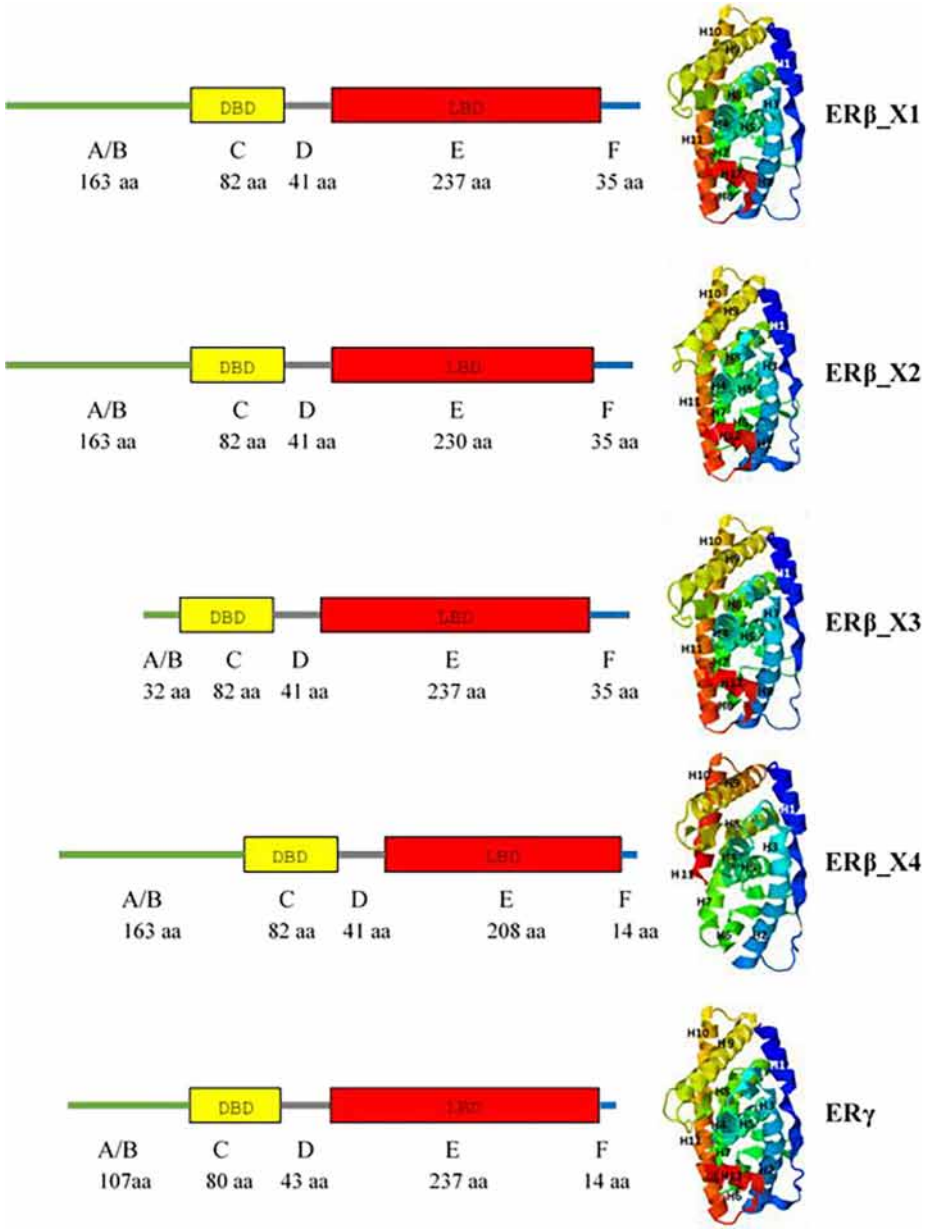

Fig. 7. Secondary structure of *Heterotis niloticus* estrogen receptors beta (ERβ_X1, ERβ_X2, ERβ_X3 and ERβ_X4) and gamma (ERγ), with the five characteristic domains and the corresponding numbers of amino acids (aa). The α-helix structure of the LBD obtained with I-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) shows that the 12 helices (H1 to H12) are conserved in ERβ_X1, ERβ_X2, ERβ_X3 and ERγ, while *H*11 is reduced and H12 is absent in ERβ_X4.

(Fig. 14b). Alignment of HnFSHR sequences with FSHRs of other teleost species showed 64.94% amino acid identity with *O. mykiss* (NP_001117799) and 71.23% with *C. carpio* (XP_018940722) (Fig. 14c).

3.1.9. β-subunit of the luteinizing hormone (LHβ)

The β-subunit of *H. niloticus* luteinizing hormone (HnLHβ) had a 423 base pairs sequence encoding a 141-aa protein (Fig. 15a). The secondary structure made it possible to distinguish a 24-aa signal peptide, three β-loop domains 1, 2 and 3 of 26, 20 and 17 amino acids, respectively, and one 18-aa seat-belt loop (Fig. 15b). Like HnFSHβ, the mature HnLHβ had twelve cysteines (Cys^1-Cys^{12}) and β-loop 1 was located between Cys^1 and Cys^4, β-loop 2 between Cys^5 and Cys^6, β-loop 3 between Cys^7 and Cys^8, and the "seat-belt loop" between Cys^{10} and Cys^{12} (Fig. 15a). A probable N-glycosylation site (NQTI) was located between Cys^1 and Cys^2. A comparison of LHβ sequences showed that HnLHβ was identical to *A. gigas* LHβ (LAO68014) and that it shared domains with *O. mykiss* (BAB17687) (79%), *S. formosus* (KPP64307) (77%), *C. carpio* (CAA42542) (74%) and *O. niloticus* (AAP49576) (68%) LHβs (Fig. 15c).

3.1.10. Luteinizing hormone receptor (LHR)

The HnLH receptor (HnLHR) was expressed in the ovary of females in vitellogenesis. It was a 2,124 base pairs transcript encoding a 708-aa protein (Fig. 16a). HnLHR had a 20-aa signal peptide (Met^1-His^{20}) followed by four extracellular domains (EC) of 361 (EC1), 21 (EC2), 19 (EC3) and 9 (EC4) amino acids, seven transmembrane domains of 23 amino acids each, and four intracellular domains (IC) of 8 (IC1), (IC2), 24 (IC3) and 65 (IC4) amino acids (Fig. 16b).

Alignment of the HnLHR sequences with the LHRs of other teleost species showed 80.37% amino acid identity with *S. formosus* (XP_018593689), 65.23% with *O. mykiss* (NP_001117798) and 62.77% with *O. niloticus* (XP_025753045) LHRs (Fig. 16c).

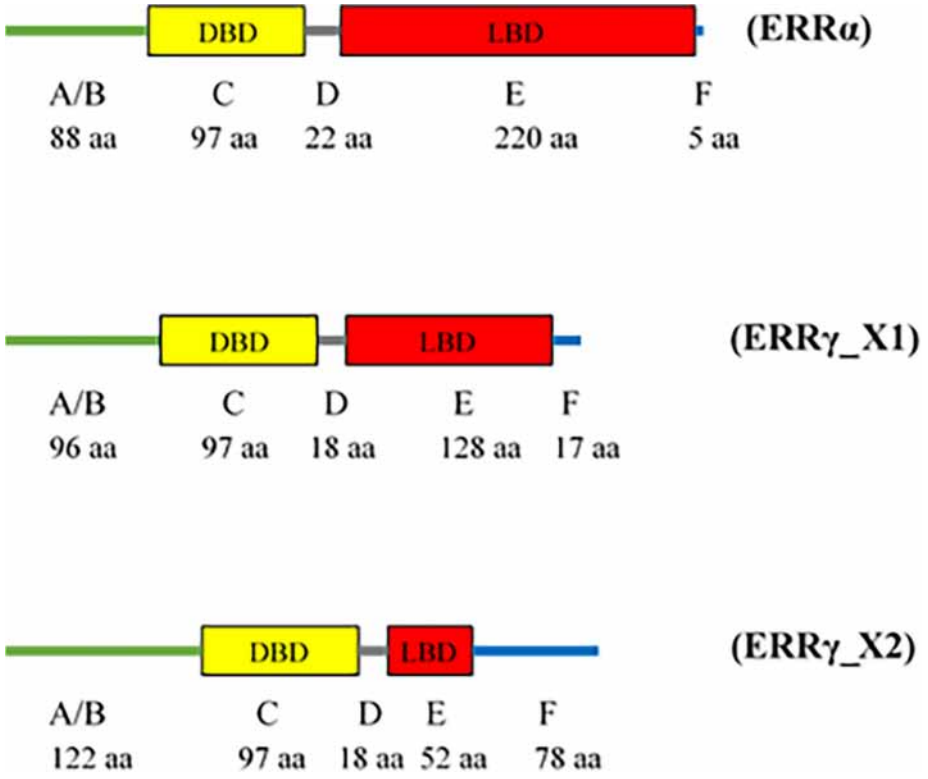

Fig. 8. Secondary structure of the receptors related to *Heterotis niloticus* estrogen receptors (ERRα, ERRγ_X1 et ERγ_X2), with the five characteristic domains and the corresponding numbers of amino acids (aa).

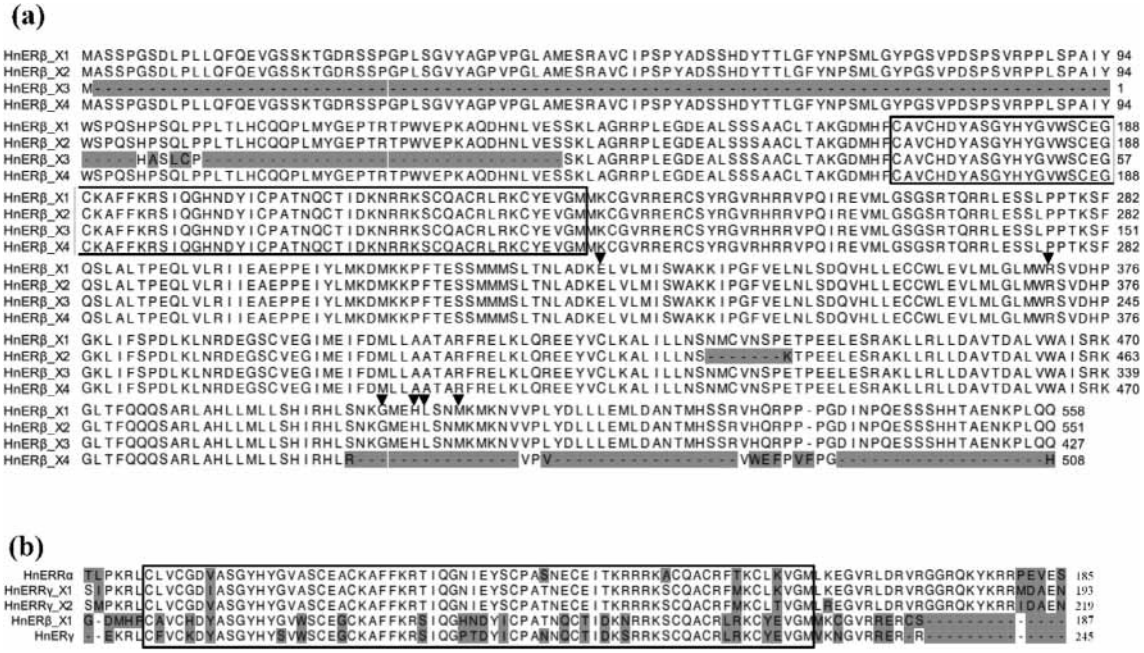

Fig. 9. Alignment of the deduced protein sequences of *Heterotis niloticus* estrogen receptors (ER) beta (a), and DNA-binding domains (DBD) of ERs and ERRs (HnERβ_X1, HnERβ_X2, HnERβ_X3, HnERβ_X4, HnERγ, HnERRα, HnERRγ_X1 and HnERRγ_X2) (b). The core of the DNA-binding domain (DBD) is in boxes, and the amino acids implied in the binding with 17 β estradiol are indicated by black triangles. The little conserved domains are in gray.

3.2. Detection of Vtg in blood and mucus

Mass spectrometry analysis detected the presence of Vtg in the blood and mucus of *H. niloticus* females in vitellogenesis (Fig. 17). No Vtg or any of its cleavage products was detected in the blood or mucus of males. Tryptic peptides of the LvI domain specific to HnVtg1 and HnVtg2 were detected in female blood samples, while only HnVtg1 peptides were detected in mucus samples (Fig. 17a). Mapping tryptic peptides from the N and C-terminal HnVtg1 sequences showed a coverage rate of 53.2% in blood proteins, and 31.9% in mucus proteins. As for the different domains of Vtg, the coverage of the peptides detected in the blood was 42.5% for LvI, 0.8% for Pv and CT, 8.2% for LvII, and

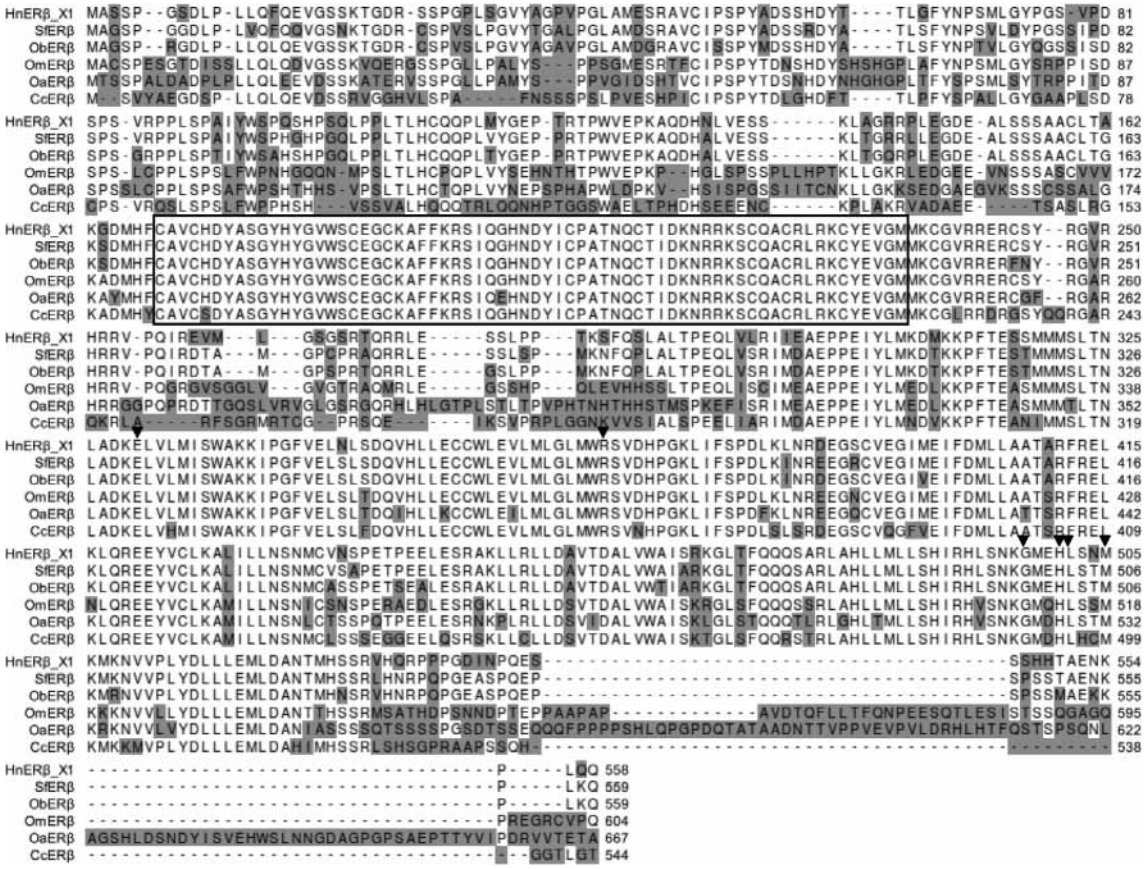

Fig. 9c. Alignment of the deduced protein sequences of *Heterotis niloticus* estrogen receptors (ER) beta (HnERβ_X1) with ERβ of *Scleropages formosus* (SfERβ), *Osteoglossum bicirrhosum* (ObERβ), *Oncorhynchus mykiss* (OmERβ and OmERγ), *Oreochromis aureus* (OaERβ and OaERγ) and *Cyprinus carpio* (CcERβ and CcERγ). The core of the DNA-binding domain (DBD) is in boxes, and the amino acids implied in the binding with 17 β estradiol are indicated by black triangles. The little conserved domains are in gray.

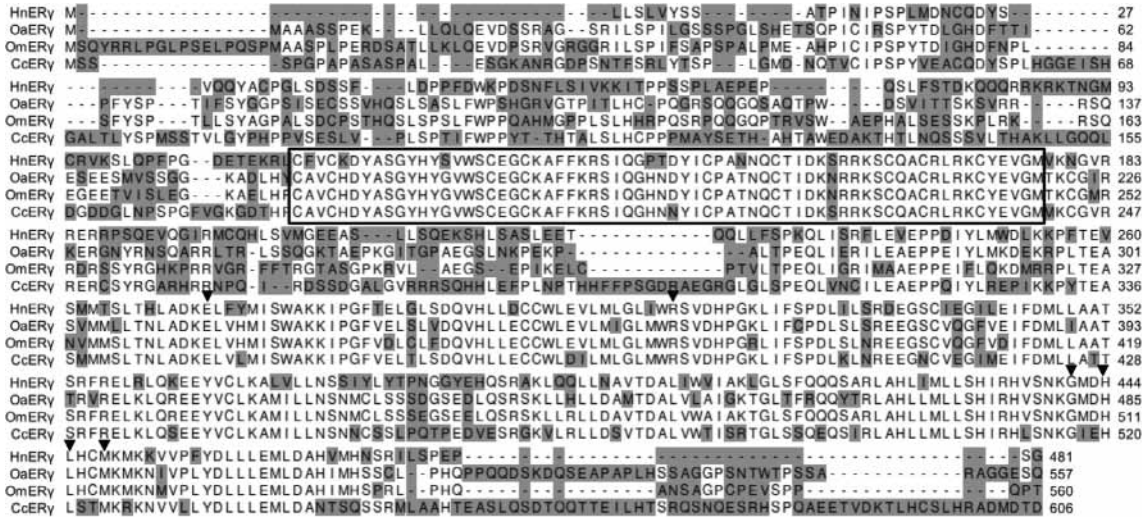

Fig. 9d. Alignment of the deduced protein sequences of *Heterotis niloticus* estrogen receptors (ER) gamma (HnERγ) with ERγ of *Oncorhynchus mykiss* (OmERγ), *Oreochromis aureus* (OaERγ) et *Cyprinus carpio* (CcERγ). The core of the DNA-binding domain (DBD) is in boxes, and the amino acids implied in the binding with 17 β estradiol are indicated by black triangles. The little conserved domains are in gray.

1% for β'-C. In mucus, the coverage of the detected peptides was 18.6% for LvI, 0.8% for Pv, 5.8% for LvII, 4.2% for β'-C, and 2.6% for CT (Fig. 17b).

4. Discussion

Vtg and some of the molecular actors involved in vitellogenesis and its regulation were characterized for the first time in the Osteoglossiforme *H. niloticus* using a combination of "omic" approaches. Comparisons were made in most cases with *S. formosus,* a

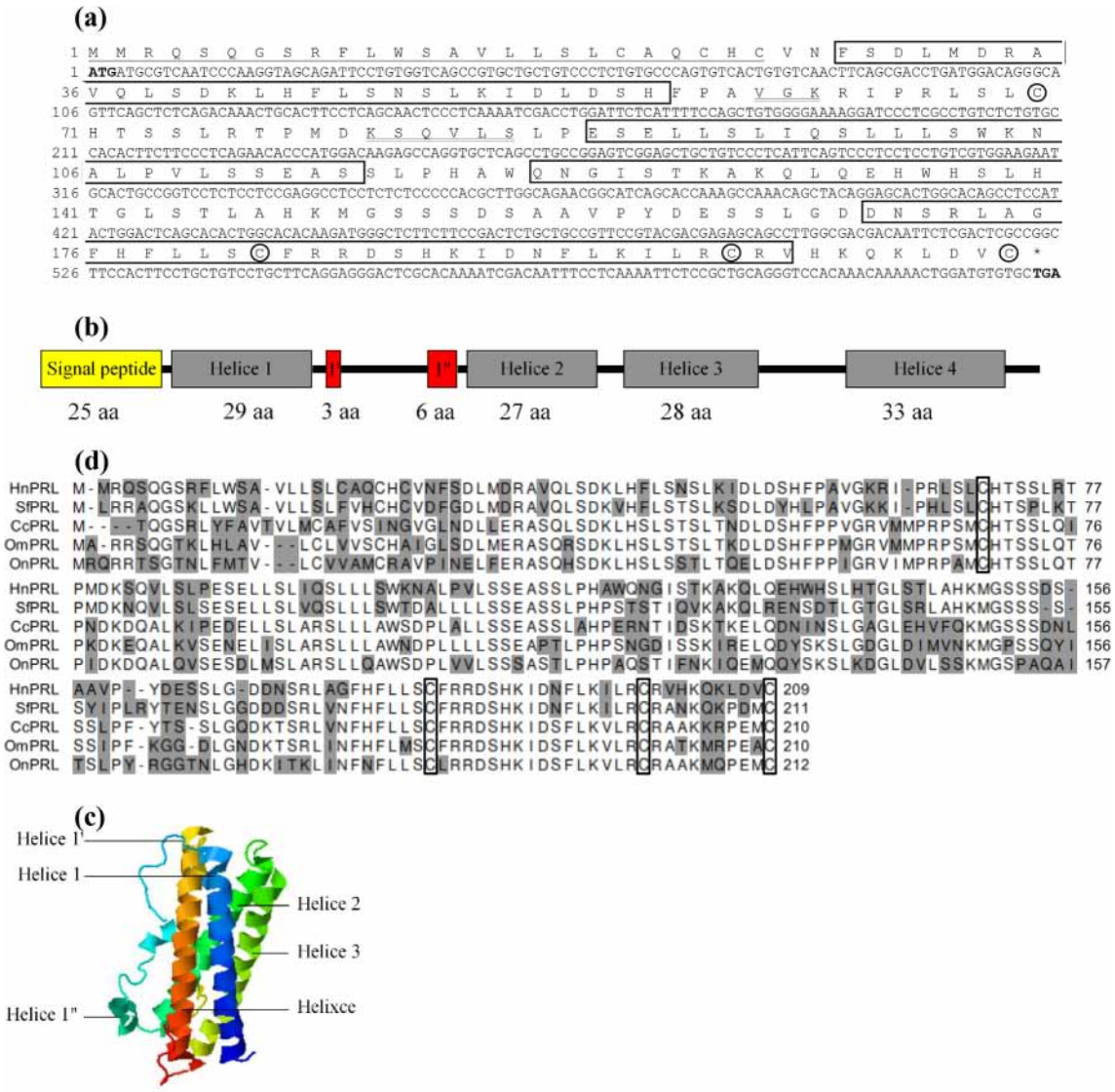

Fig. 10. Nucleotide sequences and deduced protein sequences (a) of *Heterotis niloticus* prolactin (HnPRL), secondary structure of HnPRL with the different domains (b), tertiary structure obtained with I-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) showing the α-helices of HnPRL (c), and alignment (d) with the PRLs of *Scleropages formosus* (SfPRL), *Oncorhynchus mykiss* (OmPRL), *Oreochromis niloticus* (OnPRL) and *Cyprinus carpio* (CcPRL). In figure (a), the start codon and the stop codon are in bold type, the signal peptide is underlined, the four major helices (1, 2, 3 and 4) are laid out successively in boxes, and the short helices 1′ and 1″ are double-underlined. In figure (d), the four conserved cysteines are in boxes, the little conserved domains are in gray.

phylogenetically close species, and with other phylogenetically more distant teleost species.

4.1. PRL and its receptor

Prolactin (PRL) is a polypeptide hormone secreted by the pituitary gland. Its importance in reproductive physiology has been underlined in many teleost species such as *A. gigas* (Marcos and Adalberto, 2015) and *Danio rerio* (Bu et al., 2015). Only one PRL expressed in the pituitary gland of *H. niloticus* was identified. By contrast, several isoforms issued from alternative splicing have been observed in mammals (Freeman et al., 2000) and in some non-mammalian species including teleost fish that express two distinct forms of PRL (PRL1 and PRL2) (Huang et al., 2009). As pointed out by Whittington and Wilson (2013), the expression of the PRL receptor in *H. niloticus* ovary suggests the involvement of this neurohormone in vitellogenesis and ovulation in this species. *H. niloticus* PRL (HnPRL) has a short 25-aa signal peptide like most PRLs of teleost fish (Bu et al., 2015), while mammalian (Skorupski and Kmieć, 2013) and bird (Bu et al., 2015) PRLs are made of up to 30 amino acids. The two cysteines in the N-terminal position capable of forming a disulfide bridge in the PRLs of mammals (Bu et al., 2015; Teilum et al., 2005), reptiles (Yasuda et al., 1990), Dipnoi (lungfish), chondrosteans (sturgeons) and tetrapods (Noso et al., 1993) are absent from HnPRL as in most teleosts (Bu et al., 2015; Watahiki et al., 1989). Although HnPRL shares variable amino acid identity (49 to 74%) with the PRLs of certain fish (Fig. 10d), it shares a common structure with the PRLs of vertebrate species. It contains four major α-helices (helices 1 to 4) representing more than 50% of the entire protein (Skorupski and Kmieć, 2013), and two short α-helices between helix 1 and helix 2 forming a loop (Brooks, 2012; Teilum et al., 2005) that plays a vital role in protein stabilization and ligand-receptor interaction (Goffin et al., 1996). Helix 1 and helix 4, on the one hand, and helix 2 and helix 3, on the other hand, form two antiparallel pairs (Brooks, 2012; Teilum et al., 2005). HnPRL contains four cysteines after propeller 1 that are well conserved across species and are likely to form two disulfide bridges (Brooks, 2012; Bu et al., 2015; Teilum et al., 2005). One is located between Cys^{70} and Cys^{182}, and the other at the C-terminal level of the protein between Cys^{199} and Cys^{209}. Similarly to the PRL of its sister species *A. gigas*,

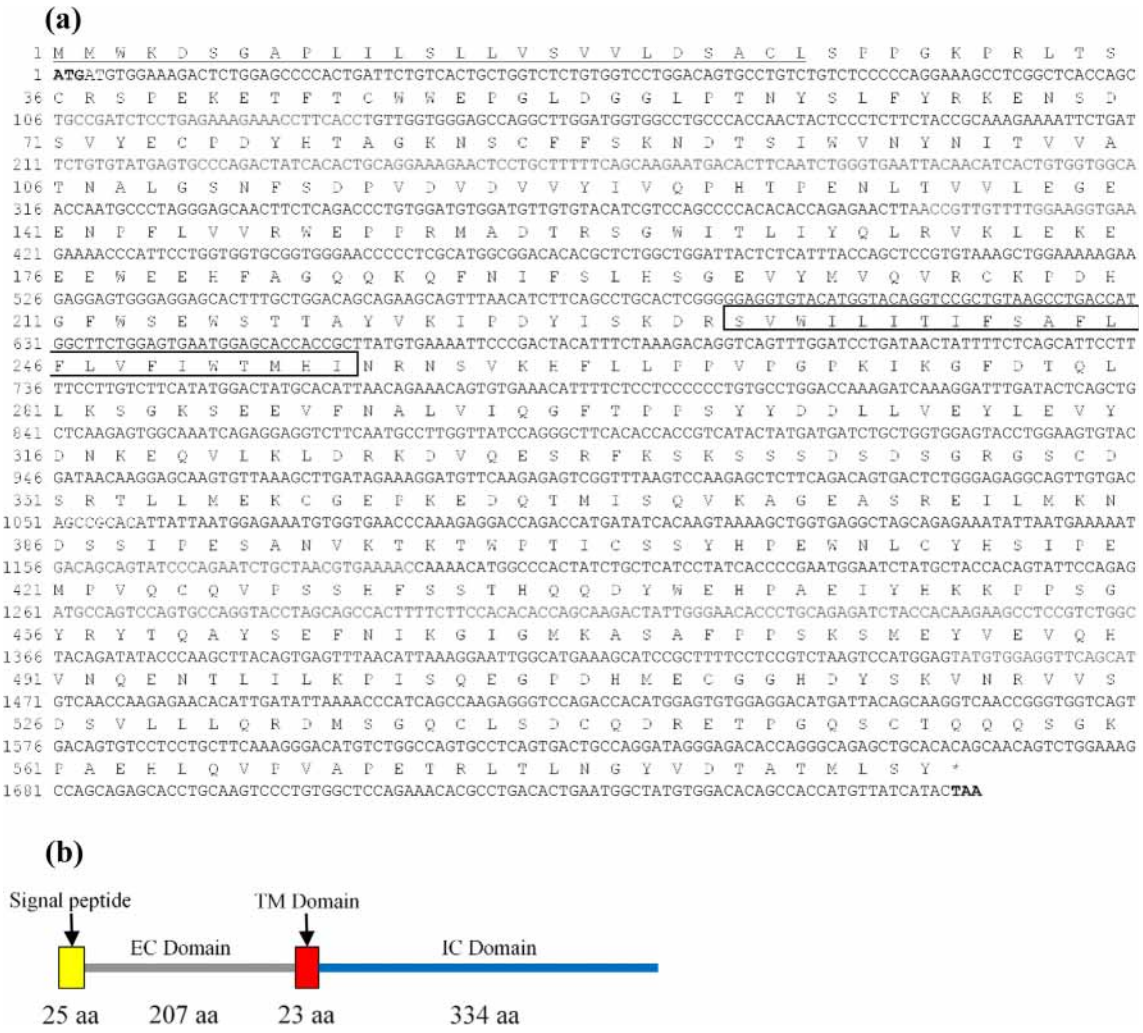

Fig. 11. Nucleotide sequences and deduced protein sequences of (a) *Heterotis niloticus* prolactin receptor (HnPRLR), and box modeling with the different domains (b). In figure (a), the start codon and the stop codon are in bold type, the signal peptide is underlined, the transmembrane domain (TM) in a box located between the extracellular domain (EC) in the N-terminal region and the intracellular domain (IC) in the C-terminal region.

where immunocytochemistry showed that PRL had little affinity with Salmonid anti-PRL anti serum (Borella et al., 2009), HnPRL shares little identity (53%) with *O. mykiss* PRL.

4.2. Gonadotropins and their receptors

H. niloticus FSH and LH expressed in the pituitary gland are heterodimeric glycoproteins like all gonadotropic hormones (GTHs), with a common α-subunit called glycoprotein hormone α (GPα) and specific β-subunits (Acharjee et al., 2015). A comparative analysis of the structure of *H. niloticus* GTH with other members of the teleost group showed that the GPα subunit consisted of three α loops and the β-subunits of three β loops with a "seat-belt" in the C terminal position. HnGPα had the ten conserved cysteines believed to be involved in the formation of five disulfide bridges, four proline residues, and two N-glycosylation sites at α loop 2 (NITS) and α loop 3 (NHTD). In catfish (*Heteropneust fossilis*), the N-glycosylation sites are located at loop 1 (NITS) and loop 2 (NHTD) (Acharjee et al., 2015). Only eight cysteines are conserved in the GPα of hagfish (Uchida et al., 2010). *A. gigas* GPα and LHβ, described by Faria et al., (2013) and Sevilhano et al., (2017), respectively, are identical to HnGPα and HnLHβ identified in this study, respectively. This result again demonstrates that there is a great phylogenetic proximity between these two species of Osteoglossomorphs previously underlined by some authors (Betancur-R et al., 2017; Guo-Qing and Wilson, 1996; Hilton, 2001, 2003; Lavoué, 2016; Lavoué and Sullivan, 2004; Nelson, 1968, 1969). This also completes the few studies reporting phylogenetic hypotheses based on molecular data (Kumazawa and Nishida, 2000). The present study shows that HnGPα shares between 61 and 82% amino acid identity with certain species of Osteoglossiformes, Salmoniformes, Perciformes and Cypriniformes. *A. gigas* GPα is identical to HnGPα and shares a higher amino acid identity (87 to 89.5%) with other teleost (Anguilliformes and Ostartiophysi) and Chondrosteans (Acipenseriformes) (Faria et al., 2013).

The position of the 12 cysteines and the only potential N-glycosylation site at LHβ β-loop 1 (NXTX) observed in most vertebrates (Bousfield and Dias, 2011; So et al., 2005) is very well conserved in *H. niloticus*. In catfish, the loop located between Cys^{10} and Cys^{11} of the "seat-belt" of LHβ contains the receptor-binding sites of this hormone, whereas the loop corresponding to Cys^{11} and Cys^{12} of LHβ stimulates FSH receptor activity (Vischer et al., 2004).

In contrast to the LHβ subunit, the FSHβ subunit was 86% conserved between *H. niloticus* and its sister species *A. gigas,* and diverged more with other species of Osteoglossiformes, Perciformes, Salmoniformes and Cypriniformes (37% to 69%). In *A. gigas*, the highest identity, based on amino acid sequences, is 61% identity with Anguilliformes for FSHβ and 76% with Cypriniformes for LHβ (Sevilhano et al., 2017). HnFSHβ had 12 conserved cysteines; the first four were located in β-loop 1 (Cys^{1} to Cys^{4}), Cys^{5} and Cys^{6} were located

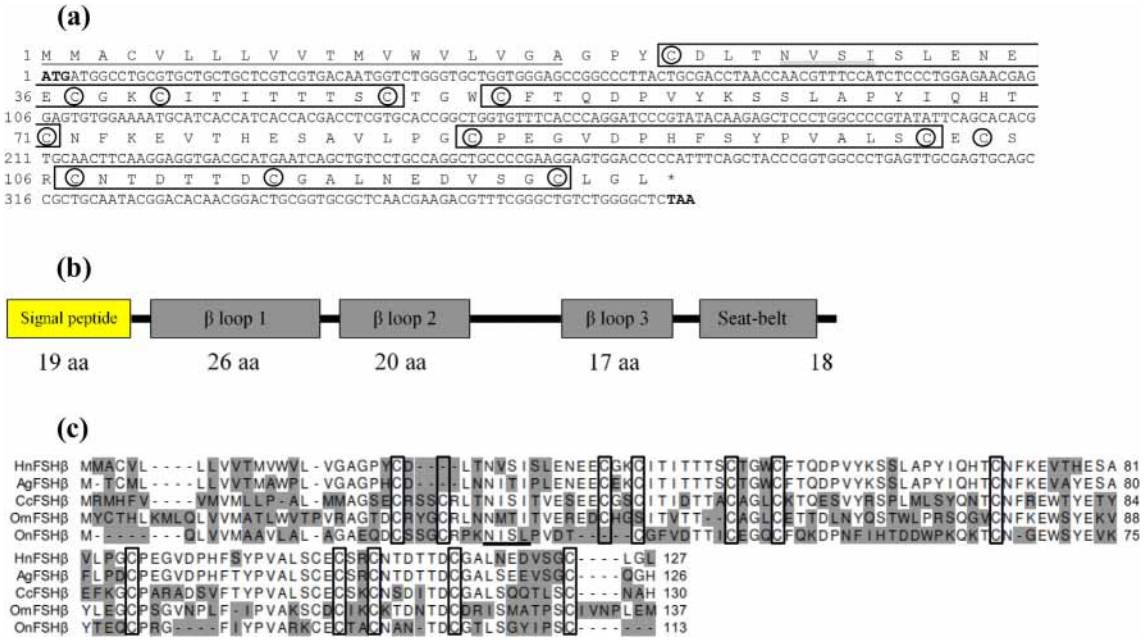

Fig. 12. Nucleotide sequences and deduced protein sequences of (a) the beta subunit of *Heterotis niloticus* follicle-stimulating hormone (HnFSHβ), box modeling of HnFSHβ with the different domains (b), and alignment (c) with *Arapaima gigas* FSHβ (Ag HnFSHβ), *Oncorhynchus mykiss* FSHβ (Om HnFSHβ), *Oreochromis niloticus* FSHβ (On HnFSHβ) and *Cyprinus carpio* FSHβ (Cc HnFSHβ). In figure (a), the start codon and the stop codon are in bold type, the signal peptide is underlined, β-loops 1, 2 and 3 and the "seat-belt" are laid out successively in boxes. The predicted N-glycosylation site is double-underlined. In figure (c), conserved cysteines across species are in boxes, the little conserved domains are in gray.

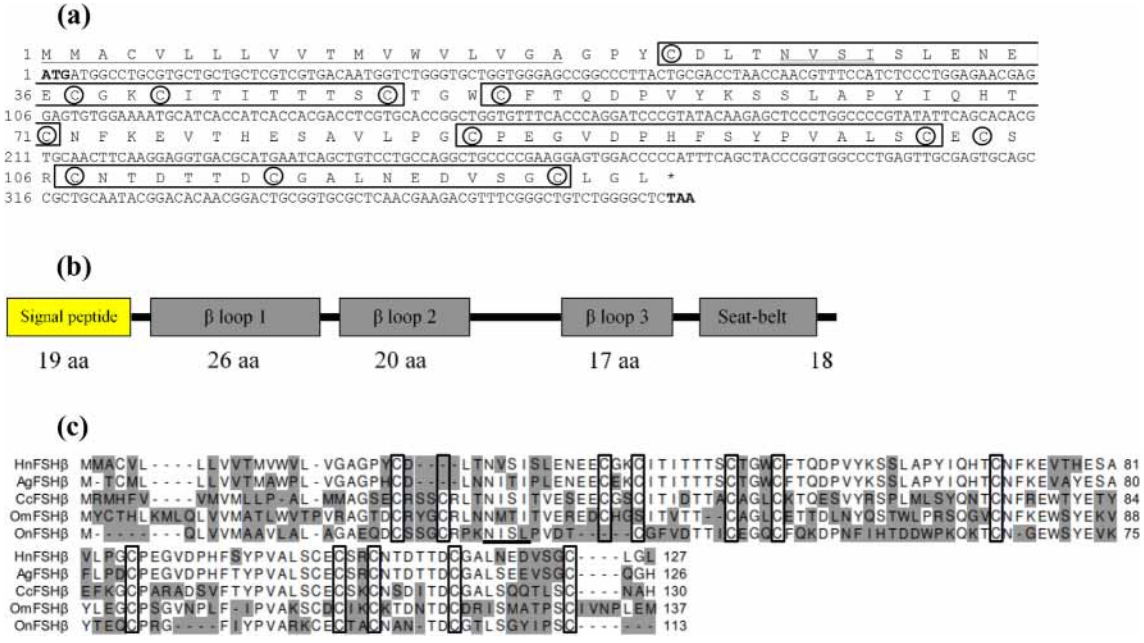

Fig. 13. Nucleotide sequences and deduced protein sequences of (a) the beta subunit of *Heterotis niloticus* follicle-stimulating hormone (HnFSHβ), box modeling of HnFSHβ with the different domains (b), and alignment (c) with *Arapaima gigas* FSHβ (Ag HnFSHβ), *Oncorhynchus mykiss* FSHβ (Om HnFSHβ), *Oreochromis niloticus* FSHβ (On HnFSHβ) and *Cyprinus carpio* FSHβ (Cc HnFSHβ). In figure (a), the start codon and the stop codon are in bold type, the signal peptide is underlined, β-loops 1, 2 and 3 and the "seat-belt" are laid out successively in boxes. The predicted N-glycosylation site is double-underlined. In figure (c), conserved cysteines across species are in boxes, the little conserved domains are in gray.

in β-loop 2, Cys^{7} and Cys^{8} in β-loop 3, Cys^{9} between β-loop 3 and the "seat-belt loop", and Cys^{10} to Cys^{12} in the "seat-belt loop". In Cypriniformes and catfish such as *H. fossilus*, good conservation of the structure of the FSHβ subunit is observed, the only difference is that there is an additional cysteine between the signal peptide and β-loop 1 in *H. fossilus*. Cys^{2} is absent in Perciformes and so is Cys^{3} in Salmonidae, but an additional cysteine residue in the N-terminal position, as in Cypriniformes and catfish, gives 12 cysteines in total. In zebrafish, Cys^{10} and Cys^{11}, thought to be the decisive loop of the FSHβ "seat-belt" (Vischer et al., 2004) are absent, and two additional cysteines exist in the N-terminal position before β-loop 1, allowing for a total number of 12 cysteines (So et al., 2005). Given the importance of the "seat-belt" region for receptor interaction and heterodimer formation, variation in the structure of FSHβ subunits among fish species may result in

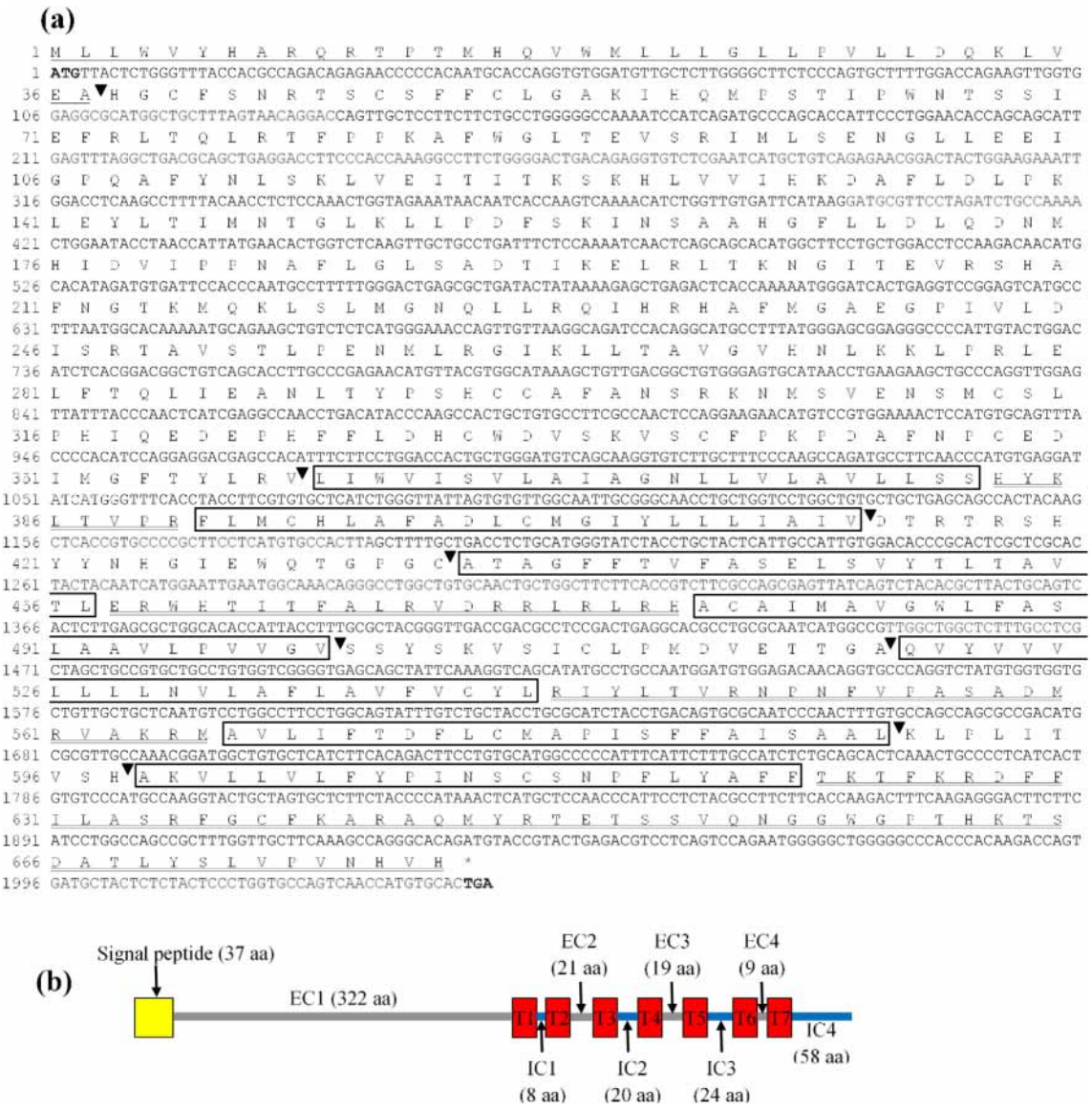

Fig. 14. Nucleotide sequences and deduced protein sequences (a) of *Heterotis niloticus* follicle-stimulating hormone receptor (HnFSHR), and box modeling (b) with the different extracellular domains (EC1 to EC4), transmembrane domains (T1 to T7) and intracellular domains (IC1 to IC4) of HnFSHR. In figure (a), the start codon and the stop codon are in bold type, the signal peptide is underlined, T domains are in boxes, IC domains are double-underlined, and EC domains are delineated by black triangles.

Fig. 14c. Alignment of deduced protein sequences of *Heterotis niloticus* follicle-stimulating hormone receptor (HnFSHR) with FSHRs of *Oncorhynchus mykiss* FSHR (OmFSHR) and *Cyprinus carpio* FSHR (CcFSHR) showing little conserved domains in gray.

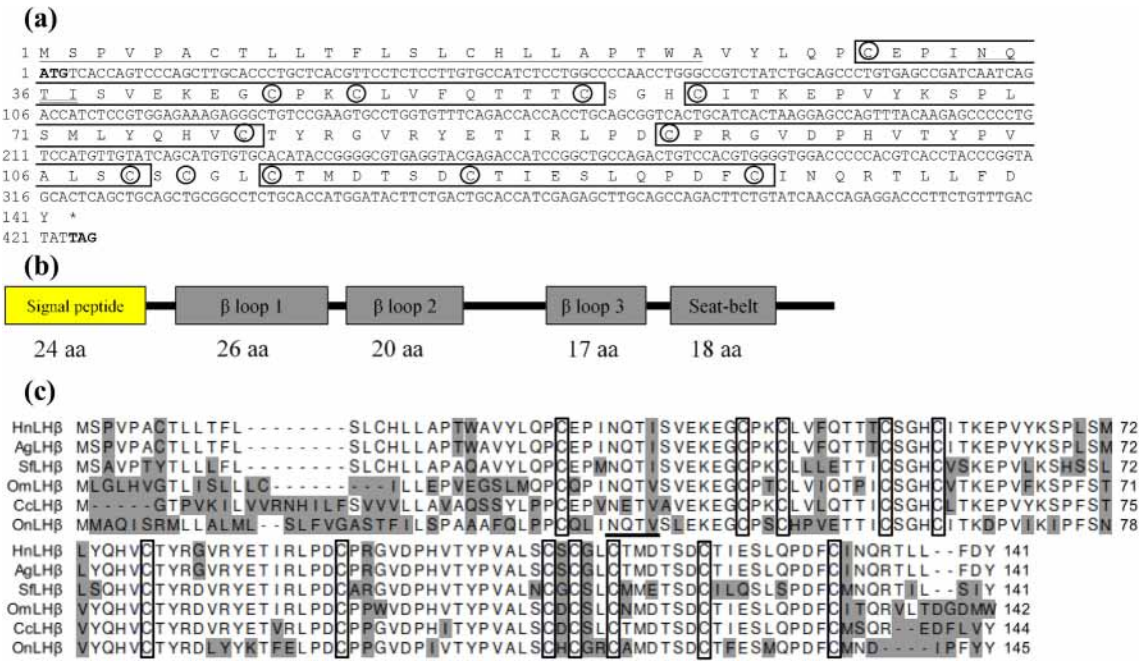

Fig. 15. Nucleotide sequences and deduced protein sequences of (a) the beta subunit of *Heterotis niloticus* luteinizing hormone (HnLHβ), box modeling of HnLHβ with the different domains (b) and alignment (c) with *Arapaima gigas* LHβ (AgLHβ), *Scleropages formosus* LHβ (SfLHβ), *Oncorhynchus mykiss* LHβ (OmLHβ), *Oreochromis niloticus* LHβ (OnLHβ) and *Cyprinus carpio* LHβ (CcLHβ). In figure (a), the start codon and the stop codon are in bold type, the signal peptide is underlined, β-loops 1, 2 and 3 and the "seat-belt" are laid out successively in boxes. The predicted N-glycosylation site is double-underlined. In figure (c), conserved cysteines across species are in boxes, the little conserved domains are in gray.

considerable differences in the nature of the receptor interactions and possibly the stability of the heterodimer (Acharjee et al., 2015; Vischer et al., 2004). Whereas tetrapods, Dipnoi (lungfish), Chondrosteans (sturgeons) and Holocephali (elephant sharks) have two probable sites of N-glycosylation at the level of β-loop 1 of their FSHβ subunit, in *H. niloticus*, only one N-glycosylation site (NVSI) was observed, as in most teleosts. GTH glycosylation is essential for bioactivity, disulfide bond formation, secretion rate, circulatory persistence, and signal transduction (Ulloa-Aguirre et al., 1999). In humans, glycosylation site 1 has been associated to signal transduction, while site 2 is linked to thermal stability, which has a greater effect on disulfide bond formation and hormone secretion (Feng et al., 1995). The loss of the second N-glycosylation of *H. niloticus* FSHβ could therefore affect interactions among receptors (Bousfield and Dias, 2011; Ulloa-Aguirre et al., 1999). Thus, the absence of a second glycosylation site in the FSHβ of most fish appears to strongly diverge from the highly conserved vertebrate model, and raises the hypothesis of the incidence on the half-life of the molecule and in turn on the absence of its detection in certain fish (Acharjee et al., 2015).

4.3. Estradiol receptors

Several authors agree on the existence of three distinct subtypes of 17 β-estradiol receptors (ERs) in fish, with quite varied nomenclatures (Choi and Habibi, 2003; Hawkins et al., 2000; Hawkins and Thomas, 2004; Marlatt et al., 2008; Métivier et al., 2002; Nelson and Habibi, 2013; Nelson and Habibi, 2010; Norris and Lopez, 2011; Sabo-Attwood et al., 2004). These are the alpha (ERα, ESR1, ER, ESR, ESRA, Era or NR3A1), gamma or beta 1 (ERγ, ESR2α, ERβ-I, ERβα, ERβ1) and beta or beta 2 (ERβ, ESR2b, ERβ- II, ERβb, ERβ2 or NR3A2) receptors.

We identified the beta subtype of the estradiol receptor (ERβ) in the liver of *H. niloticus* females in vitellogenesis, and the beta and gamma subtypes (ERβ and ERγ) in the ovary. The three ER subtypes appeared to be differentially distributed in the tissues. The alpha subtype was not present in any tissue, ERγ was expressed in the ovary, and ERβ in the liver and ovary. Investigations into the functional significance of the different ER subtypes in fish and other vertebrates showed that expression patterns differed according to sex, gametogenesis stage, species, and tissue type (Couse et al., 1997; Marlatt et al., 2008; Nelson et al., 2007) and that ER subtypes bound to E2 with different levels of affinity (Genovese et al., 2014; Nelson and Habibi, 2013). For example, in *Micropterus salmoides* (Sabo-Attwood et al., 2004), the three ER subtypes were detected in the liver, ovary and pituitary gland of females in early vitellogenesis, with higher expression levels in the ovary. Of the three *M. salmoides* ER subtypes, ERβ had the highest level of expression in the ovary and liver. ERγ was more expressed in the ovary than ERα, whereas the opposite was observed in the liver. Many studies have been devoted to the study of the regulation of ERs in the liver (Nelson and Habibi, 2013) because Vtg, the precursor of the proteins of the yolk of the egg, is mainly expressed in this tissue (Dominguez et al., 2012; Hara et al., 2016; Prat et al., 1998). The strong expression of ERβ and ERγ in the ovary of certain species suggests that these two subtypes play a greater role in gene regulation in this organ (Sabo-Attwood et al., 2004) through expression and stimulation of Vtg synthesis in some fish (Kong et al., 2014; Wang et al., 2005; Wang et al., 2010).

Two types of estradiol receptor-related receptors (ERRα and ERRγ) were also identified in the liver and ovary of *H. niloticus*. ERRs are orphan receptors for which no endogenous ligand has yet been identified, so that these receptors have not received as much attention as some endocrine receptors (Audet-Walsh and Giguère, 2015; Giguère, 1999). *H. niloticus* ERRs showed significant homology with ERs, with 68–70% amino acid identity at the core of the DBD. However, these ERRs do not bind to estrogens, but rather to hexanucleotide sequences in the form of monomers or to composite elements in the form of dimers (Giguère, 1999). As a result, ERRs do not participate directly in conventional estrogen signaling pathways or biological processes (Audet-Walsh and Giguère, 2015). The two *H. niloticus* ER subtypes contained five domains, as most nuclear receptors do (Aranda and Pascual, 2001; Nelson and Habibi, 2013; Sabo-Attwood et al., 2004). The N-terminal region (A/B) contained the first activation function (AF-1); the strong variation observed between HnERβ and HnERγ (58 to 68%) and between HnERβ_X3 and the other HnERβs contributed to a high specificity of the

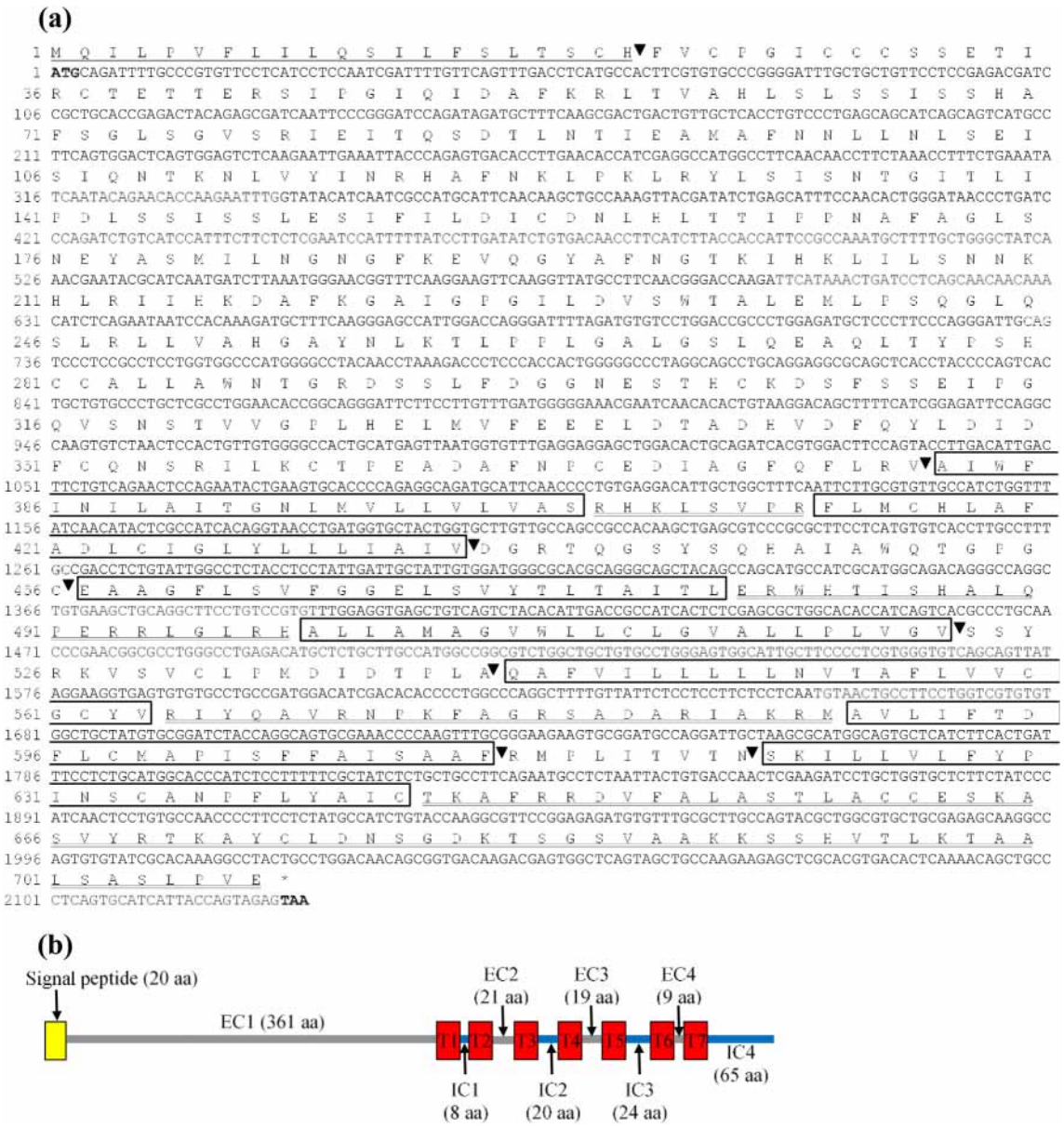

Fig. 16. Nucleotide sequences and deduced protein sequences (a) of *Heterotis niloticus* luteinizing hormone receptor (HnLHR), and box modeling (b) with the different extracellular domains (EC1 to EC4), transmembrane domains (T1 to T7) and intracellular domains (IC1 to IC4) of HnLHR. In figure (a), the start codon and the stop codon are in bold type, the signal peptide is underlined, T domains are in boxes, IC domains are double-underlined, and EC domains are delineated by black triangles.

receptors (Aranda and Pascual, 2001). The core of the DBDs comprised 66 amino acids and was by far the most conserved domain across HnERs (89%) and HnERRs (greater than94%), as in most nuclear receptors of different species (Aagaard et al., 2011). In *M. salmoides* (Sabo-Attwood et al., 2004) and *Micropogonias undulatus* (Hawkins et al., 2000), the DBD nucleus of the β and γ forms of ERs share around 70% identity, and therefore the protein sequence of ERγ is probably derived from the duplication of ERβ genes (Hawkins et al., 2000; Hawkins and Thomas, 2004). In addition to the DBD core, the C-terminal extension (CTE), composed of the first 10–30 residues following the fully conserved VGM sequence of the DBD nucleus, is an important determinant of specific DNA binding for many nuclear receptors (Aagaard et al., 2011). Unlike the DBD nucleus, CTE is more variable among nuclear receptors (Aagaard et al., 2011).

The amino acids considered to be essential for the recognition of estradiol in most vertebrate ERs (Sabo-Attwood et al., 2004) were well preserved at the level of HnERβ_X1, HnERβ_X2, HnERβ_X3 and HnERγ, while HnERβ_X4 possessed only two preserved sites.

The analysis of the predicted amino acid sequences revealed that the four forms of HnERβ had identical domains, and only diverged at the level of the A/B or LBD domains. These were four HnERβ isoforms resulting from alternative splicing of the gene coding for these estrogenic receptors. In humans five isoforms of ERβ have been identified and their divergence is located at the C-terminus of the LBD and F domain (Leung et al., 2006). The human ERβ1 isoform has the longest LBD sequence, containing the complete helices 11 and 12 where the second AF-2 activation function and four amino acids involved in binding with E2 are found; it is the only functional isoform in homodimer. The C-terminal end of the LBD is much shorter in the other isoforms because of the disorientation or absence of helices 11 and 12. As a result, they cannot form functional homodimers but rather heterodimers with the ERβ1 isoform insofar as a single functional helix 12 of the dimer is sufficient (Leung et al., 2006). Under these conditions, HnERβ_X1, HnERβ_X2, HnERβ_X3 and HnERγ, which possess helices 11 and 12 with all the amino acids involved in the binding with E2, can function in homodimers, while HnERβ_X4, which does not have helix

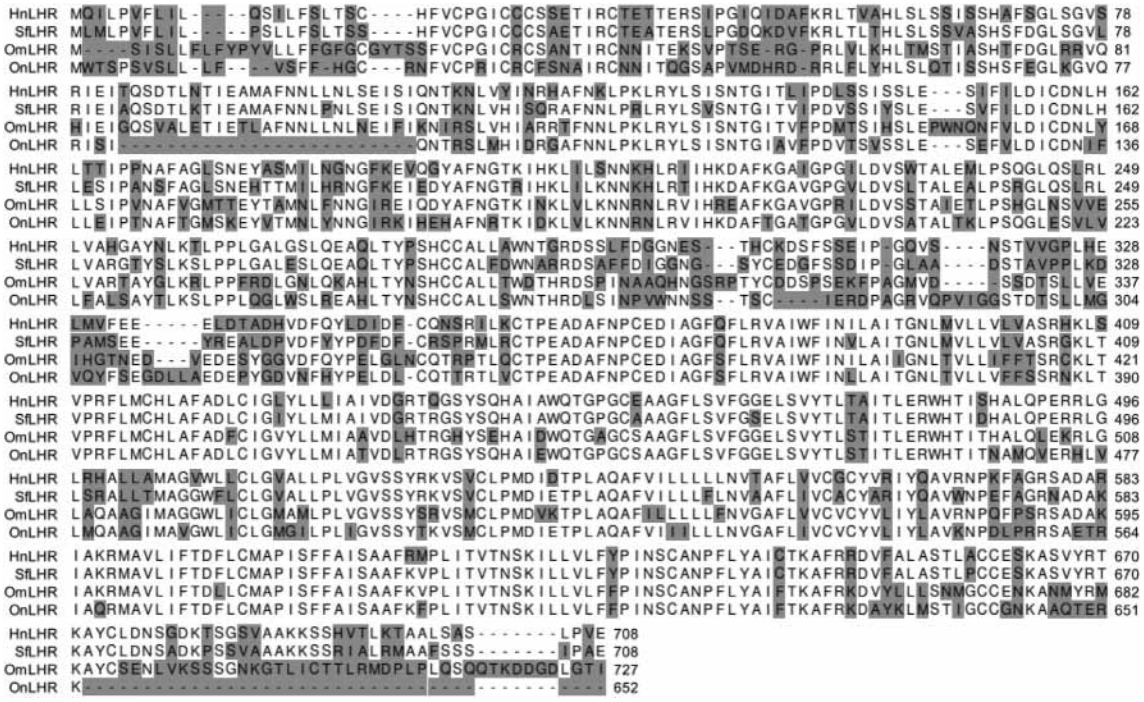

Fig. 16c. Alignment of deduced protein sequences of *Heterotis niloticus* luteinizing hormone receptor (HnLHR) with the LHRs of *Scleropages formosus* (SfLHR), *Oncorhynchus mykiss* (OmLHR), and *Oreochromis niloticus* (OnLHR) showing little conserved domains in gray.

12 and has a short helix 11 lacking the four amino acids involved in the binding with E2, can only form functional heterodimers with the other HnERβs. The D-box involved in receptor dimerization (Aranda and Pascual, 2001) is represented by different peptides, namely PATNQ in HnERβ and PANNQ isoforms in HnERγ, suggesting that these two types of ERs do not form functional heterodimers on DNA (Zechel et al., 1994).

4.4. Vitellogenins and their receptors

Three forms of Vtg identified in many species (Schilling et al., 2015; Williams et al., 2014; Yilmaz et al., 2015) are divided into two categories. One is characterized by the presence of all domains, the other by the absence of one or more domains. In the first category, Vtg are characterized by the presence of a phosvitin domain rich in phosphorylated serines (Pv), a lipovitellin heavy chain (LvI), a lipovitellin light chain (LvII), a β'-C component, and a C-terminal coding region (CT). This category of Vtg has two variants, namely VtgA and VtgB, which have more or less similar molecular weights. In the second category, VtgC lacks the Pv domain (Hara et al., 2016; Hiramatsu et al., 2006; Yilmaz et al., 2015).

We identified four Vtg in *H. niloticus*. Many fish species have several Vtg. For example, three Vtg were identified in *Larimichthys crocea* (Gao et al., 2019) and seven in *Danio rerio* (Wang et al., 2005). The fish genome can contain up to 20 copies of genes encoding vitellogenins, e.g., *Oncorhynchus mykiss* (Trichet et al., 2000). HnVtg1, HnVtg2 and HnVtg4 have a Pv domain which is used to classify them in the Vtg A or Vtg B group, whereas HnVtg3 corresponds to a VtgC devoid of the Pv, β'-C and CT domains. Alignments made with the variants of Vtg A or B of other fish species made it possible to classify HnVtg1 and HnVtg2 among VtgBs. *H. niloticus* hepatic Vtg share high amino acid identity (greater than98%) and appear to be derived from alternative splicing of the gene encoding this protein. In *O. mykiss*, the genes encoding different Vtg show 98.7% similarity (Trichet et al., 2000) whereas *Morse saxatilis* VtgC shares less than 32% identity with the A or B forms (Williams et al., 2014). The identification of Vtg transcripts in the liver and ovary of *H. niloticus* revealed that Vtg are expressed in these two organs in this species. Vtg is mainly expressed in the liver, but its expression has also been observed in the ovaries of certain fish species such as *Danio rerio* (Wang et al., 2005), *Tanichthys albonubes* (Wang et al., 2010) and *Rhodeus uyekii* (Kong et al., 2014). Ovarian Vtg, like most Vtg expressed in non-hepatic organs or tissues, represent less than 10% of the level of hepatic Vtg expression (Wang et al., 2005).

The N-ter1, N-ter2 and C-ter transcripts characteristic of HnVtg1 and HnVtg2 were not sufficient to obtain the definitive sequences of these proteins because the exact number of serines to be retained for the Pv domain was uncertain. The nineteen and twelve serines at the end of the sequence of N-ter1 and N-ter2, respectively, as well as the twenty serines at the beginning of the C-ter sequence all came from the same AGC codon, while serine is coded by five other codons (TCT, TCC, TCA, TCG and AGT). Trinity failed to connect the N-ter and the C-ter. However, alignment with the Vtg of *S. formosus*, another Osteoglossiforme, showed that the number of serines in the Pv domain was uncertain; but this uncertainty was low and probably corresponded to one or two serine residues; knowing that they may be missing or supernumerary serines, the transcripts may be overlapping. This means that apart from a few serines, the HnVtg1 and HnVtg2 sequences were complete. In other words, either some serines were missing to connect the 2 N-ter with the C-ter, or the N-ter and C-ter overlapped. But as previously stated, serines in the Pv domain of *Heterotis* Vtg are encoded by a single codon that makes assembly complex.

MS/MS analysis detected the presence of Vtg in the blood and mucus of *H. niloticus* females in vitellogenesis. Enzyme immunoassays (ELISA) have been used in many species to detect the presence of plasma Vtg (Bulukin et al., 2007; Chu-Koo et al., 2009; Ndiaye et al., 2006; Nuñez-Rodriguez et al., 1989; Nunez Rodriguez et al., 1997; Roy et al., 2004) and mucus Vtg (Arukwe and Røe, 2008; Genovese et al., 2014; Gordon et al., 1984; Meucci and Arukwe, 2005; Moncaut et al., 2003). The proteomic approach allowed us to sex *H. niloticus* adults although the species is devoid of sexual dimorphism. In the absence of direct observation of gonadal tissue, a similar sexing test was performed on blood samples from *Epinephelus lanceolatus* individuals (Om et al., 2013). The fact that we did not detect Vtg in the blood and mucus of *H. niloticus* males indicates that the sampled fish were not subject to

N-ter1 HnVtg1 (Blood)

MRAVLLALTLALAAGQQDSLTPDFANGKAYMFKYEAELLGGLPVEGLGKAGVKIVCKVLISRVSQNTYLLKLKEPQIFEYSGIRPRDDFKPAAKLTQALAAQLLIPV
KFEYIRGVVGKVFAPAAVSKTILNLHRGILNILHVNIKKTQNVYELQEAGSQGVCKTDYAISEDTKAEHIYVTKSKDLGNCQNRIMADIGIAYTETCVQCQMRNKNL
RGAATYSYVMKPTETGALITKATVQEVHQFTPFHELRGAGQVETRQEMTFVETQNDPVQPAHADYVARGSLQYEFASELLQTPLPLMKITDIKTQIEEILNHLVKNN
MGEVHEDAPLKFLQLTQLLREAKYEIMNGIWTQVKSKPVFRRWFLDTVPAVGNQDAVRFIKEKFIAGDISAAETAQALLVALHLLEANMDTVNLAGTLVFHAKLQSH
PMLREIAMLGYGSLVFKLCTKQQNCPAEVIKPVHDIAAEAISKANVEEIALVLKVLGNAGHPASIKPIMKVLPGFGSTASSLTVKVHVDAVMALRHIAKRENHKVQE
VALQLFMNRDLHPEVRMVACVVLFECKPPIGLVAMIASALQNEKSLQVASFAYSHMKALTRSTAPELAQVAAACNVAIKILSPKLDRLSYRFSKTIHMDFFNYHLMA
GAAATAHLINDAASILPRALVAKIRACMLGAAADVLEFGVRTEGLQEALLKSPAADPNADRLTRMWRVLNTLKNWKSLPASQPVASAYLKWFGQEIAFANIDRDIIE
KAVEFATGAAAQPALLKNILNMMQSGIDIQIAKPLMTSEVRRIFPTSMGFPIEISLYSAAVAAAIVKAKAILNPKPSDNFRIAQLMNTDIQLDTRVVPSIAVHKYAV
MGVNTALIQAAIEAKVKVQKVLPLKFNARININAQGHYKIETVPLHAQERILDLHMETVAVARNIENLSEAKITPVLPARLASQQSKETFASAGSGSKSSERIHEQES
SSHPMQHAVSARTDRQWCVTVASLADQACAKVTSQNAGFIRNSPLYKLIGEHSVIAAVKPVSGEAVDKIEIEMHVGPDAASKIVKTITVKDDNAKEGHAGHSPVVLK
LREILETEKKQHSRNATTSSSSSSSSRRSGQKSSSSSSSSSSSSSSSSNSSKRSKKNIKKRSSSSSSSSSSSSSSSRRRSRTENVLGGSSSSSSRSSRSQTSKAAIF
QRFTQNHIHQHETTRAASSQKTSSSPSSSSSSSNSSSSSSSSSSSSSSSSS

N-ter1 HnVtg1 (Mucus)

MRAVLLALTLALAAGQQDSLTPDFANGKAYMFKYEAELLGGLPVEGLGKAGVKIVCKVLISRVSQNTYLLKLKEPQIFEYSGIRPRDDFKPAAKLTQALAAQLLIPV
KFEYIRGVVGKVFAPAAVSKTILNLHRGILNILHVNIKKTQNVYELQEAGSQGVCKTDYAISEDTKAEHIYVTKSKDLGNCQNRIMADIGIAYTETCVQCQMRNKNL
RGAATYSYVMKPTETGALITKATVQEVHQFTPFHELRGAGQVETRQEMTFVETQNDPVQPAHADYVARGSLQYEFASELLQTPLPLMKITDIKTQIEEILNHLVKNN
MGEVHEDAPLKFLQLTQLLREAKYEIMNGIWTQVKSKPVFRRWFLDTVPAVGNQDAVRFIKEKFIAGDISAAETAQALLVALHLLEANMDTVNLAGTLVFHAKLQSH
PMLREIAMLGYGSLVFKLCTKQQNCPAEVIKPVHDIAAEAISKANVEEIALVLKVLGNAGHPASIKPIMKVLPGFGSTASSLTVKVHVDAVMALRHIAKRENHKVQE
VALQLFMNRDLHPEVRMVACVVLFECKPPIGLVAMIASALQNEKSLQVASFAYSHMKALTRSTAPELAQVAAACNVAIKILSPKLDRLSYRFSKTIHMDFFNYHLMA
GAAATAHLINDAASILPRALVAKIRACMLGAAADVLEFGVRTEGLQEALLKSPAADPNADRLTRMWRVLNTLKNWKSLPASQPVASAYLKWFGQEIAFANIDRDIIE
KAVEFATGAAAQPALLKNILNMMQSGIDIQIAKPLMTSEVRRIFPTSMGFPIEISLYSAAVAAAIVKAKAILNPKPSDNFRIAQLMNTDIQLDTRVVPSIAVHKYAV
MGVNTALIQAAIEAKVKVQKVLPLKFNARININAQGHYKIETVPLHAQERILDLHMETVAVARNIENLSEAKITPVLPARLASQQSKETFASAGSGSKSSERIHEQES
SSHPMQHAVSARTDRQWCVTVASLADQACAKVTSQNAGFIRNSPLYKLIGEHSVIAAVKPVSGEAVDKIEIEMHVGPDAASKIVKTITVKDDNAKEGHAGHSPVVLK
LREILETEKKQHSRNATTSSSSSSSSRRSGQKSSSSSSSSSSSSSSSSNSSKRSKKNIKKRSSSSSSSSSSSSSSSRRRSRTENVLGGSSSSSSRSSRSQTSKAAIF
QRFTQNHIHQHETTRAASSQKTSSSPSSSSSSSNSSSSSSSSSSSSSSSSS

C-ter HnVtg1 (Blood)

SSSSSSSSSSSSSSSSSSSSSGSSASSFENIYKKSRFLGDTVRPAAVVIVRAVRGNERRRGYQVAAYMDKANARVQVIISALADSDKWQLCADGIQLSMHKVMAKIRW
GAECQDYRAVIKAETGLLGPHPAAQLKMYWNKTPRALKRYASMIYEYVPGVALLAGFSEGRHRSGERQIKLTMAATSARTISIILRTPWMTLYKLGQVIPIALPIGA
AAARAEVEQTFAGRIHYMFVEATSAKCKLENSTVTTFNNRRYGLQMPRSCYQVVAQDCTSKLKFMVLRKGDERSEESHVIVKIADIDVDLTAEHGNIQVKVNGRVVP
ITQHNYEHPTGTISIKQKGAGISLCAPSHGLHEVYFNKNILTIQVPDWMKGNVCGLCGKADGDVRREFQGPSGHHIEDAVSFAHSWVLAAESCHDAKQCQVKQELFK
LTEPVLLNDQDMKCQSTFPVLSCLPQCSPMKTTPVTVGFHCIPIDSNVNSWSSIRKKQEDIRVTVDAHIMCNCGEECA

C-ter HnVtg1 (Mucus)

SSSSSSSSSSSSSSSSSSSSGSSASSFENIYKKSRFLGDTVRPAAVVIVRAVRGNERRRGYQVAAYMDKANARVQVIISALADSDKWQLCADGIQLSMHKVMAKIRW
GAECQDYRAVIKAETGLLGPHPAAQLKMYWNKTPRALKRYASMIYEYVPGVALLAGFSEGRHRSGERQIKLTMAATSARTISIILRTPWMTLYKLGQVIPIALPIGA
AAARAEVEQTFAGRIHYMFVEATSAKCKLENSTVTTFNNRRYGLQMPRSCYQVVAQDCTSKLKFMVLRKGDERSEESHVIVKIADIDVDLTAEHGNIQVKVNGRVVP
ITQHNYEHPTGTISIKQKGAGISLCAPSHGLHEVYFNKNILTIQVPDWMKGNVCGLCGKADGDVRREFQGPSGHHIEDAVSFAHSWVLAAESCHDAKQCQVKQELFK
LTEPVLLNDQDMKCQSTFPVLSCLPQCSPMKTTPVTVGFHCIPIDSNVNSWSSIRKKQEDIRVTVDAHIMCNCGEECA

(a)

```
HnVtg1 MRAVLLALTLALAAGQQDSLTPDFANGKAYMFKYEAELLGGLPVEGLGKAGVKIVCKVLISRVSQNTYLLKLK 73
HnVtg2 MKAAVFALALALVAGQQNILTPDFAAKKTYVYKYEAQLHSELPEEGLAKAGLKIASKVLISRADQNTYLLKLK 73
       * *   ** *** ****  ******  * *  **** *   ** *** *** **  ******  *********
```

(b)

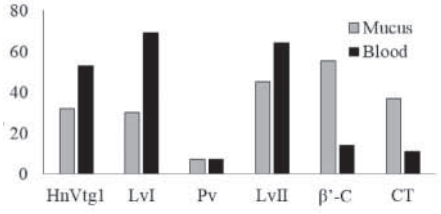

Fig. 17. Coverage of the HnVtg1 N-ter and C-ter peptides detected in *Heterotis niloticus* blood and mucus by mass spectrometry. (a) Peptides specific to HnVtg1 and HnVtg2 detected in the blood. The amino acids conserved between the two Vtg are indicated by stars. Detected peptides are in gray, the signal peptide is underlined. (b) Coverage rate (in percentage) of the HnVtg1 peptides detected in the blood and mucus compared to the whole protein (N-ter1 + C-ter) and the different domains. Only Vtg1 was detected in mucus.

endocrine disruption due to mimetic substances of natural estrogens known as xenoestrogens contained in certain foods distributed to most farmed fish (Pelissero and Sumpter, 1992; Pelissero et al., 1989) or in the aquatic environment (Arukwe and Røe, 2008; Dugué et al., 2008; Genovese et al., 2014; Le Menn, 1979).

Concerning Vtg receptors, two types of Vtg receptor transcripts apparently derived from alternative splicing of the O-glycosylated domain (Chen et al., 2016) were identified in *H. niloticus*. These were LR8 forms with eight LBRs in the LBD (Sappington and Raikhel, 1998) that have long been considered the only VLDLR capable of binding Vtg in some oviparous species (Reading et al., 2014). LR8+, which has an O-glycosylated domain, and LR8-, which does not have one, were detected in the ovary and pituitary gland of *H. niloticus*, respectively. These two forms of LR8 have also been identified in teleost fish such as *O. mykiss* (Prat et al., 1998) and *Oncorhynchus clarki* (Mizuta et al., 2013). However, LR8- is expressed primarily in the ovary of these species,

while LR8 + is expressed in both the ovary and non-ovarian tissues such as the brain, intestine and gills. Detection of LR8- in non-ovarian tissue (*H. niloticus* pituitary gland) requires further investigation. In recent years, a new VLDLR called LRP13 or LR13 + 1 has been identified in several teleost species other than *H. niloticus*, such as *M. saxatilis*, *Morone americana* (Reading et al., 2014), *Cynoglossus semilaevis*, *Oryzias latipes* (Wang et al., 2017) and *O. clarki* (Mushirobira et al., 2015), with thirteen LBRs in the N-terminal LBD and one LBR in the C-terminal LBD (Hiramatsu et al., 2013). Some fish species may have up to four forms of VLDLR, e.g., *M. Americana* (Reading et al., 2011). In the brain, VLDLRs play a role in extracellular signaling, intracellular signaling processes and central nervous system development (Trommsdorff et al., 1999).

LR8 are conserved structurally in many species (Dominguez et al., 2012; A. Li et al., 2003; Pousis et al., 2012; Prat et al., 1998) including *H. niloticus*, which shares more than 84% amino acid identity with species of Osteoglossiformes, Salmoniformes, Perciformes and Cypriniformes. The precursors of HnVLDLR have a signal peptide followed by eight or seven domains for the LR8 + and LR8- forms, respectively. LBD controls interactions between the receptor and the Vtg Lv domain (Hiramatsu et al., 2001). It is composed of eight LBRs of around 40 amino acids each, including six cysteines (Li et al., 2003; Luo et al., 2013) and negatively charged residues on its surface. The EDE signature can be found in LBR3, and the SDE signature in other LBRs well conserved across species (Agulleiro et al., 2007). Three LBRs (LBR1 to LBR3) are primarily involved in the binding of Vtg to the receptor (Li et al., 2003). In chicken (Bajari et al., 1998), the EDE tripeptide of LBR3 has a higher negative charge density than the SDE tripeptide of other LBRs and appears to be determining for binding Vtg to the receptor, whereas a role in calcium cage formation seems more likely in *O. niloticus* (Li et al., 2003). Cysteines from the three precursor domains of the epidermal growth factor (EGF) are also well conserved across species. The domain predicted to form a beta helical structure (Agulleiro et al., 2007) between EGF2 and EGF3 is composed of five 4-aa sequences of the YWS(V)D and FWA(T)D types; these are well conserved across species. Finally, the highly conserved FDNPVY sequence of the cytoplasmic domain, potentially involved in receptor internalization after ligand binding (Agulleiro et al., 2007) is also present in HnVLDLR.

5. Conclusion/outlooks

This study provides molecular data on vitellogenesis for the first time in an African Osteoglossiforme named *Heterotis niloticus*. The nucleotide sequences and protein sequences deduced from prolactin and its receptor, gonadotropins (FSH and LH) and their receptors, beta and gamma receptors to estradiol and vitellogenins and their receptors have been identified. No ERα has been identified, and the definitive sequences of two hepatic Vtg and ovarian Vtg remain to be determined by supplementary analyses. We sexed the species by mass spectrometry detection of Vtg in mucus and blood. This study can be considered as a necessary step in the functional characterization of the identified molecules. It also makes it possible to envisage studies of reproductive behavior from sexed individuals to better control reproduction and breeding of this species.

Because MS and ELISA analyses are quantitative, it is urgent to establish a relationship between the Vtg concentration in mucus and the maturity stage of females (Baumann et al., 2013; Chatakondi and Kelly, 2013) so as to implement multiple applications in fish farming in non-invasive conditions. This will preserve the welfare of farmed fish and thus produce quality, relatively stress-free fish.

On a practical level, knowledge of the primary sequences of pituitary hormones will make it possible to synthesize peptides to induce gametogenesis and control fry production continuously throughout the year (Bry et al., 1978; Mehdi and Ehsan, 2013; Mylonas et al., 2010). This is an important parameter for ensuring regular production of marketable fish.

Moreover, based on knowledge of *H. niloticus* pituitary neurohormones, it will be possible to monitor the circulating levels in this species with specific quantitative tools such as ELISA or mass spectrometry during *in vivo* experiments that remain to be undertaken. In the near future, it will be necessary to characterize the whole neuropeptidome of this species and also the whole regulatory peptidome to facilitate the interpretation of physiological studies and thus totally master breeding. The use of molecular tools could facilitate the selection of successful breeders (Hallerman, 2006) to develop sustainable, environment-friendly fish farming.

6. Thanks

We thank Sabbat Affif, former owner of the fish farm where the samples were collected for this study, and the farm manager Jacques Manlan who organized the various catches. The authors also thank the IRD itrop HPC (Montpellier SouthGreen Platform: https://bioinfo.ird.fr/ - http://www.southgreen.fr) and the ABiMS platform of the Roscoff biological station http://abims.sb-roscoff.fr) for providing the HPC resources that contributed to the search results reported in this document.

Acknowledgements

This work was funded by the IRD, the International Joint Laboratory “Evolution and Domestication of the Amazonian Ichthyofauna” (LMI EDIA) and SupAgro Montpellier (France) within the framework of the Debt Reduction and Development Contract (C2D) between France and Ivory Coast: C2D-ESA Project (Educational and institutional renovation of the École Supérieure d'Agronomie (ESA) of the Institut National Polytechnique Félix HOUPHOUËT-BOIGNY).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygcen.2020.113532.

References

Aagaard, M.M., Siersbæk, R., Mandrup, S., 2011. Molecular basis for gene-speci fi c transactivation by nuclear receptors. BBA 1812, 824–835. https://doi.org/10.1016/j.bbadis.2010.12.018.

Acharjee, A., Chaube, R., Joy, K.P., 2015. Molecular cloning and characterization of the gonadotropin subunits GP a, FSH b, and LH b Genes in the Stinging Cat fi sh Heteropneustes fossilis: phylogeny, seasonal expression, and pituitary localization. J. Exp. Zool. 323A, 567–585. https://doi.org/10.1002/jez.1949.

Adite, A., Winemiller, K.O., Fiogbe, E.D., 2006. Population structure and reproduction of the African bonytongue Heterotis niloticus in the So River-floodplain system (West Africa): implications for management. Ecol. Freshw. Fish 15 (1), 30–39. https://doi.org/10.1111/j.1600-0633.2005.00119.x.

Agulleiro, M.J., André, M., Morais, S., Cerdà, J., Babin, P.J., 2007. High transcript level of fatty acid-binding protein 11 but not of very low-density lipoprotein receptor is correlated to ovarian follicle atresia in a teleost fish (Solea senegalensis)1. Biol. Reprod. 77 (3), 504–516. https://doi.org/10.1095/biolreprod.107.061598.

Aranda, A., Pascual, A., 2001. Nuclear hormone receptors and gene expression. Physiol. Rev. 81 (3), 1269–1304 0031-9333/01.

Arukwe, A., Røe, K., 2008. Molecular and cellular detection of expression of vitellogenin and zona radiata protein in liver and skin of juvenile salmon (Salmo salar) exposed to nonylphenol. Cell Tissue Res. 331 (3), 701–712. https://doi.org/10.1007/s00441-007-0543-y.

Audet-Walsh, É., Giguère, V., 2015. The multiple universes of estrogen-related receptor α and γ in metabolic control and related diseases. Acta Pharmacol. Sin. 36, 51–61. https://doi.org/10.1038/aps.2014.121.

Bajari, T.M., Lindstedt, K.A., Riepl, M., Mirsky, V.M., Nimpf, J., Wolfbeis, O.S., et al., 1998. A minimal binding domain of the low density lipoprotein receptor family. Biol. Chem. 379 (8–9), 1053–1062.

Baumann, L., Holbech, H., Keiter, S., Kinnberg, K.L., Knörr, S., Nagel, T., Braunbeck, T., 2013. The maturity index as a tool to facilitate the interpretation of changes in vitellogenin production and sex ratio in the Fish Sexual Development Test. Aquat. Toxicol. 128–129, 34–42. https://doi.org/10.1016/j.aquatox.2012.11.016.

Betancur-R, R., Wiley, E.O., Arratia, G., Acero, A., Bailly, N., Miya, M., et al., 2017. Phylogenetic classification of bony fishes. BMC Evol. Biol. 17 (1), 162. https://doi.org/10.1186/s12862-017-0958-3.

Bhanu, S.V., Philip, B., 2011. Effect of ethanol on branchial adenosine triphosphatases in Oreochromis mossambicus (Peters). Toxicol. Int. 18 (1), 27. https://doi.org/10.

4103/0971-6580.75849.
Borella, M.I., Venturieri, A.R., Mancera, A.J.M., 2009. Immunocytochemical identification of adenohypophyseal cells in the pirarucu (Arapaima gigas), an Amazonian basal teleost. Fish Physiol. Biochem. 35, 3–16. https://doi.org/10.1007/s10695-008-9254-x.
Bousfield, G.R., Dias, J.A., 2011. Synthesis and secretion of gonadotropins including structure-function correlates. Rev. Endocrine Metab. Disorders 12, 289–302. https://doi.org/10.1007/s11154-011-9191-3.
Bray, N.L., Pimentel, H., Melsted, P., Pachter, L., 2016. Near-optimal RNA-Seq quantification. Nat. Biotechnol. 34 (5), 525–527.
Brooks, C.L., 2012. Molecular mechanisms of prolactin and its receptor. Endocr. Rev. 33 (4), 504–525. https://doi.org/10.1210/er.2011-1040.
Bry, C., Billard, R., De Montalembert, G., 1978. Induction de la maturation ovocytaire et de l'ovulation par traitement hormonal chez le brochet (Esox lucius). Bulletin Français de Pisciculture 271, 21–32. https://doi.org/10.1051/kmae:1978003.
Bryant, D.M., Johnson, K., Di Tommaso, T., Tickle, T., Couger, B.M., Payzin-Dogru, D., et al., 2017. A tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. Cell Rep. 18 (3), 762–776. https://doi.org/10.1016/j.celrep.2016.12.063.A.
Bu, G., Liang, X., Li, J., Wang, Y., 2015. General and comparative endocrinology extrapituitary prolactin (PRL) and prolactin-like protein (PRL-L) in chickens and zebrafish. Gen. Comp. Endocrinol. 220, 143–153. https://doi.org/10.1016/j.ygcen.2015.02.001.
Bulukin, E., Meucci, V., Minunni, M., Pretti, C., Intorre, L., Soldani, G., Mascini, M., 2007. An optical immunosensor for rapid vitellogenin detection in plasma from carp (Cyprinus carpio). Talanta 72 (2), 785–790. https://doi.org/10.1016/j.talanta.2006.12.007.
Carreiro, C.R.P., Furtado-Neto, M.A.D.A., Mesquita, P.E.C., Bezerra, T.A., 2011. Sex determination in the Giant fish of Amazon Basin, Arapaima gigas (Osteoglossiformes, Arapaimatidae), using laparoscopy. Acta Amazonica 41 (3), 415–419. https://doi.org/10.1590/S0044-59672011000300012.
Chatakondi, N.G., Kelly, A.M., 2013. Oocyte diameter and plasma vitellogenin as predictive factors to identify potential channel catfish, ictalurus punctatus, suitable for induced spawning. J. World Aquacult Soc. 44 (1), 115–123. https://doi.org/10.1111/jwas.12001.
Chen, Q., Takahashi, Y., Oka, K., Ma, J., 2016. Functional differences of very-low-density lipoprotein receptor splice variants in regulating wnt signaling. Mol. Cell. Biol. 36 (20), 2645–2654. https://doi.org/10.1128/MCB.00235-16.Address.
Choi, C.Y., Habibi, H.R., 2003. Molecular cloning of estrogen receptor alpha and expression pattern of estrogen receptor subtypes in male and female goldfish. Mol. Cell. Endocrinol. 204 (1–2), 169–177 S030372070200182X [pii].
Chu-Koo, F., Dugué, R., Alván Aguilar, M., Casanova Daza, A., Alcántara Bocanegra, F., Chávez Veintemilla, C., et al., 2009. Gender determination in the Paiche or Pirarucu (Arapaima gigas) using plasma vitellogenin, 17β-estradiol, and 11-ketotestosterone levels. Fish Physiol. Biochem. 35 (1), 125–136. https://doi.org/10.1007/s10695-008-9211-8.
Couse, J.F., Lindzey, J., Grandien, K.A.J., Gustafsson, J.-åke, Korach, K.S., 1997. Tissue distribution and quantitative analysis of estrogen receptor-α (ERα) and estrogen receptor-β (ERβ) messenger ribonucleic acid in the wild-type and ERα-knockout mouse. Endocrinology 138 (11), 4613–4621.
Devlin, R.H., Nagahama, Y., 2002. Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. Aquaculture 208 (3–4), 191–364. https://doi.org/10.1016/S0044-8486(02)00057-1.
Dominguez, G.A., Bisesi, J.H., Kroll, K.J., Denslow, N.D., Sabo-Attwood, T., 2014. Control of transcriptional repression of the vitellogenin receptor gene in largemouth bass (micropterus salmoides) by select estrogen receptors isotypes. Toxicol. Sci. 141 (2), 423–431. https://doi.org/10.1093/toxsci/kfu145.
Dominguez, G.A., Quattro, J.M., Denslow, N.D., Kroll, K.J., Prucha, M.S., Porak, W.F., et al., 2012. Identification and transcriptional modulation of the largemouth bass, micropterus salmoides, vitellogenin receptor during oocyte development by insulin and sex Steroids1. Biol. Reprod. 87 (3), 1–12. https://doi.org/10.1095/biolreprod.112.099812.
Dugué, R., Chu Koo, F., Alcántara Bocanegra, F., Duponchelle, F., Renno, J.F., Nuñez, J., 2008. Purification and assay of Arapaima gigas vitellogenin: potential use for sex determination by. Cybium 32 (2), 111.
Ezekiel, E.N., Abowei, J.F.N., 2013. Length-Weight Relationship and Condition Factor of Heterotis niloticus from Amassoma flood plain, Niger Delta, Nigeria. Appl. Sci. Rep. 4 (1), 164–172.
Faria, M.T., Carvalho, R.F., Sevilhano, T.C.A., Oliveira, N.A.J., Silva, C.F.P., Oliveira, J.E., et al., 2013. Isolation of the pituitary gonadotrophic α-subunit hormone of the giant amazonian fish: pirarucu (Arapaima gigas). Fish Physiol. Biochem. 39 (3), 683–693. https://doi.org/10.1007/s10695-012-9730-1.
Feng, W., Matzuk, M.M., Mountjoy, K., Bedows, E., Ruddon, R.W., Boime, I., 1995. Ehe asparagine-linked oligosaccharides of the human chorionic gonadotropin β subunit facilitate correct disulfide bond pairing. J. Biol. Chem. 270 (20), 11851–11859.
Freeman, M.E., Kanyicska, B., Lerant, A., Nagy, G., 2000. Prolactin: structure, function, and regulation of secretion. Physiol. Rev. 80 (4), 1523–1631. https://doi.org/10.1016/B978-012515400-0/50037-3.
Gao, X.M., Zhou, Y., Zhang, D.D., Hou, C.C., Zhu, J.Q., 2019. Multiple vitellogenin genes (vtgs) in large yellow croaker (Larimichthys crocea): molecular characterization and expression pattern analysis during ovarian development. Fish Physiol. Biochem. 45, 829–848. https://doi.org/10.1007/s10695-018-0569-y.
Genovese, G., Regueira, M., Da Cuña, R.H., Ferreira, M.F., Varela, M.L., Lo Nostro, F.L., 2014. Nonmonotonic response of vitellogenin and estrogen receptor α gene expression after octylphenol exposure of Cichlasoma dimerus (Perciformes, Cichlidae). Aquat. Toxicol. 156, 30–40. https://doi.org/10.1016/j.aquatox.2014.07.019.
Giguère, V., 1999. Orphan nuclear receptors : from gene to function. Endocr. Rev. 20 (5), 689–725.
Goffin, V., Shiverick, K.T., Kelly, P.A., Martial, J.A., 1996. Sequence-function relationships within the expanding. Endocr. Rev. 17 (4), 385–410.
Gordon, M.R., Owen, T.G., Ternan, T.A., Hildebrand, L.D., 1984. Measurement of a sex-specific protein in skin mucus of premature coho salmon (Oncorhynchus kisutch). Aquaculture 43 (1–3), 333–339. https://doi.org/10.1016/0044-8486(84)90034-6.
Grabherr, M., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., et al., 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat. Biotechnol. 29 (7), 644–652. https://doi.org/10.1038/nbt.1883.Trinity.
Guo-Qing, L., Wilson, M.V.H., 1996. The discovery of heterotidinae (Teleostei: Osteoglossidae) from the paleocene paskapoo formation of Alberta, Canada. J. Vertebrate Paleontol. 16 (2), 198–209. https://doi.org/10.1080/02724634.1996.10011308.
Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., et al., 2013. De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. In: Nature Protocols, vol. 8. https://doi.org/10.1038/nprot.2013.084.De.
Hallerman, E.M., 2006. Use of molecular tools for research and improvement of aquaculture stocks. Aquaculture 58 (4), 286–296.
Hara, A., Hiramatsu, N., Fujita, T., 2016. Vitellogenesis and choriogenesis in fishes. Fish. Sci. 82 (2), 187–202. https://doi.org/10.1007/s12562-015-0957-5.
Hawkins, M.B., Thomas, P., 2004. The unusual binding properties of the third distinct teleost estrogen receptor subtype ERβa are accompanied by highly conserved amino acid changes in the ligand binding domain. Endocrinology 145 (6), 2968–2977. https://doi.org/10.1210/en.2003-0806.
Hawkins, M.B., Thornton, J.W., Crews, D., Skipper, J.K., Dotte, A., Thomas, P., 2000. Identification of a third distinct estrogen receptor and reclassification of estrogen receptors in teleosts. Proc. Natl. Acad. Sci. 97 (20), 10751–10756. https://doi.org/10.1073/pnas.97.20.10751.
Hilton, E.J., 2001. Tongue bite apparatus of osteoglossomorph fishes: variation of a character complex. Copeia 2001 (2), 372–381.
Hilton, E.J., 2003. Comparative osteology and phylogenetic systematics of fossil and living bony-tongue fishes (Actinopterygii, Teleostei, Osteoglossomorpha). Zool. J. Linn. Soc. 137 (1), 1–100. https://doi.org/10.1046/j.1096-3642.2003.00032.x.
Hiramatsu, N., Fukadal, H., Sullivan, C.V., Hara, A., 2001. Simple and sensitive detection of vitellogenin receptor (s) in Sakhalin Taimen (Hucho perryi). Bull. Fish Sci. Hokkaido Univ. 52 (1), 5–9.
Hiramatsu, N., Luo, W., Reading, B.J., Sullivan, C.V., Mizuta, H., Ryu, Y.-W., et al., 2013. Multiple ovarian lipoprotein receptors in teleosts. Fish Physiol. Biochem. 39, 29–32. https://doi.org/10.1007/s10695-012-9612-6.
Hiramatsu, N., Matsubara, T., Fujita, T., Sullivan, C.V., Hara, A., 2006. Multiple piscine vitellogenins: biomarkers of fish exposure to estrogenic endocrine disruptors in aquatic environments. Mar. Biol. 149 (1), 35–47. https://doi.org/10.1007/s00227-005-0214-z.
Huang, X., Hui, M.N.Y., Liu, Y., Yuen, D.S.H., Zhang, Y., Chan, W.Y., et al., 2009. Discovery of a novel prolactin in non-mammalian vertebrates : evolutionary perspectives and its involvement in teleost retina development. PLoS ONE 4 (7), e6163. https://doi.org/10.1371/journal.pone.0006163.
Kato, S., Endoh, H., Masuhiro, Y., Kitamoto, T., Uchiyama, S., Sasaki, H., et al., 1995. Activation of the estrogen receptor through ~hosphorylation by Mi togen-activated protein kinase. Science 270, 1491–1494. https://doi.org/10.16377/j.cnki.issn1007-7731.2011.17.025.
Kong, H.J., Kim, J.L., Moon, J.Y., Kim, W.J., Kim, H.S., Park, J.Y., et al., 2014. Characterization, expression profile, and promoter analysis of the Rhodeus uyekii vitellogenin Ao1 gene. Int. J. Mol. Sci. 15 (10), 18804–18818. https://doi.org/10.3390/ijms151018804.
Kopylova, E., Noé, L., Touzet, H., 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28 (24), 3211–3217. https://doi.org/10.1093/bioinformatics/bts611.
Krogh, A., Larsson, B., Von Heijne, G., Sonnhammer, E.L.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. 305 (3), 567–580. https://doi.org/10.1006/jmbi.2000.4315.
Kumar, R., Zakharov, M.N., Khan, S.H., Miki, R., Jang, H., Toraldo, G., et al., 2011. The dynamic structure of the estrogen receptor. J. Amino Acids 2011, 1–7. https://doi.org/10.4061/2011/812540.
Kumazawa, Y., Nishida, M., 2000. Molecular phylogeny of osteoglossoids : a new model for gondwanian origin and plate tectonic transportation of the Asian Arowana. Mol. Biol. Evol. 17 (12), 1869–1878.
Lavoué, S., 2016. Was Gondwanan breakup the cause of the intercontinental distribution of Osteoglossiformes? A time-calibrated phylogenetic test combining molecular, morphological, and paleontological evidence. Mol. Phylogenet. Evol. 99, 34–43. https://doi.org/10.1016/j.ympev.2016.03.008.
Lavoué, S., Sullivan, J.P., 2004. Simultaneous analysis of five molecular markers provides a well-supported phylogenetic hypothesis for the living bony-tongue fishes (Osteoglossomorpha: Teleostei). Mol. Phylogenet. Evol. 33 (1), 171–185. https://doi.org/10.1016/j.ympev.2004.04.021.
Lazard, J., 1990. Transferts de poissons et développement de la production piscicole. Revue Hydrobiologie Tropicale 23 (3), 251–265.
Le Menn, F., 1979. Induction of vitellogenin by estradiol and androgens in a teleostean fish: Gobius niger L. Retrieved from. Comptes Rendus Des Seances de l'Academie Des Sciences 289 (4), 413–416. http://www.ncbi.nlm.nih.gov/pubmed/117947.
Leung, Y.-K., Mak, P., Hassan, S., Ho, S., 2006. Estrogen receptor (ER) beta isoforms : a key to understanding ER beta signaling. PNAS 103 (35), 13162–13167.
Li, A., Sadasivam, M., Ding, J.L., 2003. 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 (5), 2799–2806. https://doi.org/10.1074/jbc.M205067200.
Luo, W., Ito, Y., Mizuta, H., Massaki, K., Hiramatsu, N., Todo, T., et al., 2013. 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 (2), 263–271. https://doi.org/10.1016/j.cbpa.2013.06.026.
Mañanós, E., Núñez-Rodríguez, J., Le Menn, F., Zanuy, S., Carrillo, M., 2007. Identification of vitellogenin receptors in the ovary of a teleost fish, the Mediterranean sea bass (Dicentrarchus labrax). Reprod. Nutr. Dev. 37, 51–61. https://doi.org/10.1051/rnd:19970106.
Manzon, L.A., 2002. The role of prolactin in fish osmoregulation : a review. Gen. Comp. Endocrinol. 125, 291–310. https://doi.org/10.1006/gcen.2001.7746.
Marcos, P.L., Adalberto, L.V., 2015. Differentially expressed genes in the pituitary of the Amazonian fish Arapaima gigas. Int. J. Fisheries Aquacult. 7 (8), 132–141. https://doi.org/10.5897/IJFA15.0473.
Marlatt, V.L., Martyniuk, C.J., Zhang, D., Xiong, H., Watt, J., Xia, X., et al., 2008. Auto-regulation of estrogen receptor subtypes and gene expression profiling of 17β-estradiol action in the neuroendocrine axis of male goldfish. Mol. Cell. Endocrinol. 283 (1–2), 38–48. https://doi.org/10.1016/j.mce.2007.10.013.
Maruska, K.P., Fernald, R.D., 2011. Social regulation of gene expression in the hypothalamic-pituitary-gonadal axis. Physiology 26 (6), 412–423. https://doi.org/10.1152/physiol.00032.2011.
Mehdi, Y., Ehsan, S., 2013. A review of the control of reproduction and hormonal manipulations in finfish species. Int. J. Agricult. Res. 1 (1), 15–21.
Meier, F., Beck, S., Grassl, N., Lubeck, M., Park, M.A., Raether, O., Mann, M., 2015. Parallel accumulation-serial fragmentation (PASEF): multiplying sequencing speed and sensitivity by synchronized scans in a trapped ion mobility device. J. Proteome Res. 14 (12), 5378–5387. https://doi.org/10.1021/acs.jproteome.5b00932.
Métivier, R., Stark, A., Flouriot, G., Hübner, M.R., Brand, H., Penot, G., et al., 2002. A dynamic structural model for estrogen receptor alpha activation by ligands, emphasizing the role of interactions between distant A and E domaines. Mol. Cell 10, 1019–1032.
Meucci, V., Arukwe, A., 2005. Detection of vitellogenin and zona radiata protein expressions in surface mucus of immature juvenile Atlantic salmon (Salmo salar) exposed to waterborne nonylphenol. Aquatic Toxicol. (Amsterdam, Netherlands) 73 (1), 1–10. https://doi.org/10.1016/j.aquatox.2005.03.021.
Mizuta, H., Luo, W., Ito, Y., Mushirobira, Y., Todo, T., Hara, A., et al., 2013. Ovarian expression and localization of a vitellogenin receptor with eight ligand binding repeats in the cutthroat trout (Oncorhynchus clarki). Comp. Biochem. Physiol. B: Biochem. Mol. Biol. 166 (1), 81–90. https://doi.org/10.1016/j.cbpb.2013.07.005.
Moncaut, N., Nostro, F. Lo, Maggese, M.C., 2003. Vitellogenin detection in surface mucus of the South American cichlid fish Cichlasoma dimerus (Heckel, 1840) induced by estradiol-17beta. Effects on liver and gonads. Aquatic Toxicol. (Amsterdam, Netherlands) 63 (2), 127–137. https://doi.org/10.1016/S0166-445X(02)00175-3.
Monentcham, S.E., Kouam, J., Pouomogne, V., Kestemont, P., 2009. Biology and prospect for aquaculture of African bonytongue, Heterotis niloticus (Cuvier, 1829): a review. Aquaculture 289 (3–4), 191–198. https://doi.org/10.1016/j.aquaculture.2009.01.019.
Moreau, J., 1982. Exposé synoptique des données biologiques sur Heterotis niloticus (Cuvier, 1829). FAO Synop. Pêches 131, 1–45.
Mushirobira, Y., Mizuta, H., Luo, W., Todo, T., Hara, A., Reading, B.J., et al., 2015. 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 (12), 986–1000. https://doi.org/10.1002/mrd.22579.
Mylonas, C.C., Fostier, A., Zanuy, S., 2010. Reproduction and broodstock management general and comparative endocrinology Broodstock management and hormonal manipulations of fish reproduction. Gen. Comp. Endocrinol. 165, 516–534. https://doi.org/10.1002/9781444392210.ch4.
Ndiaye, P., Forgue, J., Lamothe, V., Cauty, C., Tacon, P., Lafon, P., et al., 2006. Tilapia (Oreochromis niloticus) vitellogenins: development of homologous and heterologous ELISAs and analysis of vitellogenin pathway through the ovarian follicle. J. Experiment. Zool. Part A: Comp. Experiment. Biol. 305A (7), 576–593. https://doi.org/10.1002/jez.a.290.
Nelson, E.R., Habibi, H.R., 2010. Functional significance of nuclear estrogen receptor subtypes in the liver of goldfish. Endocrinology 151 (4), 1668–1676. https://doi.org/10.1210/en.2009-1447.
Nelson, E.R., Habibi, H.R., 2013. Estrogen receptor function and regulation in fish and other vertebrates. Gen. Comp. Endocrinol. 192, 15–24. https://doi.org/10.1016/j.ygcen.2013.03.032.
Nelson, E.R., Wiehler, W.B., Cole, W.C., Habibi, H.R., 2007. Homologous regulation of estrogen receptor subtypes in goldfish (Carassius auratus). Mol. Reprod. Dev. 74 (1), 1105–1112. https://doi.org/10.1002/mrd.
Nelson, G.J., 1968. Gill arches of teleostean fishes of the division Osteoglossomorpha. J. Linn. Soc. London, Zool. 47 (312), 261–277. https://doi.org/10.1111/j.1096-3642.1968.tb00511.x.
Nelson, G.J., 1969. Infraorbital bones and their bearing on the phylogeny and geography of Osteoglossomorph fishc. Am. Mus. Novit. 2394, 1–37.
Norris, D.O., Lopez, K.H., 2011. Hormones and reproduction of vertebrates, Volume 1 - Fishes. Saudi Med. J. 1.
Noso, T., Nicoll, C.S., Polenov, A.L., Kawauchi, H., 1993. The primary structure of sturgeon prolactin: phylogenetic implication. Gen. Comp. Endocrinol. 91, 90–95. https://doi.org/10.1006/gcen.1993.1108.
Núñez-Rodríguez, J., Bon, E., Le Menn, F., 1996. Vitellogenin receptors during Vitellogenesis in the rainbow troutOncorhynchus mykiss. J. Experiment. Zool. 274, 163–170. https://doi.org/10.1002/(sici)1097-010x(19960215)274:3<163::aid-jez3>3.0.co;2-m.
Nuñez-Rodriguez, J., Kah, O., Geffard, M., Le Menn, F., 1989. Enzyme linked immunosorbent assay (ELISA) for sole (Solea vulgaris)vitellogenin. Comp. Biochem. Physiol. 92B (4), 741–746.
Nuñez, J., 2008. Arapaima gigas se dévoile. Science Au Sud 47, 1.
Nunez Rodriguez, J., Dugue, R., Oteme, Z.J., Hem, S., Le Menn, F., 1997. Vitellogenin plasma levels in two cultured African catfish species, Chrysichthys nigrodigitatus (Claroteidae) and Heterobranchus longifilis (Clariidae). Aquat. Living Resourc. 10, 231–238.
Odo, G., Nwamba, H., Eyo, J., 2009. Aspects of the biology of Heterotis niloticus Cuvier 1829 (Osteoglossiformes: Osteoglossidae) in the anambra flood river system, Nigeria. Anim. Res. Int. 6 (2), 994–1002. https://doi.org/10.4314/ari.v6i2.48131.
Oladosu, G.A., Obi, A., Oladosu, O.O., 2007. Sex determination in Heterotis niloticus (Cuvier 1829) Based on morphometric features. ASSET Ser. B 6 (1), 22–30.
Om, A.D., Jasmani, S., Ismail, N., Yeong, S.Y., Abol-Munafi, A.B., 2013. Application MALDI TOF on protein identification of vitellogenin in giant grouper (Epinephelus lanceolatus). Fish Physiol. Biochem. 39 (5), 1277–1286. https://doi.org/10.1007/s10695-013-9782-x.
Pelissero, C., Cuisset, B., Le Menn, F., 1989. The influence of sex steroids in commercial meals and fish diets on plasma concentration of estrogens and vitellogenin in cultured Siberian sturgeon Acipense baeri. Aquat. Living Resour. 2, 161–168. https://doi.org/10.1051/alr.
Pelissero, C., Sumpter, J.P., 1992. Steroids and “steroid-like” substances in fish diets. Aquaculture 107, 283–301. https://doi.org/10.1016/0044-8486(92)90078-Y.
Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8 (10), 785–786. https://doi.org/10.1038/nmeth.1701.
Plant, T.M., 2015. The hypothalamo-pituitary-gonadal axis. J. Endocrinol. 226 (2), T41–T54. https://doi.org/10.1530/JOE-15-0113.
Pousis, C., Santamaria, N., Zupa, R., De Giorgi, C., Mylonas, C.C., Bridges, C.R., et al., 2012. Expression of vitellogenin receptor gene in the ovary of wild and captive Atlantic bluefin tuna (Thunnus thynnus). Anim. Reprod. Sci. 132, 101–110. https://doi.org/10.1016/j.anireprosci.2012.03.014.
Prat, F., Coward, K., Sumpter, J.P., Tyler, C.R., 1998. Molecular characterization and expression of two ovarian lipoprotein receptors in the Rainbow Trout, Oncorhynchus mykiss 1. Biol. Reprod. 58 (5), 1146–1153. https://doi.org/10.1095/biolreprod58.5.1146.
Reading, B.J., Hiramatsu, N., Schilling, J., Molloy, K.T., Glassbrook, N., Mizuta, H., et al., 2014. Lrp13 is a novel vertebrate lipoprotein receptor that binds vitellogenins in teleost fishes. J. Lipid Res. 55 (11), 2287–2295. https://doi.org/10.1194/jlr.M050286.
Reading, B.J., Hiramatsu, N., Sullivan, C.V., 2011. Disparate binding of three types of vitellogenin to multiple forms of vitellogenin receptor in white Perch 1. Biol. Reprod. 84, 392–399. https://doi.org/10.1095/biolreprod.110.087981.
Roy, R.L., Morin, Y., Courtenay, S.C., Robichaud, P., 2004. Purification of vitellogenin from smooth flounder (Pleuronectes putnami) and measurement in plasma by homologous ELISA. Comp. Biochem. Physiol. - B Biochem. Mol. Biol. 139 (2), 235–244. https://doi.org/10.1016/j.cbpc.2004.07.006.
Sabo-Attwood, T., Kroll, K.J., Denslow, N.D., 2004. Differential expression of largemouth bass (Micropterus salmoides) estrogen receptor isotypes alpha, beta, and gamma by estradiol. Mol. Cell. Endocrinol. 218 (1–2), 107–118. https://doi.org/10.1016/j.mce.2003.12.007.
Sappington, T.W., Raikhel, A.S., 1998. Ligand-binding domains in vitellogenin receptors and other LDL-receptor family members share a common ancestral ordering of cysteine-rich repeats. J. Mol. Evol. 46 (4), 476–487. https://doi.org/10.1007/PL00006328.
Schilling, J., Loziuk, P.L., Muddiman, D.C., Daniels, H.V., Reading, B.J., 2015. Mechanisms of egg yolk formation and implications on early life history of white perch (Morone americana). PLoS ONE 10 (11), 1–20. https://doi.org/10.1371/journal.pone.0143225.
Sevilhano, T., De Carvalho, R.F., Oliveira, N.A.D.J., Oliveira, J.E., Maltarollo, V.G., Trossini, G., et al., 2017. Molecular cloning and characterization of pirarucu (Arapaima gigas) follicle-stimulating hormone and luteinizing hormone β -subunit cDNAs. PLoS ONE 12 (8), e0183545.
Skorupski, J., Kmieć, M., 2013. Structural and functional characteristics of prolactin and its gene in the American mink (Neovison vison Schreb., 1777). A rewiev / Charakterystyka strukturalna i funkcjonalna prolaktyny i jej genu u norki amerykańskiej (Neovison vison schreb., 1777), Pra. Annales UMCS, Zootechnica 30 (4). https://doi.org/10.2478/v10083-012-0035-8.
Šlechtová, T., Gilar, M., Kalíková, K., Tesařová, E., 2015. Insight into trypsin miscleavage: comparison of kinetic constants of problematic peptide sequences. Anal. Chem. 87 (15), 7636–7643. https://doi.org/10.1021/acs.analchem.5b00866.
So, W., Kwok, H., Ge, W., 2005. Zebrafish gonadotropins and their receptors : II. Cloning and characterization of zebrafish follicle-stimulating hormone and luteinizing hormone subunits — their spatial-temporal expression patterns and receptor specificity. Biol. Reprod. 72, 1382–1396. https://doi.org/10.1095/biolreprod.104.038216.
Stifani, S., Le Menn, F., Nunez Rodriguez, J., Schneider, W.J., 1990. Regulation of oogenesis: the piscine receptor for vitellogenin. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid*. Metabolism 1045, 271–279. https://doi.org/10.1016/0005-2760(90)90130-P.
Takahashi, A., Kanda, S., Abe, T., Oka, Y., 2016. Evolution of the hypothalamic-pituitary-gonadal axis regulation in vertebrates revealed by knockout medaka. Endocrinology 157 (10), 3994–4002. https://doi.org/10.1210/en.2016-1356.
Taranger, G.L., Carrillo, M., Schulz, R.W., Fontaine, P., Zanuy, S., Felip, A., et al., 2010.

Control of puberty in farmed fish. Gen. Comp. Endocrinol. 165 (3), 483–515. https://doi.org/10.1016/j.ygcen.2009.05.004.
Teilum, K., Hoch, J.C., Goffin, V., Kinet, S., Martial, J.A., Kragelund, B.B., 2005. Solution structure of human prolactin. J. Mol. Biol. 351, 810–823. https://doi.org/10.1016/j.jmb.2005.06.042.
Trichet, V., Buisine, N., Mouchel, N., Morán, P., Pendás, A.M., Le Pennec, J.P., Wolff, J., 2000. Genomic analysis of the vitellogenin locus in rainbow trout (Oncorhynchus mykiss) reveals a complex history of gene amplification and retroposon activity. Mol. Gen. Genet. 263 (5), 828–837. https://doi.org/10.1007/s004380000247.
Trommsdorff, M., Gotthardt, M., Hiesberger, T., Shelton, J., Stockinger, W., Nimpf, J., et al., 1999. Reeler/disabled-like disruption of neuronal migration in knockout mice lacking the VLDL receptor and ApoE receptor 2. Cell 97 (6), 689–701. https://doi.org/10.1016/S0092-8674(00)80782-5.
Uchida, K., Moriyama, S., Chiba, H., Shimotani, T., Honda, K., Miki, M., et al., 2010. Evolutionary origin of a functional gonadotropin in the pituitary of the most primitive vertebrate, hagfish. PNAS 107 (36), 15832–15837. https://doi.org/10.1073/pnas.1002208107.
Ulloa-Aguirre, A., Timossi, C., Damián-matsumura, P., Dias, J.A., 1999. Role of glycosylation in function of follicle-stimulating hormone. Endocrine 11, 205–215. https://doi.org/10.1385/ENDO.
Unal, G., Marquez, E.C., Feld, M., Stavropoulos, P., Callard, I.P., 2014. Isolation of estrogen receptor subtypes and vitellogenin genes: expression in female Chalcalburnus tarichi. Comp. Biochem. Physiol. Part - B: Biochem. Mol. Biol. 172–173 (1), 67–73. https://doi.org/10.1016/j.cbpb.2014.04.002.
Vischer, H.F., Marques, R.B., Granneman, J.C.M., Linskens, M.H.K., Schulz, R.W., Bogerd, J., 2004. Receptor-selective determinants in catfish gonadotropin seat-belt loops. Mol. Cell. Endocrinol. 224, 55–63. https://doi.org/10.1016/j.mce.2004.06.011.
Wallace, R.A., Hoch, K.L., Carnevali, O., 1990. Placement of small lipovitellin subunits within the vitellogenin precursor in Xenopus laevis. J. Mol. Biol. 213 (3), 407–409. https://doi.org/10.1016/S0022-2836(05)80203-7.
Wang, H., Tan, J.T.T., Emelyanov, A., Korzh, V., Gong, Z., 2005. Hepatic and extrahepatic expression of vitellogenin genes in the zebrafish, Danio rerio. Gene 356 (1–2), 91–100. https://doi.org/10.1016/j.gene.2005.03.041.
Wang, N., Wang, R., Hu, Q., Xu, W., Zhu, Y., Yan, F., Chen, S., 2017. Characterization of a low-density lipoprotein receptor, Lrp13, in Chinese tongue sole (Cynoglossus semilaevis) and medaka (Oryzias latipes). Fish Physiol. Biochem. 43 (5), 1289–1298. https://doi.org/10.1007/s10695-017-0372-1.
Wang, R., Gao, Y., Zhang, L., Zhang, Y., Fang, Z., He, J., et al., 2010. Cloning, expression, and induction by 17-β estradiol (E2) of a vitellogenin gene in the white cloud mountain minnow Tanichthys albonubes. Fish Physiol. Biochem. 36 (2), 157–164. https://doi.org/10.1007/s10695-008-9222-5.
Watahiki, M., Yamamoto, M., Yamakawa, M., Tanaka, M., Nakashima, K., 1989. Conserved and unique amino acid residues in the domains of the growth hormones. J. Biol. Chem. 264 (1), 312–316.
Whittington, C.M., Wilson, A.B., 2013. The role of prolactin in fish reproduction. Gen. Comp. Endocrinol. 191, 123–136. https://doi.org/10.1016/j.ygcen.2013.05.027.
Williams, V.N., Reading, B.J., Hiramatsu, N., Amano, H., Glassbrook, N., Hara, A., Sullivan, C.V., 2014. Multiple vitellogenins and product yolk proteins in striped bass, Morone saxatilis: molecular characterization and processing during oocyte growth and maturation. Fish Physiol. Biochem. 40 (2), 395–415. https://doi.org/10.1007/s10695-013-9852-0.
Yang, H.M., Wang, Y., Wang, Z.Y., Wang, X.X., 2017. Seasonal and photoperiodic regulation of reproductive hormones and related genes in Yangzhou geese. Poult. Sci. 96 (2), 486–490. https://doi.org/10.3382/ps/pew340.
Yasuda, A., Kawauchi, H., Papkoff, H., 1990. The complete amino acid sequence of prolactin from the sea turtle (Chelonia mydas). Gen. Comp. Endocrinol. 80 (3), 363–371. https://doi.org/10.1016/0016-6480(90)90185-O.
Yilmaz, O., Prat, F., Ibañez, A.J., Amano, H., Koksoy, S., Sullivan, C.V., 2015. Estrogen-induced yolk precursors in European sea bass, Dicentrarchus labrax: status and perspectives on multiplicity and functioning of vitellogenins. Gen. Comp. Endocrinol. 221, 16–22. https://doi.org/10.1016/j.ygcen.2015.01.018.
Zatylny-Gaudin, C., Cornet, V., Leduc, A., Zanuttini, B., Corre, E., Le Corguillé, G., et al., 2016. Neuropeptidome of the cephalopod sepia officinalis: identification, tissue mapping, and expression pattern of neuropeptides and neurohormones during egg laying. j. Proteome Res. 15, 48–67. https://doi.org/10.1021/acs.jproteome.5b00463.
Zechel, C., Shen, X., Chen, J., Chen, Z., Chambon, P., Gronemeyerl, H., 1994. The dimerization interfaces formed between the DNA binding domains of RXR, RAR and TR determine the binding specificity and polarity of the full-length receptors to direct repeats. The EMBO Journal 13 (6), 1425–1433.
Zhang, S., Dong, Y., Cui, P., 2015. Vitellogenin is an immunocompetent molecule for mother and offspring in fish. Fish Shellfish Immunol. 46 (2), 710–715. https://doi.org/10.1016/j.fsi.2015.08.011.
Zhong, L., Yuan, L., Rao, Y., Li, Z., Zhang, X., Liao, T., et al., 2014. Distribution of vitellogenin in zebrafish (Danio rerio) tissues for biomarker analysis. Aquatic Toxicol. (Amsterdam, Netherlands) 149, 1–7. https://doi.org/10.1016/j.aquatox.2014.01.022.
Zohar, Y., Muñoz-Cueto, J.A., Elizur, A., Kah, O., 2010. Neuroendocrinology of reproduction in teleost fish. Gen. Comp. Endocrinol. 165 (3), 438–455. https://doi.org/10.1016/j.ygcen.2009.04.017.