

## Supplementary Material

### **Molecular cloning, characterisation and expression analysis of the vitellogenin genes *vtgA<sub>o1</sub>* and *vtgC* during ovarian development in Chinese hook snout carp *Opsariichthys bidens***

Huayu Jiang<sup>A,\*</sup>, Daojun Tang<sup>A,\*</sup>, Xinming Gao<sup>A,\*</sup>, Chenwen Lin<sup>A</sup>, Binbin Feng<sup>A</sup>, Chen Du<sup>A</sup>, Shan Jin<sup>A</sup>, Junquan Zhu<sup>A,B</sup>

<sup>A</sup>Key Laboratory of Applied Marine Biotechnology by the Ministry of Education, School of Marine Sciences, Ningbo University, Ningbo, Zhejiang 315211, PR China.

<sup>B</sup>Corresponding author. Email: [zhujunquan@nbu.edu.cn](mailto:zhujunquan@nbu.edu.cn)

**Table S1.** Information on *O. bidens* used in this study.

**Table S2.** The Genbank accession numbers in multiple sequence alignments and NJ phylogenetic tree.

**Fig. S1.** Full-length *Ob-vtgA<sub>o1</sub>* cDNA and deduced amino acid sequence.

**Fig. S2.** Full-length *Ob-vtgC* cDNA and deduced amino acid sequence.

**Fig. S3.** Multiple sequence alignment of *Ob-VtgA<sub>o1</sub>*.

**Fig. S4.** Multiple sequence alignment of *Ob-VtgC*.

\*These authors contributed equally to this study.

**Table S1. Information on *O. bidens* used in this study**

Month age	Sampling date	Number	Average body length/cm	Average body height/g	Ovary stage
6	2019.01	15	8.79±0.65	8.5±1.8	II~III
7	2019.02	12	9.46±0.68	12.6±3.5	II~III
8	2019.03	14	9.87±0.86	14.6±3.0	II~III
9	2019.04	16	9.55±0.96	15.4±3.2	II~III
10	2019.05	12	10.52±0.72	22.6±5.3	III~IV
11	2019.06	15	12.44±0.90	28.8±7.1	III~IV
12	2019.07	12	13.59±1.16	31.1±5.2	IV~V

**Table S2. The Genebank accession numbers in multiple sequence alignments  
and NJ phylogenetic tree**

Species	Vtg type	Accession number (NCBI)
<i>Labeo catla</i>	VtgAo1	ABP04034.2
<i>Danio rerio</i>	VtgAo1	NP_001038362.3
<i>Danio rerio</i>	VtgAo2	NP_001038378.1
<i>Danio rerio</i>	VtgC	NP_571340.1
<i>Ctenopharyngodon idella</i>	VtgAo1	APB93352.1
<i>Ctenopharyngodon idella</i>	VtgAo2	APB93353.1
<i>Ctenopharyngodon idella</i>	VtgC	APB93355.1
<i>Cyprinus carpio</i>	VtgAo1	BAF73406.1
<i>Cyprinus carpio</i>	VtgAo2	BAD51933.1
<i>Gambusia affinis</i>	VtgAa	BAD93697.1
<i>Gambusia affinis</i>	VtgAb	BAD93698.1
<i>Gambusia affinis</i>	VtgC	BAD93699.1
<i>Morone americana</i>	VtgAa	AAZ17415.1
<i>Morone americana</i>	VtgAb	AAZ17416.1
<i>Morone americana</i>	VtgC	AAZ17417.1
<i>Mugil cephalus</i>	VtgAa	BAF64835.1
<i>Mugil cephalus</i>	VtgAb	BAF64836.1
<i>Mugil cephalus</i>	VtgC	BAF64837.1
<i>Thunnus thynnus</i>	VtgAa	ACX32463.1
<i>Thunnus thynnus</i>	VtgAb	ADD63987.1
<i>Thunnus thynnus</i>	VtgC	ADD63988.1
<i>Melanogrammus aeglefinus</i>	VtgAa	AAK15158.1
<i>Melanogrammus aeglefinus</i>	VtgAb	AAK15157.1
<i>Oryzias latipes</i>	VtgAa	BAB79696.1
<i>Oryzias latipes</i>	VtgAb	BAB79591.1
<i>Larimichthys crocea</i>	VtgAa	AKK31325.1
<i>Larimichthys crocea</i>	VtgAb	AKK31327.1
<i>Larimichthys crocea</i>	VtgC	AKK31326.1
<i>Gobiocypris rarus</i>	VtgC	ATX74745.1
<i>Cirrhinus molitorella</i>	VtgAo1	ADB77954.1
<i>Tanichthys albonubes</i>	VtgAo1	ABN13867.1
<i>Conger myriaster</i>	VtgAe	BAD93275.1
<i>Anguilla japonica</i>	VtgAe1	AAV48826.1
<i>Anguilla japonica</i>	VtgAe2	AAR82899.1
<i>Anguilla japonica</i>	VtgAe3	AAR82898.1

---

<i>Oncorhynchus clarkii</i>	VtgAs	AGQ04606.1
<i>Oncorhynchus mykiss</i>	VtgAs	CAA63421.1
<i>Salmo salar</i>	VtgAs	XP_014024135.1
<i>Crassostrea gigas</i>	Vtg	BAC22716.1
<i>Haliotis discus hannai</i>	Vtg	BAF98238.1
<i>Scapharca broughtonii</i>	Vtg	AYE92811.1

---

```
1      M R A V V L A L L T V A L V A S Q Q
1  GACTATCAGCCATGAGGCGTGTGGCTTGCCTGACTGAGCCCTTGTGGCAAGTCAAC
18   I N L V P E F A K D N T Y E Y K Y E A Q
61  AGATCAACCTTGTCTGAGTTGCAAGGATAACACCTATGAGTACAAGTATGAGGCTC
38   L L G G L P Q E G L A R A G M K V S C K
121 AGCTCTGGGTGGTCTTCCCAAGAGGCTCGGCCGAGCAGGTATGAAAGTCAGCTGCA
58   V H L R R L T D N T F M M K L I D P Q L
181 AGGTTCACTCAGCCCTGACAAGAACACCTTCATGATGAAGTAAATAGATCCTCAAC
78   Q E Y A G I W P Q S S Y V P A P K L T S
241 TCCAGGAGTATGCTGGTATTTGGCCAGACTTTCATATGTTCTGCCCTAAGCTCACT
98   A L A P Q L Q T P T K F E Y A N G V I G
301 CAGCTTGGCTCCTCAGCTTCAGATTCACATCAAGTTTGAATGCTAATGGTGAATTTG
118  K V F A P A G V S P T V L N L H R G I L
361 GAAAGTATTCCCCCTCAGAGGCTCCCTCCCTACAGTACTGAACTTGACAGAGGAATCC
138  N L Q L N L K K Q N I Y E L Q E A G
421 TCAACATCCTCAGCTCAACCTCAAGAACCCAGAACCTATGAGCTGCAAGAGGCTG
158  A D G V C R T Q Y V I S E D F K T K H
481 GAGCTCAGGAGCTGCGAGGCCAGTATGCTCAGTGGAGTCCAAGACCAACACAC
178  Y V T K S K D S H C Q E R I M K D V G
541 TCACCTGCACCAACTCAAGATTCAGCCACTGCCAAGAGATCATGAAGCAGCTTG
198  L A Y T E K C A G E C I E R T K S L I E T
601 GCTTGGCATACTGAGAATGCTGTAAGTACATAGAAAGGCAAGGACTGATGTAAGAA
218  A S Y N I Y I M K P G A T G V Q I T E A
661 CTGGCTTCAACTACATCAATGAACACCCAGCAACTGGTGTACAGTCACTGAAGCAA
238  V E E V M Y Q F S P F S E I H G A A Q M E
721 CAGTTGAGGAAGTATCACTTCCACCTTCACTGAGTCCAGTGGTGGCCAGATGG
258  A K Q T L A F V E I K K T P V A P I K G
781 AAGCAAAACCACTTGGCTTGTGAGATTAAGAAAACCCCTTGGCTCAATCAAG
278  D Y L A R G S L Q Y E F A T E I L Q T P
841 GTGATCTTGGCCCTGGATCCCTGCACTTGAAGTTGCAACTGAAATTTCCAGACCC
298  I Q L R V T Q A Q A Q I S E V L K H L
901 CCATTCACTCATGAAGTCACTGATGACCAAGCCAGATAGCCAGGCTTGAAGCACC
318  V E N N V A R V H E D A P L K F V Q L V
961 TTGTTGAAAACAATGGCCAGGTCATGAGGATGCTCCACTTAAGTTTGTTCAGCTCG
338  Q L L R V T L E K I D A I W S Q F K D
1021 TCCAGCTTGGCTTCCACCTTGGAGAAAATGATGCCATCTGGTCTCAGTTCAAGG
358  K P V H R R M L L D A I P A V G T P V I
1081 ACAAAACAGTTCACAGGCGCTGGCTTCTGGATGCTATTCCTGCTTGGAAACCCGGTCA
378  L K F I K E K F L A G K C T T P E F I Q
1141 TTCTGAAGTTCATCAAGGAAAGTCTGGCTGGTAAATGTAACCTCCGAGTTTATTC
398  A L V V A L Q M V P A N L D T I Q L T A
1201 AGSCTCTGCTGCTCAGAAATGCTGCAATGGTCCCTGAAATTTGGACACCACTCAGVTTGACAG
141  S L A M E R K K I A T I P A L R E V M L
1261 CTAGTTTGGCTATGGAGAGGAAATGCGCACAAATCCAGCTCTCTGTAAGTTTGTATGC
438  G H G S V I A K H C V E V P T C P A E L
1321 TTGGCATGGTTCCTGATTTGCCAAACACTGCGTTGAAGTTCCCACTTGCCTGCCGAGC
458  L R P I H E L I A E A I S K N D I P E I
1381 TCCTCAGGCCATCCATGAGCTTATTCAGAGGCAATTTCCAGAAGTACATTCCTGAAA
478  T L A L K V L G N A G H P A S L K P I M
1441 TCAGTTTGGCTGAAAGTTCGGGCAATCTGGTCACTGCTGACTTAAACCCATCA
498  K I L P L L K S A A A A M P I R V Q V D
1501 TGAAGATCCTACCTACTAATAATCAGCAGCTGCTGCTATGCCATAGAGTCCAGTTG
518  A I L A L R N I A K K E P K L V Q P V A
1561 ATGCCATCTTAGCCCTGAGGAACATGCAAGAAAGAGCCAACTGGTTCCAGCCAGTG
538  L Q L L F D R A L H P E L R M V A S I V
1621 CCCTGACGCTTTATTCAGCAGGCTTCCACCAGAAGTGGTATGGTGGCTTCTATTG
558  L F E A R P L V A L M S S L A A Y L K T
1681 TGTGTTTGGAGCAAGGCCCTTAGTGGCTCTCATGCTCCAGTCTTGGCTGCTTATTTGAAGA
578  E T N M H V A S F S Y S Y I K S L T R I
1741 CTGAGCTAACATGATGTTGCAAGCTTTAGCTACTTCACTACATCAAGTCTTGACCAAG
598  N A P D M A A V A G A A N V A I K L M S
1801 TCAATGCTCCTGATATGGCAGCTGTTGGTGGTCAAGTAACTGCTATCAAGCTCATGT
618  R K L E R L S F R F S R A L Q D F Y H
1861 CCCGAACTAGAGAGGCTTAGCTTCCGTTTCCAGCAGAGCCCTTCCAGCAGGCTTCTATC
638  T P L H I G V A G S A Y M I N D A T I
1921 ACACCTCTTAATGATGGAGTCTGAGTGGCTTCACTGATGATCAATGATCCGCCACCA
658  L P R A V V A K R A Y L A G A S A D V
1981 TCTTGGCCAGAGCTGTGTAGCTAAAGCAGCTTCTTCCCTGGCTGGAGCTTCTGCTGATG
```

```
678  A E I G V R T E G I Q E A L L K S P A A
2041 TTGCTGAGATTGGTGTGAGAACTGAAGGAATCCAGAGGCTCTCTGAAAATCTCCTGCTG
698  D E S V D S S T K V K R T L K A L A N W
2101 CAGATGAAAGTGTGACAGTAGCACAAGGTAAAGCTTACCTGAAAGCACTTGCAAACT
718  R T L P T N Q P L A S A Y V K L L G Q E
2161 GGAGGACCTTGCCAAACCACTCAGCCATGGCTTCAGCATATGCAAAATCTTGGACAAG
738  V A H L I I D K T I I E E A I P T I T G
2221 AAGTGGCTCATTAACTCATTGACAAGACCATCATTGAGAAGCAATACCAACTATCAGT
758  P K S R E V L K G A L K A L Q E G I S F
2281 GACCAAACTCAGTGAAGTGTGAAAGGTGCCCTCAAGCCCTGCAAGAGGAATTTCTC
778  Q Y T K P L L A S E V R R I L P T A V G
2341 TTCAGTACCAAAACCCCTGCTGCTCATGAGTGGCTGCTATCTTCCCAACAGCCTTG
798  V P I E L S L Y T S A V A A T S L K V K
2401 GTGGCCCAICGAGCTCAGTTTGTACACTTCTGCTGTCGTCGTCACAGCTTCAAGTTA
818  A T I T P P L P G D I D T M T L G R L K
2461 AGCCCACTATTACACCTCCTCCCTGGGGATATTGACACCATGACTTGGGGCGGTGA
838  R T D V E L Q A E A R P S V A L Q T F A
2521 AGCCCACTGATGTTGAATCCAGCTGAAAGTGAAGTGAAGTGGTCTTCCAGCAGCTTG
858  V M G V N T A L I Q A A V T T A K G K I H
2581 CTGTGATGGAGTAGAACCTGCTTGAATCCAGCTGCTGTTACGGGCAAGAAAGATTC
878  V N A P G K V V A R A D I L K G N F K V
2641 ATGTAATGCCCTGAAAAGTGGTGAAGAGCAGATATCTCAAGGCAACTTTAAGG
898  E L L P V E V P G H I A A V S F E T F A
2701 TGGAGCTTGGCTGTGAGTTCAGGCTTGGAGACACTGCTGCTGAGACTTGGAGCTTTG
918  V V R N I E N P T S E R I V P L V P E V
2761 CTGTGCTGAGAACTGAAAATCCTACTTCTGAAAAGATTTGCCCTTGAAGCTGAGG
938  S Q Q K S Q T S Y D D S S E M P D E T
2821 TGTCCCGCAAAAATCCAGACATCATATGATGATGCTGATGAGTCAAGGCGCAGATGAGA
958  P V R A P A L F K Q I C I N I L G Y M
2881 CTCTGTGAGAGCTCCTGCTCTATTAAAGCACTGTTTAAATATCTTGGTGTCAAGG
978  C F E V K S H N A A F I R K A A A L Y M
2941 TGTGCTTGGAGTGAATTCACAAATGCTGCTTATCAGAAAAGCTGCTTACTACTA
998  I G K H S A R A E A E A R G E G P E V E R
3001 TGATTTGAAAGCACTCAGCCCTGCTGAAGAGGCAAGAGTGAAGCTTCCAGGTTGAA
1018  L E L E V Q V G F R A A E R L G R Q N R
3061 GACTGGAGCTTGAAGTCCAAGTGGTCTCAGAGCTGCTGAGAGCTCGTGAAGCAAAA
1038  L M D E F G T P E G K A F I L K L Q T L
3121 GCCTCATGGATGAGGGACTCCAGATGGAAGGCCCTTCTGTGAAGCTGAAGCAATTC
1058  D I G K T V E E I G E I V E G
3181 TGGATATTGAGGTAAGCTGTGCTGAATCTTCTGAAAGCAGCAGCAGCAGCGGCA
1078  H N Q M R V T D T A T T
3241 GCACAAACGCACTAAGGCTTCCATGTCAGCTCTGCTGAGCCGATCTGCCACCA
1098  M E P P R K F H K D Q Y N E T H R T S K
3301 CCATGAGGCTTCCAGAACTCCACAAGTCACTGATGAAAGCCATGCCACTTCAA
1118  G R S S S T T G S S F E A V Q K N L G
3361 AGGGAAGACGATGGAAGTGCACATCTAGCTTTCAGGCTTCCAGAAACAGATCTG
1138  N D N P V V F A V I A R V R A D E P L
3421 GAAATGATACTCAACTGTTTGGCTGCTATGCTGCTGTTGGAGAGCTCAGCAGCCG
1158  G Y Q L A A Y F D K P T A R V L L V I S
3481 TGGCTACCAACTTGGGCTTACTTTCAGCAGCACTGCAAGAGTCTGCTGCTATT
1178  S V A E N D N Q K I C V D G L L S K H
3541 CCTCCGCTGAAATGACAACCAAGATCTGGCTGAGGCTTACGTTCTGAGGCAAGC
1198  K A A A K V S W G P E C K Q Y S V T A K
3601 ACAAGCTGTCGCAAGTCTTGGGCTCCAGATGCAAAAGTATTCCTGACTGCTA
1218  V D A G V I L G R Y P A A R I F L E W E R
3661 AAGTTGATGCGGCTGCTGGTATGATCCCGCAATGCTGTCAGGATATTCCTGACTGAG
1238  L P V I F T T Y A K M L S K H I P T M A
3721 GGCTGCAAGTATTTCAACCACTAGCCAAATGCTGTCAGGATATTCCTGACTGAG
1258  C Q T G F S A E R A K N S E K E I E L T
3781 CTGCGCAGAGGCTTGTGCTGAAAGAGCAAGACAGCGAAGAAAGAAATGAACTGA
1278  A A L P T Q R F L N V I V R V E M T L
3841 CTGAGCCTTGGGACTCAGAGATTTGAAATGCTATGTTAGAGTTCCAGAGATGAC
1298  S R M A I P L P G A V P I N P D G T L S
3901 TGTCAAGGATGGCTTCTTCCAGGCGCTTCCCAATCAATCCAGATGGAATCTT
1318  I H V D E D I R S W I Q Y I E E E *
3961 CCATTCATGTAGTAGGAGCATTGCTCTGGATCCAGAAATATATCGAGGAAGAAATGAG
4021 AAATTAATACATTTTACATGTTCACTACATTAATTAATGAAATCAATTAAGTGTGCA
4081 TGTAAATGAAATCACTAATGCTAATAAGACCAGCAAGTATAAAAAAAAAAAAA
```

**Fig. S1.** Full-length *Ob-vtgAol* cDNA and deduced amino acid sequence. Blue letters above nucleotide sequence represent the corresponding amino acid sequence. Start (ATG) and stop (TAA) codons are indicated with grey boxes. The first 15 amino acids (underlined in blue) form the signal peptide sequence. Yellow highlights indicate the predicted Vtg receptor-binding region, and green highlights indicate the polyserine site of the Pv domain. Black double underlines indicate polyadenylation signal and poly-A sites.

1                               M R W L Y L C L L V A L A A S  
 1 ACCTCTGGACTTAATCAEATCAGTGCAGGGGTGGTGTATCTCTCTCTCTGCTGCTGTCCTGGTGGCCCTGGCTGCCAG  
 16 E T V N Y E P F L N F K K T Y E Y K Y E  
 61 CGAGACTGTCAATTATGAGCCTTCTTGAACCTTTAAAAAAACGTACGAGTATAAATGAG  
 36 G L V R V G R E L P D L V E S A V K L R  
 121 GGGACTGGTACGATGGGACGGGAATGCCCAGACCCTGGTCGAGTCCGAGTGAAGCTGAG  
 56 C T F E I T G E S P Q T F V L Q I S N V  
 181 GTGCACCTTGAATACTCAGTGGTGAATGCCAACAATTTGTACTTCAGATCTCAAATGT  
 76 D F E D F S G I P G K S D F S S S Q K L  
 241 AGACTTTGAGACTTCAGTGGCATCTGGGAAAAGTACTTTAGCTTCCCAAAAAGCT  
 96 T K R L S A E F R Q P I V F E F S K G Q  
 301 CACTAAGCGTCTATCTCCGAGTTCAGACAACCAATCGTTTTCGAATTCCTCAAGAGGACA  
 116 I T D T I R T V P G V S N S V V N T V R G  
 361 AATTACTGCATTCGGACTTCAGTGGGTCTCAAATTCAGTGTAAATATTGGAGGGG  
 136 V L G F L Q V T V K T T Q S F Y E L V E  
 421 GGTCCCTGGATTTTACAAGTACACAGTCAAAACCAACCAAAAGTTTATGAACTGGTGGGA  
 156 L G I H G V T C Q S S Y T V E E D S K A K  
 481 GTTGAAGATCATGGTGTATGTCAAGAGTTCACTGTTGAAGAAGACTCCAAGGCCAA  
 176 V L L V F T K M V D T N C Q Q P A A W Y  
 541 AGTGTGATAGTACCAAAAGGGTGTAGTGCACCAACTGTCAACAGCCCACTGCTGTGTA  
 196 S G M A L P E D K L S R Q R G E S V V  
 601 CAGCGGTATGGCTCTGCCCTGAGACAACTTAGCAGACAGAGGCGTGAAGTGTGT  
 216 S T V K H N Y T V K S T E N G G L I T K  
 661 TTCACAGTAAACACACTACAGTGAAGCTCCAGGAAAACGCGGTCTAATCACTAA  
 236 A F A Q E R R Q S F S P F N V K G G N S R  
 721 AGCGTTGCTCAGGAGCCTAATATTCTCCGTTTAACTAAAGGAGGAAACTCTCG  
 256 L L A L R D I E L L K V S E I E K K I V  
 781 AATTGGCTTACCGGCACTTGAAGCTTCAAAGTTCGAAATAGAGAAAATAGT  
 276 I G K V L S R G N L D P L P K I L D L Y I F  
 841 GATCGGAAGGTACAAGTAGAGGCCAATTAATGTACAGACAGAGAAAGACTCCAACK  
 296 I Q V L M I N L D P L P K I L D L Y I F  
 901 AATCCGTGATGANTCAACTGAAGCACTTACTTAAGATTTAGATTTAATTA  
 316 R L A Q A N I Y H V N S A N S T D I L D  
 961 GCGCTGGCTCAGGCAACATAATGCAAGTCAAAACAGCAGAGATTTCTGGA  
 336 L I Q L R V A T L E N L K Q I W K Q V  
 1021 TCTAATTGATTCCTGCCAGTTCGAAAGCTTGAANAATCAAAAGCAGATTGGAAGCAGGT  
 356 S G N D E H R R W F L D L V V E V T D E  
 1081 CTCAGGAATGATGAGCAGAGGAGTTCGCACTGGTGTGGAGGTAAACAGATGA  
 376 R I L K F L P G Y S A G S E K L P I R V Q  
 1141 AAGGATCCTCAAATTCCTGAGACTAAGTCAAAGCAGGAGCACTTACAGTGAATGAGG  
 396 G Q A I V V A F N H L P A K P V S V A L  
 1201 AGGACAGGCACTTGGTGGCATTTAACCACCTTGCCTGCCAAGCTGTATGATGGCATT  
 416 A Q V F L T I P F S K S Q P L L W N T V  
 1261 GCGTCAAGTGTTCGACGATTTCTTCAGTAATCCCAACCTCTCTCTGGCAACACTG  
 436 V L A Y G S L Y R Y C V Y T D P C P V  
 1321 CGTTTTGGCATATGATCTTCTTATACAGATCTGGTGTATAGTATCCCTGCCCTGT  
 456 T V V Q P L L D M A A S S L S K N S E V  
 1381 CACTGTGGTGCAGCCATTAAGTGCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT  
 476 D M V L A L K L G N A A H P S S I K T  
 1441 GGAATGGCTCTGCATTAAGCTTTGGGAATGCAGCTCATCCCTCCAGCATCAAGAC  
 496 L L K F L P G Y S A G S E K L P I R V Q  
 1501 TCTCTAAAGTCTCTCCAGGACTCACTGGATCTGAAAAGTTCACATCAGATGACA  
 516 G A A V Q A F R L L A G R N P H N V Q D  
 1561 GGGTCTGCAGTCCAGGCAATTTAGTCTGCTGCAAGGACCCCAACATGTCAGGA  
 536 I V L N L F V K K T L P A E T R M L A C  
 1621 TATTCCTTGAACCTCCTTGTAAAGAGACTTCTCTGTAATCCGATCTTGGCCTG  
 556 M V L L D T K P S M V L I S A I T N V L  
 1681 CMTGGCTTCTAGACACCAAGCCATCCATGGTTTGTCTCAGCAATAACTAATGTCT  
 576 L E E N D L Q V A S F S Y S L L K G I A  
 1741 TTTAGAAGAGATGATCTCAGTGTGGAGCTTCTCCACTGCCGATAATTTGCCATGAAGT  
 596 K S R T P D N Q H L S T A R N I A M K I  
 1801 CAAGTCCCGCAACCTGTAAATCAAGTCTCCTGCTCAGCAATAACTAATGTCT  
 616 L T R K K L G H L S Y R Y S K S M H F D W  
 1861 TCTTACCGCAAAACCTGGCCACTGAGTTATCGTTACAGCAAGAGCATGCAATTTGACTG  
 636 F H D D F L F G T S T D V Y M L Q N E S  
 1921 GTTCCATGATGATTTTTATTTGGGACATCCACATGATTTTTATATGCGCAAATGAAA  
 656 L I P S K L M L K G K I H F I G R I L Q  
 1981 TCTCATCCCTCCAACCTATGCTAAGGGAATAATTCATTCATTTGGCAGAATATTGCA  
 676 F I E V G V H A D G I K E L L A G K I P  
 2041 ATTATAGAGGTTGGTGTCCACGACAGTGGAAATAAAGAGTGTTCAGCAGAAAATCC  
 696 Q L K K D L G D F A T I L N I L S D W Q  
 2101 TCAACTAAGAAGATTTAGGAGATTTTGGCAGCATCTGATATTTCTGACTGGCA  
 716 K L P K D R P L L T T Y V R M F G Q E A  
 2161 GAAGTACCGAAAAGACAGACCCTACTGACCACATAATGACGCAATGTTTGCAGGAAG  
 736 F L M D V N G D F I Q S I T K S F S P T  
 2221 CTTTTAATGGATTAACCGAGATTTCATCAGATATCAAAAAGTCAATTAGCCCTAA  
 756 P G K E S K V W A K I Q D V Q K G T S W  
 2281 TCCAGGGAAGGAGATAAAGTTTGGGCAAGATTCGAAGTTCACAAAAGGAACTTCG  
 776 H W T K P H L V Y E A R F M Q P T C L G  
 2341 GCACTGGACCAAGCCGACCTTGTATATGAAGCTGATTCAGAGCCAATGTTTGGG  
 796 L P V E I S K Y F S I V N A V T M K V K  
 2401 TCTCCAGTTGAAATCAGAAATATTTCATATGTAATGCTGTTTACAAGAAAGTTAA  
 816 A E I S P P P K E H L R E L L S S V I S  
 2461 GCGTAAATGATCCCTCCCAAGGAAATCTGCGTGAAGCTGCTAAGCTCAGTCAATTC  
 836 L Q T D G F A G V T K D H F V F H G I N  
 2521 CCGCAACAGATGGCTTTGCTGGAGTACAAGGATCATTTTGTCTTCCATGAACTCAA  
 856 T D L F Q C G V E M K T K M V T G L P W  
 2581 CAGATCTTTTCCAGTGTGCTGTGAGATGAAGCAATAAGTACTGCCTGCCATG  
 876 E F F G L K I N I K E Q K Y Q M N L T P S  
 2641 GNAATCGCGCTGAAAATAACATAAAGAGCAAAATGATCAATGACTTGACTCCGAG  
 896 K T D T E L F S V V I N S N A V L R N I  
 2701 CAAAACAGACAGAGTATTTCAGTCAATTCATATTCGCTCTTGGAGAATAT  
 916 E D P S L S K I T P M M P E S E D S Q Q  
 2761 TGAGGATCCATCTTATCCAAAATACTCCCATGATGCCTGATCAGAGGATCCAGCA  
 936 G L P L A R I I L P T S R G Q S N S  
 2821 GGGTTGGCATTGGCAGAAATATTTCACCACTCAAGGATGAACAGATGAAGCACT  
 956 E V K F R Q C A E A K I Y G T A C I E  
 2881 TGAGTGAATTCAGGCAATGTGCTGAGGCAGATATTGGAATGCCAATGCAATGTA  
 976 A E A K R S H Y L H E Y P L Y F G E L G  
 2941 GCGAGACCCAAACGATCACACTACATGATATGATTCCTCGTATATTTCCTGGAGT  
 996 T R F S Y K L E P A K S A K T I E I Q  
 3001 TACCCGCTTTCATATAAACAGAAAGACCAAGAGTGCAAACCTTGGAAAAATCA  
 1016 I E V T A G I K H P Q A F S E M M D L S  
 3061 GATTGAGTCACTGCAGGCAAAAACACCTCAAGCATTGAGTGAATGATGATGATCTCAG  
 1036 R R V F K D T R D E I T T C R E H N L S  
 3121 TCGCAGAGTATCAAGGACACAAAGAGAGAGGATCACCCTGCTGTAAGCAACATCTATC  
 1056 S S F L A N Q D L G L T P D P V T V K  
 3181 CAGTTCATTCCTGCAATCAGGACCTTGGTACTCTCAGCAAGCTGATGATCAGATAA  
 1076 A I S L S P P G K S L F G A V A F Y L  
 3241 AGCATTGCTCAAGTCCAGGAAATCTTGGGCTATGAGCTGATGAGCTTTTATCT  
 1096 P T A Q R D N V E M I V S E V G E A N  
 3301 GCCAACAGCACAAGAGCAATGTTGAATGATGCTCCTGSSAATGGTGAAGAAGCAAA  
 1116 W K M C A N A N A N V D K H S A K A H L  
 3361 CTGAAAATGTTGCCAATGCAAATGTAGATAAGAGTATTCATCAGCAAAAGGCCATCT  
 1136 R W G A E C Q T Y D V S M K V S S A C Q  
 3421 CAGATGGGTCAGAGTGTGAGCAATATGATGTATGTGAAGGTTCTCCAGCATGCA  
 1156 P E S K T S L Y T K I K W G A L P S V F  
 3481 GCCAGATCAAACTCTCTATACACCAAGATATAAGGGGAGCTGCCCTCAGTGT  
 1176 T I G Q T I Q E Y I P G L S Y T M G F Y  
 3541 CACAATGTCACAAATCAGAGTACATTTCCGCTTATCTTATACTATGAGGTTCTA  
 1196 Q K Y E K N P E R Q A A V I V V A S S T  
 3601 CCGAATATGAGAAATCCGAAACCTCAGACGCTCAGCTCAGTGTGTAGCTTCTCAC  
 1216 E T F D M K V F I P E R T I D K K A I F  
 3661 AGACACCTTGCATGAAAGTGAATAATCCAGAGCAACTCGACAAGGCAACTCC  
 1236 S P I E L V G F E A V N L T T \*  
 3721 CTCACAAATCGAACTTGGGGTTTGAAGTGTGAACTCACCAATGAGCCATAACGGC  
 3781 TGATGTCAAGAGCAGCTCTGGAAGCAAAATAACAAACACCAACATCGGAAAATA  
 3841 TGACATTTGATTACTGACTGATTTGGATATCAAATAATGATGCTATCAATTAATCT  
 3901 AAAATATGCACAGTGAATACAGATTCATTGAGAGTCAATTAATTCATCTCTGGA  
 3961 TGTGCTGTATTTCCAGCCTGAGTGTATTATATTATTTACCTAGAGATCAGCAACAT  
 4021 TCCCTTACATGTCAAAGGATTTTGCATAGACTGCAAAATCTGGAGAATGAATATATC  
 4081 ATAAAGACAGATATAGATGGGTGTTATGGTTGCTCCTGGACATCTTGGCTGAAGA  
 4141 TGAGACTACAGTGTAAATAGGTCAAGGAGGATGCAAAAGTAAAATGCTGATA  
 4201 AACTTTGTGAGGAACTGTCACTGTGACTCGAAATAATGTTGGAGTCTTTTGT  
 4261 AGTTACACATGCACATGCAAGTGTCTATCTGCTTTTGCAGCAACTGTGATATATGA  
 4321 TGTAAACAATGCAACAATGTTGCTAAGCAGATTTTTAATTAAGTACAATAATCAT  
 4381 GCAAAAAAAAA

Fig. S2. Full-length *Ob-vtgC* cDNA and deduced amino acid sequence. Gray boxes indicate start (ATG) and stop (TGA) codons. The predicted Vtg receptor-binding region is highlighted in yellow. The first 15 amino acids (blue underline) form the signal peptide sequence. Black double underlines show the polyadenylation signal and poly-A sites.







**Fig. S3.** Multiple sequence alignment of *Ob-VtgAol*. Black arrow indicates the cathepsin D hydrolysis site, the black line separates VtgAol signal peptide from the three yolk protein domains, red box is the conserved motif RGILN. Red arrow indicates the cathepsin D hydrolysis site, the red line separates the five yolk protein domains of the VtgAa signal peptide, and the red box is the conserved motif CGXC (X represents any amino acid). The predicted Vtg receptor-binding region is indicated with grey shadow.

	1	Signal peptide	↓	Lipovitellin heavy chain	100
<i>Opsariichthys bidens</i> VitC	(1)	MRWYI		LDVA	MAE
<i>Gobiocypris rarus</i> VitC	(1)	HWQYI		LDVA	MAE
<i>Danio rerio</i> VitC	(1)	HWRLC		LDVA	MAE
<i>Ctenopharyngodon idella</i> VitC	(1)	HWRYI		LDVA	MAE
<i>Gambusia affinis</i> VitC	(1)	WRGLL		LDVA	MAE
<i>Morone Americana</i> VitC	(1)	WRGLL		LDVA	MAE
<i>Thunnus thynnus</i> VitC	(1)	WRGLL		LDVA	MAE
<i>Mugil cephalus</i> VitC	(1)	WRGLL		LDVA	MAE
Consensus	(1)	M		GLLLCLL	VALATSQSVNYEPFLNPKTYEYKYEGLVNVGRGLFNLVESAVKLRCTVKIIGESPQTFVQLVQSNLDFEDNFGFPFGKS PSASPKLTRKIS
<i>Opsariichthys bidens</i> VitC	(101)	AEFR		LVSR	SKG
<i>Gobiocypris rarus</i> VitC	(101)	AEEF		LVSR	SKG
<i>Danio rerio</i> VitC	(101)	AEEF		LVSR	SKG
<i>Ctenopharyngodon idella</i> VitC	(101)	AEEF		LVSR	SKG
<i>Gambusia affinis</i> VitC	(101)	AEEF		LVSR	SKG
<i>Morone Americana</i> VitC	(101)	AEEF		LVSR	SKG
<i>Thunnus thynnus</i> VitC	(101)	AEEF		LVSR	SKG
<i>Mugil cephalus</i> VitC	(101)	AEEF		LVSR	SKG
Consensus	(101)	AELSOP		IMFEYSK	GIITDIRTAPGVSNVTVVNI
<i>Opsariichthys bidens</i> VitC	(201)	APDKL		SKRG	ESV
<i>Gobiocypris rarus</i> VitC	(201)	APDKL		SKRG	ESV
<i>Danio rerio</i> VitC	(201)	APDKL		SKRG	ESV
<i>Ctenopharyngodon idella</i> VitC	(201)	APDKL		SKRG	ESV
<i>Gambusia affinis</i> VitC	(201)	APDKL		SKRG	ESV
<i>Morone Americana</i> VitC	(201)	APDKL		SKRG	ESV
<i>Thunnus thynnus</i> VitC	(201)	APDKL		SKRG	ESV
<i>Mugil cephalus</i> VitC	(201)	APDKL		SKRG	ESV
Consensus	(201)	AVEDKLS		KQRGESV	ISTVHKVHYTVKVSADGGLITKAFGQERQHFSPFNKGGSEFKM ALKRDVLLKVS
<i>Opsariichthys bidens</i> VitC	(301)	INLNP		AFPL	LR
<i>Gobiocypris rarus</i> VitC	(301)	IDLNP		AFPL	LR
<i>Danio rerio</i> VitC	(301)	IDLNP		AFPL	LR
<i>Ctenopharyngodon idella</i> VitC	(301)	INLNP		AFPL	LR
<i>Gambusia affinis</i> VitC	(301)	KLDNP		AFPL	LR
<i>Morone Americana</i> VitC	(301)	ONLNP		AFPL	LR
<i>Thunnus thynnus</i> VitC	(301)	ONLNP		AFPL	LR
<i>Mugil cephalus</i> VitC	(301)	ONLNP		AFPL	LR
Consensus	(301)	QNLNDP		VPEKILDLKRLQ	ANIYHDSATSDTIKLYQLLRV TYENLEQLKQVSGNDEHRRWFLLD
<i>Opsariichthys bidens</i> VitC	(401)	YAFNL		ISAK	Y
<i>Gobiocypris rarus</i> VitC	(401)	YAFNL		ISAK	Y
<i>Danio rerio</i> VitC	(401)	YAFNL		ISAK	Y
<i>Ctenopharyngodon idella</i> VitC	(401)	YAFNL		ISAK	Y
<i>Gambusia affinis</i> VitC	(401)	YAFNL		ISAK	Y
<i>Morone Americana</i> VitC	(401)	YAFNL		ISAK	Y
<i>Thunnus thynnus</i> VitC	(401)	YAFNL		ISAK	Y
<i>Mugil cephalus</i> VitC	(401)	YAFNL		ISAK	Y
Consensus	(401)	YAFNHL		A	PVSVLALQMLTTPPSKSNPLLNWTVLAYGSLVYKYCVVYDPCPVTVVQPLDMA SSLRNNS
<i>Opsariichthys bidens</i> VitC	(501)	RCYSI		AE	ST
<i>Gobiocypris rarus</i> VitC	(501)	RCYSI		AE	ST
<i>Danio rerio</i> VitC	(501)	RCYSI		AE	ST
<i>Ctenopharyngodon idella</i> VitC	(501)	RCYSI		AE	ST
<i>Gambusia affinis</i> VitC	(501)	RCYSI		AE	ST
<i>Morone Americana</i> VitC	(501)	RCYSI		AE	ST
<i>Thunnus thynnus</i> VitC	(501)	RCYSI		AE	ST
<i>Mugil cephalus</i> VitC	(501)	RCYSI		AE	ST
Consensus	(501)	EGYSAG		VKLP	RVQSAVQVFRLLAARDPHSVQDIVL LFLQK LPABIRMLACMILLDCKPSMALVSVVT VLQEEKLDQVAFSFSYSLKGI
<i>Opsariichthys bidens</i> VitC	(601)	YAFNL		ISAK	Y
<i>Gobiocypris rarus</i> VitC	(601)	YAFNL		ISAK	Y
<i>Danio rerio</i> VitC	(601)	YAFNL		ISAK	Y
<i>Ctenopharyngodon idella</i> VitC	(601)	YAFNL		ISAK	Y
<i>Gambusia affinis</i> VitC	(601)	YAFNL		ISAK	Y
<i>Morone Americana</i> VitC	(601)	YAFNL		ISAK	Y
<i>Thunnus thynnus</i> VitC	(601)	YAFNL		ISAK	Y
<i>Mugil cephalus</i> VitC	(601)	YAFNL		ISAK	Y
Consensus	(601)	DNQHL		STACNIAMK	ILTRKLGRLS



Fig. S4. Multiple sequence alignment of *Ob-VtgC*. Arrow indicates the cathepsin D hydrolysis site, and the straight line separates the *Ob-VtgC* signal peptide from the yolk protein domain. The predicted Vtg receptor-binding region is indicated with grey shadow.