Identification, expression analysis and the regulating function on C/EBPs of KLF10 in Dalian purple sea urchin, *Strongylocentrotus nudus*

Journal:	<i>Genome</i>
Manuscript ID	gen-2017-0033.R2
Manuscript Type:	Article
Date Submitted by the Author:	22-May-2017
Complete List of Authors:	Wu, Kaikai; Northwest Agriculture and Forestry University, College of Animal Science and Technology; Jia, Zhiying; Northwest Agriculture and Forestry University Wang, Qi`ai; Northwest Agriculture and Forestry University Wei, Zhenlin; dezhou university, Biological sciences Zhou, Zunchun; Liaoning Ocean and Fisheries Science Research Institute, Liaoning Key Lab of Marine FisheryMolecular Biology Liu, Xiaolin; Northwest A&F University, College of Animal Science
Is the invited manuscript for consideration in a Special Issue? :	This submission is not invited
Keyword:	Dalian purple sea urchin (<i>Strongylocentrotus nudus</i>), KLF10, C/EBPs (C/EBP α , C/EBP γ , C/EBP ζ), lipogenesis

SCHOLARONE™
Manuscripts

1 **Identification, expression analysis and the regulating function on**
2 **C/EBPs of KLF10 in Dalian purple sea urchin, *Strongylocentrotus***
3 ***nudus***

4 Kaikai Wu^a, Zhiying Jia^a, Qi`ai Wang^a, Zhenlin Wei^b, Zunchun Zhou^c, Xiaolin Liu^{a,*}

5

6 ^a *College of Animal Science and Technology, Northwest A&F University, Shaanxi Key Laboratory*
7 *of Molecular Biology for Agriculture, Yangling 712100, China*

8 ^b *Biological Science Department, Dezhou University, Dezhou, Shandong 253023, China*

9 ^c *Liaoning Key Lab of Marine Fishery Molecular Biology, Liaoning Ocean and Fisheries Science*
10 *Research Institute, Dalian, Liaoning 116023, China*

11

12

13

14

15

16

17

18

19 * Corresponding author: Tel.: +86 029 87054333; fax: +86 029 87092164.

20 E-mail address: liuxiaolin@nwsuaf.edu.cn (Xiaolin Liu).

Abstract

Accumulating evidence indicates that Krüppel-like factors (KLFs) play important roles in fat biology via the regulation of CCAAT/enhancer binding proteins (C/EBPs). However, KLFs and C/EBPs have not been identified from *Strongylocentrotus nudus*, and their roles in this species are not clear. In this study, the full-length cDNA of *S. nudus* KLF10 (SnKLF10) and three cDNA fragments of *S. nudus* C/EBPs (SnC/EBPs) were obtained. Examination of tissue distribution and expression patterns during gonadal development implied that SnKLF10 and SnC/EBPs play important roles in gonadal lipogenesis. The presence of transcription factor-binding sites (TFBSs) for KLFs in SnC/EBPs, and the results of an over-expression assay, revealed that SnKLF10 negatively regulates the transcription of SnC/EBPs. In addition, the core promoter regions of SnC/EBPs were determined, and multiple TFBSs for transcription factor (TFs) were identified, which are potential regulators of SnC/EBP transcription. Taken together, these results suggest that SnC/EBP genes are potential targets of SnKLF10, and that SnKLF10 plays a role in lipogenesis by repressing the transcription of SnC/EBPs. These findings provide information for further studies of KLF10 in invertebrates and provide new insight into the regulatory mechanisms of C/EBP transcription.

Key words

Dalian purple sea urchin (*Strongylocentrotus nudus*); KLF10; C/EBPs (C/EBP α , C/EBP γ , C/EBP ζ); lipogenesis

44 **Introduction**

45 Sea urchins are a valuable food resource for humans, and are farmed on a large
 46 scale for their edible gonads (Ernst 2011; Briggs and Wessel 2006; Ding et al. 2007).
 47 The Dalian purple sea urchin, *Strongylocentrotus nudus*, is an important economic
 48 species in Chinese aquaculture because of the nutritional value of various unsaturated
 49 fatty acids in its gonad (Castell et al. 2004). A better understanding of the mechanisms
 50 of lipogenesis in the *S. nudus* gonad may contribute to fattening and sustainable
 51 farming, thereby satisfying the human demand for seafood.

52 In recent years, a complex network of transcription factor (TFs) has been found to
 53 regulate adipogenesis and lipogenesis by coordinating the expression of numerous
 54 proteins (Rosen 2000a; 2000b). These TFs include multiple activators, co-activators,
 55 and repressors, such as Krüppel-like factors (KLFs), CCAAT/enhancer binding
 56 proteins (C/EBPs), and peroxisome proliferator-activated receptor γ (PPAR γ) (Farmer
 57 2006; Brey et al. 2009; Laprairie et al. 2016). KLFs, a family of 17 zinc finger TFs,
 58 play diverse roles in the regulations of cellular growth, differentiation, and
 59 development (Swamynathan 2010, Presnell et al. 2015). Furthermore, there is
 60 accumulating evidence that KLFs have an important role in fat biology. KLF15, KLF4,
 61 KLF5, KLF6, and KLF9 have been shown to be involved in the positive regulation of
 62 adipogenesis and lipogenesis, and KLF2 and KLF3 have been shown to be involved
 63 in the negative regulation of adipogenesis and lipogenesis (Kaczynski et al. 2003;
 64 Brey et al. 2009; Wu and Wang 2012; Matsubara et al. 2013). C/EBPs are a family of
 65 leucine-zipper TFs, which play key roles in adipogenesis and lipogenesis (Nerlov
 66 2007; Reddy et al. 2016). Thus far, several C/EBPs (C/EBP α , C/EBP β , C/EBP γ ,
 67 C/EBP δ , C/EBP ϵ , and C/EBP ζ) have been identified in humans (Darlington et al.
 68 1998; Madsen et al. 2014; Brey et al. 2009). However, KLF10 and C/EBPs have not

69 been identified in *S. nudus*, and their roles in this species are unclear.

70 In this study, the full-length cDNA of *S. nudus* KLF10 (SnKLF10) and three
71 cDNA fragments of *S. nudus* C/EBPs (SnC/EBPs) were obtained. Moreover, three
72 promoter sequences, for SnC/EBP α (1154 bp), SnC/EBP γ (1494 bp), and SnC/EBP ζ
73 (823 bp), were cloned. To investigate the potential roles of SnKLF10 and SnCEBPs in
74 lipogenesis, the mRNA expression profiles of SnKLF10 and SnCEBPs were
75 examined in gut, muscle, tube feet, and male and female gonads, and their RNA
76 expression patterns were examined in two developmental stages of male and female
77 gonads. To determine the roles mechanism of SnKLF10 and SnC/EBPs in lipogenesis,
78 the effects of SnKLF10 overexpression on the transcription of SnC/EBP genes were
79 investigated in 293T cells. Furthermore, to delineate the regulatory mechanisms of
80 SnC/EBPs, the core promoter regions of SnC/EBPs were determined by luciferase
81 reporter gene assays using a series of deletion vectors. The results of this study will
82 provide a better understanding of the roles of KLF10 and C/EBP genes, and the
83 regulation of C/EBP transcription in *S. nudus*.

84 **Materials and methods**

85 *Sample collection*

86 Mature *S. nudus* (second- and third-instar) were collected along the coast of
87 Dalian, Liaoning, China. Based on age and gender, the sea urchins were placed into
88 four groups, each containing three samples of the same gender and developmental
89 stage. The animals were acclimatized under laboratory conditions for 1 week in
90 seawater in aerated aquaria at 22°C, and fed twice daily. The gonad samples were
91 snap-frozen in liquid nitrogen for 24 h and stored at -80°C until use.

92 The samples were homogenized in TRIzol reagent (Invitrogen) and total RNA
93 was isolated according to the manufacturer's instruction. Total RNA was incubated

94 with RNase-free DNase I (Roche) to remove contaminating genomic DNA before
95 being reverse transcribed into cDNA using random hexamer primers and M-MLV
96 reverse transcriptase (Promega).

97 *Cells preparation*

98 The 293T cell line, kindly provided by Biotechnology Laboratory Animal Science,
99 Northwest A & F University, was cultured in DMEM (Invitrogen) containing 10%
100 fetal bovine serum (FBS) (Biosource), 100 IU/mL penicillin, and 100 µg/mL
101 streptomycin sulfate (Sigma). The cells were incubated at 37°C and 5% CO₂
102 humidity.

103 *Cloning the full-length cDNA of SnKLF10*

104 To identify the SnKLF10 cDNA sequence, KLF7F/KLF8R primers (Table 1)
105 were designed based on the sequence of the *S. nudus* transcriptome. To obtain the 3'
106 ends of the SnKLF10 gene, primer pairs KLF87F/UPM and KLF88F/NUP (Table 1)
107 were employed for the 3' RACE method (Clontech). Similarly, the 5' end of SnKLF10
108 gene was obtained by a 5' RACE system (Invitrogen), using the primers
109 KLF85R/AAP and KLF86R/AUAP (Table 1). The full-length cDNA sequence was
110 confirmed by primers KLF89F/KLF99R (Table 1). PCR was performed in a 25-µL
111 volume containing 1 µL of cDNA as a template, 0.5 µL of each primer (10 pmol/ml),
112 10.5 µL of PCR-grade water, and 12.5 µL Mix (CWBIO). PCRs were performed with
113 denaturation at 94°C for 4 min; 30 cycles of 94°C for 30 s, 56°C for 30 s, and 72°C
114 for 2 min; 72°C for 5 min. PCR products were purified using a Universal DNA
115 Purification Kit (TIANGEN) and cloned into a pUC-T vector using a pUC-T Ligasing
116 Kit (CWBIO). Positive clones were sequenced (Genescript).

117 *Confirming cDNA sequences of SnC/EBPs*

118 The primers C/EBP109F/C/EBP108R (Table 1) were designed to identify the

cDNA sequence of SnC/EBP α . The primers C/EBP113F/C/EBP31R and C/EBP193F/C/EBP194R (Table 1) were designed to identify the cDNA sequence of SnC/EBP γ . The primers C/EBP117F/C/EBP198R (Table 1) were designed to identify the cDNA sequence of SnC/EBP ζ . The PCR and protocols were performed as described above.

Sequence analysis

Nucleotide and protein sequence similarities were searched using the BLAST program (<http://www.ncbi.nlm.nih.gov/blast>). Transcription start sites (TSS) were predicted by the Berkeley Drosophila Genome Project (http://www.fruitfly.org/seq_tools/promoter.html). The deduced amino acid sequence was analyzed with the Expert Protein Analysis System (<http://www.expasy.org/>) and the Sequence Manipulation Suite programs (<http://www.bioinformatics.org/sms/>). The PSORT tool (<http://psort.hgc.jp/form.html>) was used to predict protein localisation sites in cells. Transcription factor-binding sites (TFBSs) in full-length promoter vector sequence were predicted using online software (<http://www.genomatix.de>). MicroRNA (miRNA) target sites in the 3'-untranslated region (UTR) of SnKLF10 were predicted on the website <http://genie.weizmann.ac.il/>. Protein domains were predicted using the Simple Modular Architecture Research Tool (SMART) (<http://smart.emblheidelberg.de/>). Amino acids sequences were obtained from NCBI (<http://www.ncbi.nlm.nih.gov/>). Phylogenetic trees were constructed based on multiple sequence alignments of amino acids by the neighbor-joining (NJ) algorithm embedded in the Mega 5.0 program.

Tissue expression analysis

Real-time PCR was used to analyze the tissue distribution of SnKLF10 and SnC/EBP mRNA in second-instar *S. nudus*. Tissues, including gut, muscle, tube feet,

male and female gonads, were collected from three males and three females. Gut, muscle, and tube feet samples from females and males were mixed. Total RNA extraction and cDNA synthesis were performed as described above. β -actin ([GenBank accession number KU143835](#)) served as an internal reference gene to normalize mRNA expression. The primers used for qRT-PCR are listed in Table 2. UltraSYBR Green Mix (CWBIO) was used in 20- μ L PCR. All qRT-PCR was performed in three duplicates on a Bio-Rad CFX96 Real-Time PCR Detection System (BioRad). The qRT-PCR conditions were as follows: 95°C for 10 min, 39 cycles of 95°C for 15 s, 57°C for 30 s, and 72°C for 20 s. Amplification efficiency and melting curve analysis were used to confirm the accuracy and specificity of PCR. The relative expressions of SnKLF10 and SnC/EBPs were calculated using the comparative cycle threshold (Ct) method with the formula $2^{-\Delta\Delta Ct}$ [$\Delta\Delta Ct = \Delta\Delta Ct$ (Test) - $\Delta\Delta Ct$ (Control)] where tube feet mRNA was used as the control.

Quantification of gene expression in two developmental stages

To delineate the function of the SnKLF10 and SnC/EBPs gene products in lipogenesis in male and female gonadal tissues, mRNA expression of these genes were quantified in second- and third-instar *S. nudus*. β -actin (Sn β -actin) served as an internal reference gene to normalize the mRNA expression. The specific primers used for qRT-PCR are listed in Table 2. Second-instar gonadal tissues were selected as the control. The following processes were performed as described above.

Cloning the 5'-flanking sequences of SnC/EBPs

First, primers (Table 1) were designed to amplify the intron sequences from genomic DNAs in the 5'-UTRs of SnC/EBPs. PCR and protocols were the same as those described above. Based on the obtained gene sequences of SnC/EBP α , SnC/EBP γ , and SnC/EBP ζ , three adjacent reverse primers (SP1, SP2, and SP3) (Table

3) were designed to amplify the 5'-flanking sequences of SnC/EBPs from genomic DNA, according to the protocol of the Genome Walking Kit (TaKaRa) with four universal primers (AP1–AP4) (Table 3). Genomic DNA was extracted from *S. nudus* muscle tissue using a Marine animal tissue genomic DNA extraction kit (TIANGEN).

Overexpression of *SnKLF10*

The open reading frame (ORF) of *SnKLF10* was amplified using LA Taq™ DNA polymerase (TaKaRa) by specific upstream EGFP-KF (containing a XhoI site) and downstream EGFP-KR (containing a KpnI site) primers (listed in Table 4). The PCR product (location [nt]: 45–2118) was purified, and digested with XhoI and KpnI enzymes (TaKaRa); the purified pEGFP-C1 plasmid was also digested using the same enzymes. Then, the target fragment was inserted into the pEGFP-C1 plasmid. Finally, the recombinant plasmid was obtained and verified by sequencing. The recombinant plasmid was named pEGFP-KLF10. Primers (pGL3-α1/pGL3-α8, pGL3-γ1/pGL3-γ8, and pGL3-ζ1/pGL3-ζ8; listed in Table 4) were used in the PCR, with *S. nudus* genomic DNA as a template to construct the full-length promoter vectors of SnC/EBPα (location [nt]: -1144–172), SnC/EBPγ (location [nt]: -1396–127), and SnC/EBPζ (location [nt]: -727–831). The primers used to construct vectors containing the KpnI (forward primer) and XhoI (reverse primer) restriction sites are listed in Table 4. Similarly, the target fragments were separately inserted into the pGL3-basic vector, which contained a luciferase reporter gene. The recombinant plasmids were named C/EBPα-1, C/EBPγ-1, and C/EBPζ-1. The PCR and protocols were performed as described above. All plasmid constructs were verified by sequencing.

The 293T cells (4×10^4 cells/mL) were transfected in 96-well plates with Lipofectamine 2000 (Invitrogen), in accordance with the manufacturer's instructions. Briefly, for each well, 0.3 μL of Lipofectamine 2000 and 100 ng of DNA (EGFP-,

pGL3-, and pRL-TK) were mixed in 125 μ L of FBS-free and pen/strep-free Opti-MEM I medium (Promega) for 20 min. To normalize the transfection efficiency, the pRL-TK plasmid vector (Promega), which carries a *Renilla* luciferase gene, was co-transfected with the reporter construct. The experiments were performed in triplicate for each construct. At 36-h post transfection, EGFP expression was assessed by imaging with a fluorescence microscope (Bio-Rad, USA). Cells were harvested 48-h post-transfection, and firefly and *Renilla* luciferase activities were measured using the Dual-Luciferase Reporter Assay System and a BHP9504 Fluorescent Analytic Instrument (Hamamatsu). Firefly luciferase activity was normalized using the *Renilla* luciferase activity in each well. Data represent the average of three replicates. Finally, total cellular RNA isolation and cDNA synthesis were performed as described above. mRNA expression was normalized to that of *Homo sapiens* GAPDH (HsGAPDH). The primers are listed in Table 2.

Exploration of promoter activity

To analyze the core promoter regions in the 5'-flanking sequence of SnC/EBP α (location [nt]: -1144–172), SnC/EBP γ (location [nt]: -1396–127), and SnC/EBP ζ (location [nt]: -727–831), primers (listed in Table 4) were used to obtain truncated promoter vectors. The truncated promoter vectors were generated using the full-length promoter vector as described previously. All plasmid constructs were verified by sequencing.

Cells were prepared as previously described, and then 293T cells were transfected in 96-well plates. Firefly and *Renilla* luciferase activities were measured after transfection for 48 h. The experiments were performed in triplicate for each construct. Data represent the average of three replicates.

Results

cDNA sequence analysis

The SnKLF10 cDNA (GenBank accession number [KU058873](#)) was 3526-bp long with an ORF of 1476-bp encoding 491 amino acids (aa), a 138-bp 5'-UTR and a 1912-bp 3'-UTR; “1” marked the TSS (Fig. 1A). The putative protein had an estimated molecular weight (MW) of 53.9247 kDa and a predicted isoelectric point (PI) of 9.66. Three highly conserved C2H2 Zinc finger (ZnF_C2H2) domains (371–395, 401–425, and 431–453 aa) were found to be located at the carboxyl-terminal (C-terminus) end of the SnKLF10 protein by domain analysis (Fig. 1B and Fig. 2). The predicted protein localization sites showed that SnKLF10 was located in the endoplasmic reticulum. Furthermore, based on the results obtained from the miRNA target site prediction programs, we identified a few putative target sites in the 3'-UTR of SnKLF10 (Fig. 1A).

BLASTP analysis revealed that SnKLF10 had the highest similarity to *S. purpuratus* KLF10 (81%) (Protein ID., [XP_794951](#)), followed by *S. kowalevskii* KLF11 (80%) (Protein ID., [XP_002731613](#)). Moreover, SnKLF10 shared sequence similarities with *H. sapiens* KLF11 (68%) (Protein ID. [AAH74922](#)) and KLF10 (67%) (Protein ID. [ACE87526](#)). The phylogenetic analysis and sequence similarities were consistent (Fig. 2). Phylogenetic analysis demonstrated that all vertebrate KLF10 and KLF11 genes clustered separately together. However, SnKLF10 clustered with *S. purpuratus* KLF10 and *S. kowalevskii* KLF11 belonged to a relatively independent clade (Fig. 2).

The SnC/EBP α (GenBank accession number [KU133955](#)), SnC/EBP γ (GenBank accession number [KU133956](#)), and SnC/EBP ζ (GenBank accession number [KU133957](#)) cDNA fragments with an ORF encoded complete amino acids (aa).

243 Additionally, SnC/EBP α (208–272 aa), SnC/EBP γ (110–174 aa), and SnC/EBP ζ
244 (986–1054 aa) proteins contained a BRLZ domain located at the C-terminus (Fig. S2).
245 The promoter sequences of SnC/EBP α (1154 bp), SnC/EBP γ (1494 bp), and
246 SnC/EBP ζ (823 bp) were cloned (Fig. S1).

247 *Tissue expression analysis of SnKLF10 and SnC/EBPs*

248 As shown in Fig. 3, SnKLF10 and SnC/EBPs were widely expressed in all the
249 examined tissues, and were expressed predominantly in male and female gonads.

250 *Gonadal mRNA expression of SnKLF10 and SnC/EBPs*

251 The mRNA expression levels of the SnKLF10 and SnC/EBPs genes differed in
252 second- and third-instar gonads. In male gonads, the mRNA levels of SnKLF10 ($P <$
253 0.01) and SnC/EBP α ($P < 0.01$) were up-regulated from the second to the third instar,
254 while SnC/EBP γ ($P < 0.05$) and SnC/EBP ζ ($P < 0.05$) were down-regulated (Fig. 4A).
255 In female gonads, the mRNA levels of SnKLF10 ($P < 0.01$), SnC/EBP α ($P > 0.05$),
256 and SnC/EBP ζ ($P > 0.05$) were up-regulated from the second- to the third-instar,
257 while SnC/EBP γ ($P < 0.05$) was down-regulated (Fig. 4B).

258 *Role of SnKLF10 in SnC/EBP gene regulation*

259 We further investigated the effects of SnKLF10 overexpression on the regulation
260 of SnC/EBP α , SnC/EBP γ , and SnC/EBP ζ genes using the Dual Luciferase Reporter
261 Assay. To determine SnKLF10 expression in 293T cells, the expression of EGFP
262 was assessed by fluorescence microscopy. The EGFP gene was strongly expressed in
263 293T cells, suggesting that the pEGFP-KLF10 construct can be used to initiate gene
264 expression in 293T cells (Fig. S4). In addition, SnKLF10 mRNA expression in 293T
265 cell following overexpression of the pEGFP-control or pEGFP-KLF10 vectors was
266 detected by qRT-PCR. The results further showed that SnKLF10 mRNA was only
267 expressed in 293T cells co-transfected with pEGFP-KLF10 (Fig. S5). Therefore,

overexpression of SnKLF10 driven by pEGFP-KLF10 was suitable for the regulation of gene transcription. Here, as shown in Fig. 5, luciferase activity of C/EBP α -1 in pEGFP-KLF10-transfected cells was inhibited 0.81-fold ($P < 0.05$) relative to that in pEGFP-Control-transfected cells. Luciferase activity of C/EBP γ -1 in pEGFP-KLF10-transfected cells was inhibited 0.63-fold ($P < 0.01$) relative to that in pEGFP-Control-transfected cells. Luciferase activity of C/EBP ζ -1 in pEGFP-KLF10 transfected cells was inhibited 0.74-fold ($P < 0.05$) relative to that in pEGFP-Control-transfected cells. Taken together, these results showed that SnKLF10 overexpression led to lower C/EBP α -1, C/EBP γ -1, and C/EBP ζ -1 promoter activity relative to their pEGFP-controls.

Prediction of TFBSs

To further delineate the regulatory mechanism of SnC/EBPs, TFBSs were predicted in SnC/EBP α (location [nt]:-1144–172) (Fig. 6A), SnC/EBP γ (location [nt]:-1396–127) (Fig. 6B), and SnC/EBP ζ (location [nt]: -727–831) (Fig. 6C). The highest core matrix similarities of these TFBSs at SnC/EBP α (Table. S1), SnC/EBP γ (Table. S2), and SnC/EBP ζ (Table. S3) were analyzed by the Genomatix software. Moreover, various TFs, including KLFs, were found to potentially bind the promoter sequences of SnC/EBPs (Fig. 6). The positions and directions of parts of these TFBSs are labeled with arrowheads (Fig. S1).

Promoter analysis

We then sought to determine the core promoter region required for promoter activity within the upstream SnC/EBPs gene sequences. Deletion constructs were transiently transfected into the 293T cell lines, and luciferase activity was detected. C/EBP α -1 to C/EBP α -6 displayed elevated promoter activity relative to pGL3-basic, with C/EBP α -6 having the highest promoter activity (Fig. 7A), leading us to conclude

that a potential core promoter region is located within nucleotides -204 to -17 (Fig. S1A). C/EBP γ -, except for C/EBP γ -7, all displayed lower promoter activity relative to the pGL3-control. Furthermore, the promoter activity gradually increased from C/EBP γ -1 to C/EBP γ -3, and decreased in C/EBP γ -4 (Fig. 7B). These results indicated that two potential core promoter regions are located within nucleotides -1396 to -1179 (core promoter region 1) and nucleotides -742 to -525 (core promoter region 2), respectively (Fig. S1B). C/EBP ζ - displayed higher promoter activity than the pGL3-control, and promoter activity rapidly increased from C/EBP ζ -1 to C/EBP ζ -5, and decreased markedly from C/EBP ζ -6 to C/EBP ζ -7 (Fig. 7C), which indicated the presence of an enhanced promoter region located at nucleotides 164 to 609 (Fig. S1C). Moreover, the core promoter region was located completely in intron 1 (Fig. S1C).

Discussion

Accumulating evidence indicates that KLFs have an important role in fat biology, with five members involved in the positive regulation of adipogenesis and two members involved in its negative regulation (Brey et al. 2009). KLF10, a KLF family member, is involved in cellular differentiation and multiple disease processes (Subramaniam et al. 2010). However, little is known about the function of KLF10 in adipogenesis.

In the present study, a KLF10 gene was identified and characterized from *S. nudus*. Three highly conserved ZnF_C2H2 domains were found in SnKLF10 (Fig. 1B and Fig. 2). Zinc finger domains are common motifs in transcription factors. All members of the KLFs possess three highly conserved zinc finger motifs at the C-terminus ends of the proteins, and a DNA-binding domain (DBD) consisting of three zinc finger motifs. This is important for KLF proteins to bind to similar sites in the promoter regions of downstream target genes (Presnell et al. 2015; Bonnefond et

al. 2011; McConnell and Yang 2010). Furthermore, many putative miRNAs target sites were discovered in the 3'-UTR of the SnKLF10 gene, these are emerging as mediators in the regulation of numerous biological functions, including fat biology, through their ability to bind to a variety of genes (McGregor and Choi 2011; Lages et al. 2012; Stroynowska-Czerwinska et al. 2014). miR-27a is a negative regulator of adipocyte differentiation, and acts by suppressing the transcriptional regulation and expression of PPAR γ (Kim et al. 2010). miR-125a inhibits preadipocyte differentiation by targeting estrogen-related receptor α (ERR α) (Ji et al., 2014). The positions of three miR-27as and five miR-125as were discovered and marked on SnKLF10 (Fig. 1). This result implied that SnKLF10 plays a role in fat biology as a target gene for adipogenic miRNAs. However, further study is needed to confirm the interaction between miRNAs and SnKLF10.

SnKLF10 and SnC/EBPs genes were found to be predominantly expressed in male and female gonads (Fig. 3). Unlike vertebrates, gonadal tissue generates and stores lipids, and plays crucial roles in providing energy for growth, development, and reproduction in *S. nudus* (Arafa et al. 2012; Gaitán-Espitia et al. 2016). The high expression in fat-relevant tissues implies that these genes play important roles in gonadal lipogenesis. Conversely, mRNA expression of the SnKLF10 and SnC/EBPs genes were up- or down-regulated from the second- to the third-instar in male and female gonads (Fig. 4), which further showed that SnKLF10 and SnC/EBPs genes play roles in gonadal adipogenesis.

C/EBPs are key players in the transcriptional networks controlling adipogenesis and lipogenesis (Moseti et al. 2016). KLFs, an emerging new frontier in fat biology, are involved in adipogenesis by regulating fat-related transcription factors (Brey et al. 2009). Overexpression of KLF3 or KLF7 blocks adipocyte differentiation by

inhibiting C/EBP α expression. KLF4 induces C/EBP β expression, and knockdown of KLF4 reduces C/EBP β levels. RNA interference of KLF15 reduces the expression of PPAR γ , KLF15, and C/EBP α , and increases the expression of PPAR γ following a decrease in C/EBP β and C/EBP δ levels, PPAR γ is further able to reduce or elevate C/EBP α levels (Brey et al. 2009; McConnell and Yang 2010). The results of the present study suggest that the expression of SnC/EBPs is associated with SnKLFs. With respect to SnKLF10 and SnC/EBPs, we found that the expression of SnKLF10 followed a very similar pattern to that of SnC/EBP α (Fig. 3A and B). Furthermore, a significant increase in the expression of SnKLF10 from the second- to the third-instar was accompanied by a significant increase in the expression of SnC/EBP α at the same stages of development (Fig. 4), which demonstrates that the expression of SnC/EBP α is associated with SnKLF10 in the control of lipogenesis.

Multiple potential TFBSs were discovered in the 5'-flanking sequences of SnC/EBPs, some TFBSs recognized for KLFs were present in all SnC/EBP genes (Fig. 6), which suggests that SnC/EBPs are downstream target genes of KLFs. In general, the C2H2 zinc finger of KLFs binds to downstream target gene promoters to regulate their transcriptional activation (Zhang et al. 2014). KLF8 binding sites were identified in the C/EBP α promoter by site mutation analysis, and overexpression of KLF8 induced C/EBP α promoter activity, suggesting that KLF8 is an upstream regulator of C/EBP α (Lee et al. 2012). KLF4 binds directly to the C/EBP β promoter, and knockdown of KLF4 downregulates C/EBP β levels (Birsoy et al. 2008). A mechanistic study identified C/EBP as the target gene of KLF2 or KLF3 in lipid metabolism in *Caenorhabditis elegans* (Ling et al. 2017). In the present study, we found that overexpression of SnKLF10 was able to suppress SnC/EBP genes transcription (Fig. 5), The amino-terminal regions of KLFs vary significantly,

allowing them to bind different co-activators and co-repressors, resulting in functional diversity and specificity (McConnell and Yang 2010). KLF10 contains three distinct repression sites in amino-terminal regions of the proteins, and KLF10 acts as a transcriptional repressor (Subramaniam et al. 2010). However, a negative regulatory role of SnKLF10 is not consistent with the above results, which show that the expression of SnC/EBP α is associated with SnKLF10 in the control of lipogenesis (Fig. 3 and 4). In brief, C/EBPs have multiple roles in the regulation of cellular proliferation and differentiation, and in the expression of many inflammatory and immune genes (Van et al. 2015; Moseti et al. 2016). C/EBPs are downstream target genes controlled by multiple TFs (Anand et al. 2013; Jakobsen et al. 2013). Thus, the expression of C/EBPs is regulated by numerous TFs (Siersbæk and Mandrup 2011; Siersbæk et al. 2012). In summary, these results indicate that SnC/EBP genes may be targets of SnKLF10 and SnKLF10 controlled lipogenesis by as a repressor in regulating SnC/EBPs genes transcription. Furthermore, to determine whether SnKLF10 negatively regulates the expression of SnC/EBPs, the use of a SnKLF10 knock-down would be of interest.

Despite the importance of C/EBPs in many biological functions, little is known about their regulation. To identify the possible TBFSs in the C/EBP α (nucleotides -1144–172), C/EBP γ (nucleotides -1396–127), and C/EBP ζ (nucleotides -727–831), we analyzed the sequences using web-based software. Here, we discovered numerous TFBSs for TFs located in the promoter region and in the first exon, such as KLFs, Doublesex, and mab-3 related transcription factors (DMRTs), Nuclear factor NF-kappa-B (NFKB1), Runt-related transcription factor 2 (Runx2), PPAR γ , C/EBPs, TATA-box binding protein (TBP), and Sp1. The presence of these TFBSs in promoter regions indicated that SnC/EBPs are subjected to a high level of transcriptional

control involving many TFs (Huang et al. 2011; Huang et al. 2013). Future research should address the interaction between these TFBSs to provide a better understanding of SnC/EBPs regulation.

To analyze the effect of TFs on the activity of SnC/EBPs transcription, a series of deletion constructs was generated and Luciferase reporter gene assays were performed. The core promoter regions of SnC/EBPs were determined, and found to contain many potential TFBSs for TFs. The core promoter region (nucleotides -204 to -17) of SnC/EBP α contained TFBSs for KLF5 and NF κ B1 (Fig. S1A). The core promoter region 1 (nucleotides -1396 to -1179) of SnC/EBP γ contained TFBSs for C/EBP α , TBP, and activating transcription factor 3 (ATF3), while core promoter region 2 (nucleotides -742 to -525) of SnC/EBP γ contained TFBSs for Dmrt3, KLF1, and KLF16 (Fig. S1B). The core promoter region (nucleotides 164 to 609) of SnC/EBP ζ contained TFBSs for SOX9, KLF16, and KLF14 (Fig. S1C). Thus, these core promoter regions contain one or more regulatory DNA sequence elements, termed core promoter elements or motifs, which are capable of markedly affecting their transcriptional regulation (Zehavi, et al., 2014). In addition, the core promoter region of SnC/EBP ζ was found to be located in an intron (Fig. S1C), which showed that introns are also regulated by many TFs.

In conclusion, a full-length cDNA of SnKLF10 and three cDNA fragments of SnC/EBPs were identified, and three promoter sequences of SnC/EBPs were cloned. SnKLF10 and SnC/EBPs genes were predominantly expressed in male and female gonads, and the mRNA expression levels were up- or down-regulated from the second-to the third-instar in male and female gonads, which implied that SnKLF10 and SnC/EBPs play important roles in gonadal lipogenesis. Furthermore, some TFBSs recognized for KLFs were discovered in 5'-flanking sequences of SnC/EBPs, and

over-expression of SnKLF10 revealed that it acted as a negative regulatory factor in the transcription of all SnC/EBPs genes. Thus, we inferred that SnKLF10 participates in lipogenesis by suppressing the transcription of SnC/EBPs genes. In addition, the core promoter regions of SnC/EBPs were determined, in which we discovered a multitude of TFBSs for TFs; these TFs are potentially key regulators of SnC/EBP transcription. These findings will help to clarify the functions of KLF10 in *S. nudus* and provide insight into the regulatory mechanisms of C/EBPs gene transcription in lipogenesis. Further research is needed to confirm the roles of SnKLF10 and SnC/EBPs in lipogenesis and the mechanisms regulating the transcription of SnC/EBPs genes.

Acknowledgements

We thank Miss Nana Yan and Mr Yongzhen Huang for their technical advice and assistance in experiments. This work was supported by the National Natural Science Foundation of China (Grant No. 31472313 and 31272704).

References

- Anand, S., Hasan, T., and Maytin, E.V. 2013. Mechanism of differentiation-enhanced photodynamic therapy of cancer: Upregulation of coproporphyrinogen oxidase by C/EBP transcription factors. *Mol. Cancer. Ther.* 12(8) 1638–1650.
- Arafa, S., Chouaibi, M., Sadok, S., and El Abed, A. 2012. The influence of season on the gonad index and biochemical composition of the sea urchin *Paracentrotus lividus* from the Gulf of Tunis. *Sci. World.* 2012: 815935.
- Birsoy, K., Chen, Z., and Friedman, J. 2008. Transcriptional regulation of adipogenesis by KLF4. *Cell Metab.* 7(4): 339–347.
- Bonnefond, A., Lomberk, G., Buttar, N., Busiah, K., Vaillant, E., and Lobbens, S. 2011. Disruption of a novel Kruppel-like transcription factor p300-regulated pathway for insulin biosynthesis revealed by studies of the c.-331 INS mutation found in neonatal diabetes mellitus. *J. Biol. Chem.* 286:

- 445 28414–28424.
- 446 Brey, C.W., Nelder, M.P., Hailemariam, T., Gaugler, R., and Hashmi, S. 2009. Krüppel-like family of
447 transcription factors: an emerging new frontier in fat biology. *Int. J. Biol. Sci.* 5: 622–636.
- 448 Briggs, E., and Wessel, G.M.. 2006. In the beginning animal fertilization and sea urchin development.
449 *Dev. Biol.* 300: 15–26.
- 450 Castell, J.D., Kennedy, E.J., Robinson, S.M.C., Parsons, G.J., Blair, T.J., and Gonzalez-Duran, E. 2004.
451 Effect of dietary lipids on fatty acid composition and metabolism in juvenile green sea urchins
452 (*Strongylocentrotus droebachiensis*). *Aquaculture*, 242: 417–435.
- 453 Darlington, G.J., Ross, S.E., and MacDougald, O.A. 1998. The role of C/EBP genes in adipocyte
454 differentiation. *J. Biol. Chem.* 273: 30057–30060.
- 455 Ding, J., Chang, Y.Q., Wang, C.H., and Cao, X.B. 2007. Evaluation of the growth and heterosis of
456 hybrids among three commercially important sea urchins in China: *Strongylocentrotus nudus*, *S.*
457 *intermedius* and *Anthocidaris crassispina*. *Aquaculture*, 272: 273–280.
- 458 Ernst, S.G. 2011. Offerings from an Urchin. *Dev. Biol.* 358: 285–94.
- 459 Farmer, S.R. 2006. Transcriptional control of adipocyte formation. *Cell. Metab.* 4: 263–273.
- 460 Gaitán-Espitia, J.D., Sánchez, R., Bruning, P., and Cárdenas, L. 2016. Functional insights into the testis
461 transcriptome of the edible sea urchin *Loxechinus albus*. *Sci. Rep.* 6: 36516.
- 462 Huang, Y.Z., Zhan Z.Y., Sun, Y.J., Wang, J., Li, M.X., Lan, X.Y., Lei, C.Z., Zhang, C.L., and Chen, H.
463 2013. Comparative analysis of the IGF2 and ZBED6 genes variants and haplotypes reveals
464 significant effect of growth traits in cattle. *Genome*, 56(6): 327–334.
- 465 Huang, Y.Z., Zhan, Z.Y., Sun, Y.J., Wang, J., Li, M.X., Lan, X.Y., Lei, C.Z., Zhang, C.L., Zhang, E.P.,
466 Wang, J.Q., and Chen, H. 2011. Haplotype combination of SREBP-1c gene sequence variants is
467 associated with growth traits in cattle. *Genome*, 54: 507–516.
- 468 Jakobsen J.S., Waage, J., Rapin, N., Bisgaard, H.C., Larsen, F.S., and Porse, B.T. 2013. Temporal
469 mapping of CEBPA and CEBPB binding during liver regeneration reveals dynamic occupancy and
470 specific regulatory codes for homeostatic and cell cycle gene batteries. *Genome. Res.* 23(4):
471 592–603.
- 472 Ji, H. L., Song, C. C., Li, Y. F., He, J. J., Li, Y. L., Zheng, X. L., and Yang, G. S. 2014. miR-125a
473 inhibits porcine preadipocytes differentiation by targeting ERR α . *Mol Cell Biochem.* 395(1–2):
474 155–165.

- 475 Kaczynski, J., Cook, T., and Urrutia, R. 2003. Sp1- and Krüppel-like transcription factors. *Genome Biol.*
476 4: e206.
- 477 Kim, S. Y., Kim A. Y., Lee, H. W., Son, Y. H., Lee, G. Y., Lee, J. W., Lee, Y. S., and Kim, J. B. 2010.
478 miR-27a is a negative regulator of adipocyte differentiation via suppressing PPARgamma
479 expression. *Biochem Biophys Res Commun.* 392(3): 323–328.
- 480 Lages, E., Ipas, H., Guttin, A., Nesr, H., Berger, F., and Issartel, J.P. 2012. MicroRNAs: molecular
481 features and role in cancer. *FRONT BIOSCI.* 17: 2508–2540.
- 482 Laprairie, R.B., Denovan-Wright. E.M., and Wright. J.M. 2016. Divergent evolution of cis-acting
483 peroxisome proliferator-activated receptor elements that control the tandemly-duplicated fatty
484 acid-binding protein genes, *fabp1b.1* and *fabp1b.2* in zebrafish. *Genome*, 59(6): 403–412.
- 485 Lee, H., Kim, H.J., Lee, Y.J., Lee, M.Y., Choi, H., Lee, H., and Kim, J. 2012. Krüppel-Like factor
486 KLF8 plays a critical role in adipocyte differentiation. *PLoS ONE*, 7(12): e52474.
- 487 Ling, J., Brey, C., Schilling, M., Lateef, F., Lopez-Dee, Z.P., Fernandes, K., and Hashmi, S. 2017.
488 Defective lipid metabolism associated with mutation in *klf-2* and *klf-3*: important roles of
489 essential dietary salts in fat storage. *Nutr. Metab.* 14: 22.
- 490 Madsen, M.S., Siersbæk, R., Boergesen, M., Nielsen, R., and Mandrup, S. 2014. Peroxisome
491 proliferator-activated receptor γ and C/EBP α synergistically activate key metabolic adipocyte
492 genes by Mol. Cell. Biolassisted loading. *Mol. Cell. Biol.* 34: 939–954.
- 493 Matsubara, Y., Aoki, M., Endo, T., and Sato, K. 2013. Characterization of the expression profiles of
494 adipogenesis-related factors, ZNF423, KLFs and FGF10, during preadipocyte differentiation and
495 abdominal adipose tissue development in chickens. *Comp. Biochem. Physiol. B.* 165: 189–195.
- 496 McConnell, B.B., and Yang, V.W. 2010. Mammalian Krüppel-Like factors in health and PHYSIOL.
497 REVdiseases. *Physiol. Rev.* 90(4): 1337–1381.
- 498 McGregor, R., and Choi, M. 2011. microRNAs in the regulation of CURR MOL MEDadipogenesis and
499 obesity. *Curr. Mol. Med.* 11(4): 304–316.
- 500 Moseti, D., Regassa, A., and Kim, W.K. 2016. Molecular regulation of adipogenesis and potential
501 anti-adipogenic bioactive molecules. *Int. J. Mol. Sci.* 17(1): 124.
- 502 Nerlov, C. 2007. The C/EBP family of transcription factors: a paradigm for interaction between gene
503 expression and proliferation control. *Trends. Cell. Biol.* 17: 318–324.
- 504 Presnell, J.S., Schnitzler, C.E., and Browne, W.E., 2015. KLF/SP transcription factor family evolution:

- expansion, diversification, and innovation in eukaryotes. *Genome. Biol. Evol.* 7: 2289–2309.
- Reddy, K.C., Dunbar, T.L., Nargund, A.M., Haynes, C.M., and Troemel, E.R. 2016. The *C. elegans* CCAAT-enhancer binding protein gamma is required for surveillance immunity. *Cell. Rep.* 14: 1581-1589.
- Rosen, E.D., and Spiegelman, B.M., 2000a. Molecular regulation of adipogenesis. *Annu. Rev. Cell Dev. Biol.* 16: 145–171.
- Rosen, E.D., Walkey, C.J., and Puigserver, P. 2000b. Transcriptional regulation of adipogenesis. *Genes. Dev.* 14: 1293–1307.
- Siersbæk. R., Mandrup. Sand . 2011. Transcriptional networks controlling adipocyte differentiation. *Cold Spring. Harb. Symp. Quant. Biol.* 76: 247–255.
- Siersbæk. R., Nielsen. R., and Mandrup. S. 2012. Transcriptional networks and chromatin remodeling controlling adipogenesis. *Trends. Endocrin. MeTrends Endocrin. Met.* 23(2): 56-64.
- Stroynowska-Czerwinska, A., Fiszer, A., and Krzyzosiak, W.J. 2014. The panorama of miRNA-mediated mechanisms in mammalian cells. *CELL MOL LIFE SCI.* 71(12): 2253–2270.
- Subramaniam, M., Hawse, J.R., Rajamannan, N.M., Ingle, J.N., and Spelsberg, T.C. 2010. Functional role of KLF10 in multiple disease processes. *BioFactors (Oxford, England)*, 36(1): 8–18.
- Swamynathan, S.K., 2010. Krüppel-like factors: Three fingers in control. *Hum. GenomicsHum. Genom.* 4: 263–270.
- Van der Krieken, S.E., Popeijus, H.E., and Mensink, R.P., Plat, J. 2015. CCAAT/enhancer binding protein β in relation to ER stress, inflammation, and metabolic disturbances. *Biomed. Res. Int.* 324815.
- Wu, Z., and Wang, S. 2012. Role of krüppel-like transcription factors in adipogenesis. *Dev. Biol.* 373: 2235–2243.
- Zehavi, Y., Kuznetsov, O., Ovadia-Shochat, A., and Juven-Gershon, T. 2014. Core promoter functions in the regulation of gene expression of *Drosophila* dorsal target genes. *J. Biol. Chem.* 289: 11993–12004.
- Zhang, Y., Lei, C.Q., Hu, Y.H., Xia, T., Li, M., Zhong, B., and Shu, H.B. 2014. Krüppel-like factor 6 is a co-activator of NF- κ B that mediates p65-dependent transcription of selected downstream genes. *J. Biol. Chem.* 289(18): 12876–12885.

Table 1

Primers for genes confirming and introns cloning.

Primer name	Sequence (5'→3')	Purpose
KLF7F	AATCGGGTTCCTGTCATCAGCG	Confirming SnKLF10 sequence
KLF8R	AACTGCGATGGTCTTTGGTATT	
KLF87F	CGCTGTACCGTACTTGGTTTTACATAG	3' RACE
KLF88F	TGGAGTGTAAGCGGGACTTGCA	
KLF86R	GTGGCATCAACTGCACGGGGCTAG	5' RACE
KLF85R	TTGTGGATTGTTCCGTACTCTTAGCG	
KLF89F	ATAGTCAATGAACCAGTCTAACGT	Confirming full-length SnKLF10 sequence
KLF99R	TTTATTCACCCCTTTACAACATGA	
C/EBP109F	CAATTGACACTTCTTGACGTGATC	Confirming SnC/EBP α sequence
C/EBP108R	AGAACACAAATAGGACCAGAAAAAT	
C/EBP113F	AATACCCCTGGAAAAAGGAGCGT	Confirming SnC/EBP γ sequence
C/EBP31R	CAGCCTTCAAATAACTTCTTCT	
C/EBP193F	CCTGGATTCAACAACCAAGTAAGA	
C/EBP194R	GCTCCATTACATGCTCTTATTC	
C/EBP117F	CATAGTAGTACATGTGCTTCACG	Confirming SnC/EBP ζ sequence
C/EBP198R	GCTATGACATAATCAAATTGTGGC	
C/EBP114F	CTCTACGGCGAGGAGAAACGGTG	SnC/EBP γ intron1 detection
C/EBP116R	CGAGTTGCTTCCTACCTTCCCGC	
C/EBP118F	ACCAAGTCTTCTGTGACCGGAA	SnC/EBP ζ intron1 and intron2 detection
C/EBP120R	TGTCTTTCTTCTGAGTTTATCCT	
3' RACE universal adaptor primer		
UPM	Long: CTAATACGACTCACTATAGGGCAAGCAG TGGTATCAACGCAGAGT	3' RACE
NUP	Short: CTAATACGACTCACTATAGGGC AAGCAGTGGTATCAACGCAGAGT	
5' -RACE adaptor primer		
AAP	GGCCACGCGTCGACTAGTACGGGIIGGGI IGGGIIG	5' RACE
AUAP	GGCCACGCGTCGACTAGTAC	

Table 2

Primers for qRT-PCR.

Primer	Sequence (5'→3')	Product size (bp)
Sn β -actin		
ActinF	GGAACACCCCGTCCTCCTTACT	335
ActinR	CACGCACGATTTCACGCTCA	
HsGAPDH		
GAPDH-F	AGCCACATCGCTCAGACAC	66
GAPDH-R	GCCCAATACGACCAAAATCC	
SnKLF10		
KLF7F	AATCGGGTTCCTGTCATCAGCG	241
KLF85R	TTGTGGATTGTTCCGTACTCTTAGCG	
SnC/EBP α		
C/EBP26F	ATGGATTGCGCTGCTGCTAACT	167
C/EBP125R	TCCGTCCTTCTAGAGCTAGATACG	
SnC/EBP γ		
C/EBP29F	GGTCCCAGACTCTACAGGCATCG	160
C/EBP138R	GCGACTCTTCCGCACAGCCTC	
SnC/EBP ζ		
C/EBP118F	CATCGCCGCATTCTACCTG	271
C/EBP120R	TGGATGTTTTGAATGCCCTATG	

9
10
11

Table 3

Primers for Genome walking

Primer name	Sequence (5'→ 3')	Location (nt)	Product size (bp)
C/EBPα			
SP3-C/EBP126R	ATTGTTGCTTCTCTTCGCAATCGTA	250 ~ 275	1429
SP2-C/EBP125R	TCCGTCCTTCTAGAGCTAGATACG	368 ~ 393	1547
SP1-C/EBP124R	AGTTGGGGCAGGCGTTTGGATGACC	480 ~ 505	1659
C/EBPγ			
SP3-C/EBP129R	TTAGCACCGTTTCTCCTCGCCGTAG	126 ~ 151	1645
SP2-C/EBP128R	CACATATCTAAACGTGCTTATACA	200 ~ 224	1718
SP1-C/EBP127R	ACAATCTCCATAGAAACCAGTAACC	506 ~ 531	2025
C/EBPζ			
SP3-C/EBP133R	TGCCGTCTTCATTGACATTGTGCGTGC	266 ~ 293	1116
SP2-C/EBP131R	CAATCAAAGAGTGTGTGAGAATCAA	464 ~ 489	1312
SP3-C/EBP134R	ATGGATACTGCTCACGAGACCAGATAA	804 ~ 831	1654
Genome Walking universal primers			
AP1	Nnknown	For cloning the 5'-flanking sequences of SnC/EBPs	
AP2	Nnknown		
AP3	Nnknown		
AP4	Nnknown		

12
13

Table 4

Primers for overexpression, protein expression and promoter deletion fragments vectors construction.

Primer name	Sequence (5'→ 3')	Location (nt)	Product size (bp)
KLF10			
EGFP-KF	F: TCAG ctcgagATAGTCAATGAACCAGTCTAACGTAA	45 ~ 71	2073
EGFP-KR	R: TCAG ggtaccTAGGATTAGTAACTAGACCAGTGTTA	2092~2118	
C/EBPα			
pGL3-α1	F: TCAG ggtaccTGACTATAGAATACTCAAGCTATG	-1144 ~ -1120	1316
pGL3-α2	F: TCAG ggtaccGTAAACGAAGCTGGAGGTAAGACTA	-956 ~ -932	1128
pGL3-α3	F: TCAG ggtaccGTGAATACATGTAGAATACATCAT	-767 ~ -744	940
pGL3-α4	F: TCAG ggtaccTTTATTAACCATATTGTTTAAATT	-580 ~ -556	752
pGL3-α5	F: TCAG ggtaccCACATTCTAAATATAGCGTCATTG	-392 ~ -368	564
pGL3-α6	F: TCAG ggtaccTAAAATTCATCGTATATATATCGG	-204 ~ -180	376
pGL3-α7	F: TCAG ggtaccCACGTAGTTGCATATTGTCCAGTA	-16 ~ 8	188
pGL3-α8	R: TCAG ctcgagCAAAATTCGCAGATAGACCTCACGG	147 ~ 172	-
C/EBPγ			
pGL3-γ1	F: TCAG ggtaccGCTCGTCGCTCGTCTGTACGCCAT	-1396 ~ -1371	1523
pGL3-γ2	F: TCAG ggtaccATGTTGTACAAAGCTTTCATTCT	-1178 ~ -1154	1305
pGL3-γ3	F: TCAG ggtaccGGTACACATTTTTTTTGGTTTTAA	-960 ~ -936	1087
pGL3-γ4	F: TCAG ggtaccAGACACACAGAGAGGCGATCGAGG	-742 ~ -718	869
pGL3-γ5	F: TCAG ggtaccTGGGATGAAACTTGAAGGGGGAG	-524 ~ -500	651
pGL3-γ6	F: TCAG ggtaccACTCATTTTCTACATTCTTTCTC	-306 ~ -282	433
pGL3-γ7	F: TCAG ggtaccCATTCAGTGTGTGTGTGTGTGTGT	-88 ~ -64	215
pGL3-γ8	R: TCAG ctcgagGAGCAACGGTACGTGTCAAGTTCAG	103 ~ 127	-
C/EBPζ			
pGL3-ζ1	F: TCAG ggtaccCTTCTCGTACCTCAAACGGGACACT	-727 ~ -702	1558
pGL3-ζ2	F: TCAG ggtaccTCTCGATCTGGTAATGAAAGAAAA	-504 ~ -480	1335
pGL3-ζ3	F: TCAG ggtaccTTTTTTTATTGAAGATGATTTCCA	-282 ~ -258	1113
pGL3-ζ4	F: TCAG ggtaccCCCGGAAACTTTCGGCATGATT	-58 ~ -35	890
pGL3-ζ5	F: TCAG ggtaccATTATTTCATTTATTATGTGGCTT	164 ~ 188	667
pGL3-ζ6	F: TCAG ggtaccATCATGAATATATTTTACTGC	387 ~ 411	444
pGL3-ζ7	F: TCAG ggtaccTAACTGTTTATAAAACGGGGGCC	610 ~ 634	221
pGL3-ζ8	R: TCAG ctcgagATGGATACTGCTCACGAGACCAGATAA	804 ~ 831	-

Note: F: forward primer; R: reverse primer; The primer used for constructing vectors pEGFP-KLF10

17 and pGL3-. pEGFP-KLF10 contains *XhoI* (forward primer) and *KpnI* (reverse primer) restriction sites,
18 pGL3- contains *KpnI* (forward primer) and *XhoI* (reverse primer) restriction sites. The nucleotides of
19 restriction sites indicated by the lower-case letters. The protective base “TCAG” were added in order to
20 express in the 5' terminal of primer had also deliberately added black bold font. The location and size
21 of each 5'-deletion fragment was indicated to the left of each bar relative to TSS.

Draft

1 **Figure legends**

2 **Fig. 1. SnKLF10 sequence.** A: Nucleotide and deduced amino acid sequences of the SnKLF10 gene.
3 The nucleotide and amino acid sequences are numbered on the left; “1” marked the TSS. The start
4 codon (ATG) is boxed and the stop codon (TGA) is marked with an asterisk. On the deduced amino
5 acid sequences, three ZnF_C2H2 domains (371–395, 401–425, and 431–453aa) are indicated by a gray
6 background. B: Schematic drawing of predicted SnKLF10 protein domains. The “491 aa” indicates the
7 numbers of amino acids. The five potential miR-125as (location [nt]: 1735–1742, 2870–2877,
8 2935–2942, 3205–3312, and 3391–3398 bp) are indicated with a wavy line. The three potential
9 miR-27as (location [nt]: 1846–1853, 1938–1945, and 2358–2365 bp) were indicated with a double line.

10 **Fig. 2. Phylogenetic relationships of KLF10s and KLF11s.** Amino acid sequences of known or
11 predicted KLF10s and KLF11s in GenBank, were aligned using ClustalW and used to construct a
12 phylogenetic tree by the neighbor-joining algorithm in the MEGA 5.0 program. The bar indicates the
13 distance. SnKLF10 is marked with a red triangle.

14 **Fig. 3. Tissue distribution of KLF10 (A), C/EBPα (B), C/EBPγ (C), and C/EBPζ (D) mRNA in**
15 *Strongylocentrotus nudus*. β-actin was used as an internal reference gene. The five examined tissues
16 were gut, muscle, tube feet, and male and female gonads. mRNA expression in tube feet was used for
17 normalization. Error bars indicate standard deviation (n=3).

18 **Fig. 4. SnKLF10, SnC/EBPα, SnC/EBPγ, and SnC/EBPζ expression profiles in different**
19 **developmental stages of male (A) and female (B) gonads.** β-actin was used as an internal reference
20 gene. Asterisk (*) indicates a significant difference between different gender groups or different age
21 groups (*P < 0.05, **P < 0.01). Error bars indicate standard deviation (n=3).

22 **Fig. 5. Luciferase activity of SnC/EBPα-1, SnC/EBPγ-1, and SnC/EBPζ-1 in 293T cell lines**
23 **overexpressing pEGFP-KLF10.** Light blue solid bar shows luciferase activities from cell lines
24 co-transfected with pEGFP (control). Light red solid bar represents luciferase activities from cell lines
25 transfected with pEGFP-KLF10. Values represent the mean ± SE of three replicates. Asterisk (*)
26 indicates a significant difference between the experimental group and the control group (*P < 0.05, **P
27 < 0.01).

28 **Fig. 6. TFBSs predicted in the full-length promoter vector sequence of SnC/EBPα-1 (location [nt]:**
29 **-1144–172) (A), SnC/EBPγ-1 (location [nt]: -1396–127) (B), and SnC/EBPζ-1 (location [nt]:**
30 **-727–831) (C).** Many putative TFBSs for TFs were identified. The TFBSs for KLFs are marked by red

arrow. The initial position of all deletion vectors is indicated.

Fig. 7. Recombinant vectors used to assay the upstream sequence activity of SnC/EBPs.

Luciferase activity of the SnC/EBP α (A), SnC/EBP γ (B), and SnC/EBP ζ (C) upstream deletion constructs in 293T cells. The location and size of each 5'-deletion fragment is indicated to the left of each bar relative to the predicted TSS. Note that the SnC/EBP ζ (C) upstream deletion constructs contained intron 1. Values represent the mean $\pm$ SE of three duplication.

Fig. 1.

A

```

1 GTGCATGCAAGGACAGTACACACCGCTCCTTGCTCGAATCTTTTATAGTCAATGAACCA
61 GTCTAACGTAAACAACCTGCGAATGATAACGTGCATCTGCCTAAGGATTTTCTACTGTGT
1 M E F V L P S P P S T P P T
121 TTTACCGTGTTTTTTGTAATGGAATTCGTTTACCATCGCCACCATCCACACCACCAACT
15 L L I S E T S L S V K N I P T I T A T A
181 CTGCTAATTTGGAACGAGTCTATCCGTGAAAAACATCCCTACTATCACCGCTACTGCT
35 G C F T M S P R P I E K S D F D A V Q T
241 GGTGCTTCACCATGAGTCCTCGCCCAATCGAGAAATCTGACTTCGACGCCGTTTCAGACC
55 L L S M R S V P S Q V T I K R S D S P V
301 CTCTTGTCATGCGGAGCGTTCCTTCTCAAGTCACCATCAAACGGAGCGATTCTCCCGTG
75 P S S S P V P T S F N A P L S P V S I D
361 CCGTCGTCATCGCCGTTCCAACCAGCTTCAACGCCCGCTCTCACCGGTCTCTATTGAC
95 D E N S Q H H Q P A E M M E F T P A R M
421 GACGAGAACAGTCAGCATCATCAACCTGCAGAGATGATGGAATTCACACCGGCCAGAATG
115 R G M D T P P L T P P P S K P T V V P G
481 AGGGGGATGGACACACCGCCTCTCACACCCCCACCTAGCAAACCAACAGTCGTGCCCGGT
135 M P M S H T F S M P Q A S S A I V N T Q
541 ATGCCAATGAGCCACACATTCTCTATGCCACAAGCCAGTAGCGCCATCGTCAACACGCAG
155 R I A S C N L V S I T P S I M A S K E I
601 AGGATAGCTTCTTGCAATCTGGTCAGCATTACCCCAAGCATCATGGCATCTAAGGAGATC
175 P S K W T R M E T T P S Q S V Q S R L S
661 CCGAGTAAATGGACAAGGATGGAAACCACTCCGTCACAGTCTGTTCAATCGCGTCTCTCT
195 P V P Q T Q T T S F Y N R V P V I S E T
721 CCCGTCCCACAAACGCAAACGACGTCTTTTACAATCGGGTTCCTGTCATCAGCGAGACG
215 R R D H M S V P A A S A G S S P S P V Q
781 AGGAGAGATCACATGTCTGTGCCAGCGCGTCGGCGGGATCGAGTCCTAGCCCCGTGCAG
235 L M P Q M Q E S R N S V Q Q P C E P Q P
841 TTGATGCCACAGATGCAAGAATCCAGAACTCTGTCCAGCAACCTTGCAGCCCCAACCA
255 K Y I A V N G S F Q I P V S C Q N G P S
901 AAATACATTGCCGTCAATGGGTCTTTCCAGATCCCAGTTTCGTGCCAAAATGGACCTTCT

```

69 275 A M L A K S T E Q S T N V Q T V S F I Q
70 961 GCAATGCTCGCTAAGAGTACGGAACAATCCACAAACGTGCAAACGGTATCTTTTATCCAG
71 295 I P Q Q T T Q K S S S N G Q S D K L K T
72 1021 ATTCCTCAGCAGACCACTCAAAAGTCAAGCAGCAATGGACAGTCTGATAAACTCAAAACG
73 315 V V M A V P A N V M V V V N G M A K S E
74 1081 GTCGTGATGGCAGTCCCGGCAATGTCATGGTTGTTGTGAACGGAATGGCCAAGTCGGAA
75 335 G Q K L C P L A P A P S R C T S P M S D
76 1141 GGACAGAAGCTGTGCCCTCTTGCCCCTGCACCGTCTCGGTGCACATCACCAATGTCTGGAC
77 355 K A P M S P A A T E F S R R R N H I C T
78 1201 AAGGCACCCATGTCCCCTGCAGCTACAGAGTTCTCAAGAAGGAGAAATCACATCTGTACC
79 375 F P N C G K T Y F K S S H L K A H V R T
80 1261 TTCCCGAATTGTGGCAAGACCTACTTTAAGAGCTCCCATCTCAAGGCCACGTCAGAACT
81 395 H T G E K P F H C T W E G C D K R F A R
82 1321 CATACAGGAGAGAAAACCGTTCCACTGTACATGGGAAGGCTGTGACAAGCGATTGCGCCGA
83 415 S D E L S R H K R T H T G E K K F L C P
84 1381 TCTGACGAACCTCTCAAGACACAAGCGTACTCACACGGGCGAGAAGAAGTTCCTCTGTCCC
85 435 M C D R R F M R S D H L T K H A R R H M
86 1441 ATGTGTGATCGCCGCTTCATGAGATCGGACCACTTAACCAAGCATGCCCGTCGTACATG
87 455 A A K K V P N W Q L E V S K L S T M A A
88 1501 GCAGCCAAGAAGGTTCCCTAACTGGCAATTGGAGGTGAGCAAACTCTCGACGATGGCAGCC
89 475 E N R Q Q P Q Q M V P M I I T S S *
90 1561 GAGAATCGACAACAACCTCAGCAAATGGTGCCCATGATCATCACCAGCTCATGAACACTA
91 1621 CATCAGAGAGACTTATGAACCGAGCAATGACTCAGGATATCCTCCAGCCGTTGCCGATGT
92 1681 GTGCTGGCAACGATCATGAATGTCAAATACATCTTCTCTTCAGATGTGACTGAATCGCGG
93 1741 GGAGATTGGCATGGATTGACAAGCAGGGCTAACTTTGAACTCTAAGCATGCAATGAGACT
94 1801 GACTGAAAGATATTTATATAGAGAGGTGTATATTAACATTATATCACAGTGGAAATATCA
95 1861 TTCCTGCACGTTTACTAGAAAAGAACAACCTTTTGTGTCTTTTTCTAGTCAAATGTGCTT
96 1921 TAGACATACAAACATGCATTTTGAAATACTTGTGATTTGTCTTCATTTCAAGTAGACAG
97 1981 TAGCAAGATATACATGTAGTTAGATTGATAAACAACATCCTACGCATAATGTACACTG
98 2041 TACATGTTGCAGTGCCTCTGAATCTTTTAGGTGTTCTGAAGCAAAAGTAATGTAACACTG
99 2101 GTCTAGTTACTAATCCTATTGTAAGGAAAGAATCGAGGAAAAGCAGACTTATCTTTTCAG
100 2161 ATGTGATTAGGGCTTGTGTATACAAAGCACACGTACGCAAAGTCCGGTGAGCAAACCTAAC
101 2221 TTTATTATTTATGCTGCTATATATTTTATACGACGACTTTGAAATATTAAGTAAACTGTG
102 2281 CTTTTATGTCGTAAAGAATTTGATTTCTGTGCTGGTATCACGAAGACAAACAAATCTCGTG
103 2341 TAGATAAGTATTTATTTAAATGTACAACAAGATATTTCACTTCATAAATTTCAAAGTATC
104 2401 AAATATGAAAACGGAATGCTTGGACCCTTTTTTGCCAAATTATTTGTGCCGCTGTACCG
105 2461 TACTTGGTTTTACATAGCAACCGTTTTATTGTCAATTTGTCTGATCTTAGCATGCTACTAC
106 2521 TTTGAAAAATGACAAAGCTAAAGCATTGAACAATAACAACCTTTTAATTTTGTAAATTTAT
107 2581 TTGAATCCAGTAGAGATAATCTCTAAAAAAGGATTGAGAAAGTTGCACCGAAACCGAACC
108 2641 ATTTTTTTCTTCAATGATTAAGGGATTTAATGTGCTGAAGAAAAAAGAAAAA
109 2701 AAACGAATTCAAAACGATTTTTTTTTTTTTTCTTCTTCTGTAACAACCTAATATTTTT

110 2761 GCCATTAAAAAAATTTGTTTGTACAAAATTTTCGGGTAAATTGGAGTGTAAAGCGGGACT
111 2821 CATGAAAATACTTTTTTTTTTGGTCTCGATTTTCTTGATATCACTGTTGACTCGGGGATC
112 2881 AAAATGAACTGAAAGGTTGATTATAATATTATATATTAGGTAACAAATGTATGCTTAGG
113 2941 GTAAATTATTTATGAATAATTTAAATAAATGTGTGTGTTTAAATACCAAAGACCATCGCAG
114 3001 TTCTGATTTGAAAGCATGAAAAGAAATATGAGACTAAGTCTCTATTTTTTAAAATTAGAA
115 3061 ATGTATACCCAACCAACAATCTTGCGTTAAAGATATTGCCAATTTCCATATTTTTTTCTG
116 3121 TGCTTCTCTGATGCAAATGTACCACACAAATGCATTTTATATGTATTTGTTTTCTTTCC
117 3181 TAATGTACTACTTAAATCTAATCACATACCTGTGAATCCTTCGGTCACAATTTGTAGGCA
118 3241 TTTCTCTCTGTACCTTTATTTTCCGTAACCTTATTTCCAACAAAACCTTTGTCGGTATTACCA
119 3301 GTGGTGTCCGTAATGTACAGTAATGATACAGTATCATATACCACCAATATGTTAACATA
120 3361 CTTTTATACAAAGATTTTCAAATAACATTACTCATCTTGTTACATTTCTATGCTTCGTTT
121 3421 ATTTTATGACGTCAGCAACACTTCTTTTTCTTTCAAGTGAGTTTCATGTTGTAAAGGGGT
122 3481 GAATAAATCTTCATTCTCTTTGAATTGTTAAAAAACATAAAAAAA
123

B

124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156

157 **Fig.2**

158

Fig. 3.

160

Fig. 4.

162

163

164

Fig. 5.

Fig. 6.

Fig. 7.