Theriogenology xxx (2014) 1-12



Contents lists available at ScienceDirect

# Theriogenology



journal homepage: www.theriojournal.com

# Molecular characterization, sexually dimorphic expression, and functional analysis of 3'-untranslated region of *vasa* gene in half-smooth tongue sole (*Cynoglossus semilaevis*)

Jinqiang Huang <sup>a,b,c</sup>, Songlin Chen <sup>b,\*</sup>, Yang Liu <sup>d</sup>, Changwei Shao <sup>b</sup>, Fan Lin <sup>b</sup>, Na Wang <sup>b</sup>, Qiaomu Hu <sup>b</sup>

<sup>a</sup> College of Marine Life Sciences, Ocean University of China, Qingdao, China

<sup>b</sup> Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao, China

<sup>c</sup> College of Animal Science and Technology, Gansu Agricultural University, Lanzhou, China

<sup>d</sup> College of Fisheries and Life Science, Dalian Ocean University, Dalian, China

### ARTICLE INFO

Article history: Received 15 November 2013 Received in revised form 23 March 2014 Accepted 25 March 2014

Keywords: Vasa Flatfish Primordial germ cells Sex differentiation Gonadal development

### ABSTRACT

Vasa is a highly conserved ATP-dependent RNA helicase expressed mainly in germ cells. The vasa gene plays a crucial role in the development of germ cell lineage and has become an excellent molecular marker in identifying germ cells in teleosts. However, little is known about the structure and function of the vasa gene in flatfish. In this study, the vasa gene (Csvasa) was isolated and characterized in half-smooth tongue sole (Cynoglossus semilaevis), an economically important flatfish in China. In the obtained 6425-bp genomic sequence, 23 exons and 22 introns were identified. The Csvasa gene encodes a 663-amino acid protein, including highly conserved domains of the DEAD-box protein family. The amino acid sequence also shared a high homology with other teleosts. Csvasa expression was mainly restricted to the gonads, with little or no expression in other tissues. Real-time quantitative polymerase chain reaction analysis revealed that Csvasa expression levels decreased during embryonic and early developmental stages and increased with the primordial germ cell proliferation. A typical sexually dimorphic expression pattern of Csvasa was observed during early development and sex differentiation, suggesting that the Csvasa gene might play a differential role in the proliferation and differentiation of male and female primordial germ cells (PGCs). Csvasa mRNA expression levels in neomales were significantly lower than those in normal males and females, indicating that the Csvasa gene might be implicated in germ cell development after sex reversal by temperature treatment. In addition, medaka (Oryzias latipes) PGCs could be transiently labeled by microinjection of synthesized mRNA containing the green fluorescence protein gene and 3'-untranslated region of Csvasa, which confirmed that the Csvasa gene has the potential to be used as a visual molecular marker of germ cells and laid a foundation for manipulation of PGCs in tongue sole reproduction.

© 2014 Elsevier Inc. All rights reserved.

### 1. Introduction

E-mail address: chensl@ysfri.ac.cn (S. Chen).

In sexually reproducing organisms, germ cells, derived from primordial germ cells (PGCs), play important roles in gonadal differentiation and transmission of genetic information through the generations [1,2]. Primordial germ cells differentiate from blastomeres during the early stages of

<sup>\*</sup> Corresponding author. Tel.: +86 532 85844606; fax: +86 532 85811514.

<sup>0093-691</sup>X/\$ – see front matter © 2014 Elsevier Inc. All rights reserved. http://dx.doi.org/10.1016/j.theriogenology.2014.03.017

embryogenesis and migrate into the presumptive gonads, giving rise to sperm in males and eggs in females. Several molecular markers have recently been used to study PGCs in Metazoa, including *vasa*, *dnd*, and *nanos* [3]. Among these, the *vasa* gene was the first molecular marker for PGCs in teleosts, and researchers have developed a great interest in it [4].

The vasa gene encodes an ATP-dependent RNA helicase belonging to the DEAD (Asp-Glu-Ala-Asp)-box family, which was originally reported in fruit fly (*Drosophila melanogaster*) and considered as the maternal-effect gene required for abdominal segment formation and the proliferation and differentiation of germ cells [5]. Many studies revealed that the vasa gene was specifically expressed in germline cells and played important roles in germline development [6–17]. Sequence analysis of the vasa gene demonstrated a high conservation of gene structure during evolution. The vasa gene is therefore an excellent germ cell marker and has been used extensively for researching the formation, proliferation, and differentiation of PGCs in teleosts [12–17].

Recent years have seen the development of techniques to trace live PGCs in teleosts using microinjection of a fusion gene containing regulatory regions of *vasa* gene and green fluorescence protein (GFP) gene [17–22]. The 3'-untranslated region (3'UTR) of *vasa* has been verified to be sufficient for regulation of specific expression in the PGCs of zebrafish (*Danio rerio*) [18], medaka (*Oryzias latipes*) [21,22], Atlantic salmon (*Salmo salar*) [17], and a range of teleost species [20]. Migration patterns of PGCs in embryos and proliferation of PGCs during development can now be analyzed by tracing live PGCs visually, and labeled PGCs from some teleost species have been successfully isolated for cryopreservation and surrogate broodstock of endangered and commercial species [23,24].

Half-smooth tongue sole (Cynoglossus semilaevis) is one of the most valuable marine flatfish in China, which possesses a ZZ/ZW genetic sex determination system. Females of the species grow two to three times bigger than males [25], the phenomenon of natural sex reversal in genetic female tongue sole was observed [26], and it was also found that physiological sex of juvenile tongue sole was subject to reversal by an influence of temperature [27]. Artificial gynogenesis was carried out to obtain the super-females and some progress has been made [28]. Understanding the sex determination mechanism will accelerate the development of monosex female production. It was reported that germ cells played a crucial role in female sex determination in zebrafish [2], medaka [29], and three-spined stickleback (Gasterosteus aculeatus) [30], but in loach (Misgurnus anguillicaudatus) and goldfish (Carassius auratus), germ cells were not the primary factor for sex determination [31,32]. The role of germ cells in gonadal sex differentiation remains unclear, partly because of limited knowledge about the expression and functional analysis of germ cell makers in flatfish species. In addition, the tongue sole broodstocks cannot guarantee eggs production under artificial culture conditions. Primordial germ cell transplantation and surrogate broodstock might be employed to improve reproduction, yet no molecular marker for PGCs has been reported in this species. Therefore, information about the tongue sole vasa gene (Csvasa), a germ cell and PGC molecular marker, would not only broaden our understanding of sex determination but also provide the basic knowledge of PGC labeling, cryopreservation, transplantation, and surrogate broodstock in this species.

In this study, we cloned and characterized cDNA and genomic sequence of *Csvasa*, studied *Csvasa* mRNA expression levels in different tissues and embryonic stages by real-time quantitative PCR (qPCR), and sex-linked differential expression of *Csvasa* was also detected during early developmental and sex differentiation stages. The function of *Csvasa* 3'UTR was also analyzed by micro-injecting synthesized mRNA into fertilized medaka eggs.

### 2. Materials and methods

#### 2.1. Fish and samples

The half-smooth tongue sole used in this study were obtained from Haiyang High-Tech Experimental Base (Haiyang, Shandong Province, China) during the spawning season. Tissue samples (which included heart, liver, gill, skin, blood, kidney, intestine, brain, spleen, muscle, pituitary, and gonads) were collected from 1-year-old tongue sole, and immediately dipped into liquid nitrogen and stored at -80 °C until further analysis. Twenty body trunks (4–66 days post-hatching, dph) and ten gonads (80-150 dph) from different developmental stages of tongue soles were collected and stored in liquid nitrogen. To determine the genetic sex, the fins of these samples were placed in 100% ethanol for DNA extraction. The embryos that were cultured about 40 hours between fertilization and hatching under the temperature of 23  $\pm$  0.5 °C and larvae of different developmental stages were placed in liquid nitrogen and stored at -80 °C until RNA extraction. The sex reversal induction of tongue sole fry by temperature treatment was carried out as previously described [27]. Briefly, the fry at 28 dph were cultured in tanks with a temperature of 28 °C until 100 days and then grown into adults under a natural water temperature.

The work described in this article have been carried out in compliance with the EU Directive 2010/63/EU for animal experiments; http://ec.Europa.eu/environment/chemicals/ lab\_animals/legislation\_en.hem; Uniform Requirements for manuscripts submitted to Biomedical journals http:// www.icmje.org.

### 2.2. Physiological sex and genetic sex identification

The physiological sex of each tongue sole was identified using gonad-specific expression genes (*dmrt1* and *cyp19a1a*) [27]. Genomic DNA of each sample was extracted using the traditional phenol chloroform method [28]. A pair of sex-linked SSR primers (CseF-SSR1 and SChen-1) was used for genetic sex identification [28]. The genetic male produced a 206-bp band and the genetic female produced two DNA bands of 206 and 218 bp. Fish with female genetic sex but male phenotypic sex were regarded as neomales.

# 2.3. Cloning full-length cDNA and genomic DNA sequences of Csvasa gene

Total RNA was isolated from the gonads of adult halfsmooth tongue sole using TRIzol reagent (Invitrogen, USA)

according to the manufacture's instructions. First-strand cDNA was synthesized from 1 µg of total RNA, using transcript first-strand cDNA synthesis kit (Takara, Dalian, China). A 1097-bp cDNA fragment of the Csvasa gene was amplified by reverse transcription polymerase chain reaction (RT-PCR) using primers Vasa-mid-A and Vasa-mid-S (Table 1), which were designed according to the partial sequence of Csvasa from whole-genome and transcriptome sequencing [33]. Then rapid amplification of cDNA ends (RACE)-specific primers Vasa-5'GSP and Vasa-3'GSP(Table 1) were designed based on the partial cDNA fragment. Using a BD SMART RACE cDNA amplification kit (Clontech, Mountain View, CA, USA), 5'- and 3'-RACE were employed to isolate a full-length cDNA sequence according to the manufacturer's protocol. The PCR conditions were as follows: five cycles at 94 °C for 30 seconds and 72 °C for 3 minutes, five cycles at 94 °C for 30 seconds, 70 °C for 30 seconds and 72 °C for 3 minutes, then 27 cycles at 94 °C for 30 seconds, 68 °C for 30 seconds, and 72 °C for 3 minutes. The amplification products were separated using 1.0% agarose gel using the DL2000 marker. Purified fragments were cloned into pMD18-T simple vector (Takara) for sequencing. Sequences were assembled using the Vector NTI software package (Invitrogen). The Csvasa cDNA sequence was analyzed by bioinformatics, and three pair of primers (Vasa-DNA-S1, Vasa-DNA-A1; Vasa-DNA-S2, Vasa-DNA-A2; and Vasa-DNA-S3, Vasa-DNA-A3) (Table 1) were designed to amplify the genomic sequence. Polymerase chain reactions were performed as follows: initial step at 94 °C for 3 minutes, followed by 35 cycles for 30 seconds at 94 °C, 30 seconds at 60 °C, and 3 minutes at 72 °C, and a final extension at 72 °C for 10 minutes. DNA fragments were ligated into pMD18-T vector and sequenced.

### 2.4. Sequence analysis and alignment

Homology searches were performed using an online Website (http://www.ncbi.nlm.nih.gov/BLAST/). Exon-intron boundaries were determined by comparing *Csvasa* cDNA and

#### Table 1

Primers and their sequences used in this study.

DNA sequence. A phylogenetic tree was constructed with Mega 5.2 software using the neighbor-joining method with a bootstrap analysis of 1000 replications [34]. The deduced amino acid sequences were aligned using the AlignX program in Vector NTI software. Sequences used for alignment were as follows: medaka (BAB61047), zebrafish (AAI29276), human (*Homo sapiens*) (AAF72705), mouse (*Mus musculus*) (EDL18409), fruit fly (AAF53438), frog (*Xenopus laevis*) (AAC03114), and chicken (*Gallus gallus*) (BAB12337).

### 2.5. Real-time quantitative PCR of Csvasa gene

Owing to the difficulty of dissecting the gonads from fish at 4 to 66 dph because of their small size, each body trunk was divided into several parts. The putative gonad parts (4-66 dph) and gonads were identified and confirmed using gonad-specific expression genes (dmrt1 and cyp19a1a), respectively [27]. Thirty embryos or three to six gonads (gonad parts) of the same gender and age were pooled as one sample for RNA isolation. Three samples were collected from each stage for replicate experiments. Total RNA was extracted from each sample of different developmental stages and tissues of six healthy adults and reversely transcribed as described above. cDNA was used as the template for qPCRs. The expression of each sample was performed in triplicate using gPCR. Quantitative PCR was carried out on an ABI PRISM 7500 real-time PCR system (Applied Biosystems, Foster City, CA, USA) in 20 µL reactions using SYBR Premix Ex Tag (Takara). Reactions were performed using specific primers (Vasa-RT-A, Vasa-RT-S; and Actin-RT-A, Actin-RT-S) (Table 1) based on the manufacturer's manual, and  $\beta$ -actin was used as the internal reference [35]. Cycling parameters were as follows: 95 °C for 10 seconds, 40 cycles at 95 °C for 5 seconds, 60 °C for 34 seconds, followed by disassociation curve analysis to check amplification specificity. Csvasa expression was analyzed using the relative quantification method and 7500 system SDS software (Applied Biosystems).

Primer	Primer sequences (5'-3')	Usage
CseF-SSR1	GAGGCCGACAGGATCGTAC	Genetic sex identification
SChen-1	TACGACGTACTCCGGTGGTTTT	Genetic sex identification
P450a-A	CAGGTAGGAGGTTGCTGGGTA	Gonad identification
P450a-S	CAGGAGGAAGAACTTGGGATTT	Gonad identification
Dmrt1-A	GGTGAGGATGTGACCCAGTGT	Gonad identification
Dmrt1-S	ACGGGCTGAAATCGCAAG	Gonad identification
Vasa-mid-A	TGACGACATACTGGTGAACG	Partial fragment clone
Vasa-mid-S	TGGAAATCAGAGGGCGAG	Partial fragment clone
Vasa-5' GSP	GAGTCGCACAGGGCAGCCTCCTCAAAT	5' RACE PCR
Vasa-3' GSP	GCATTGATGAGTATGTTCACCGCATTGGG	3' RACE PCR
Vasa-DNA-A1	GAATATGGATGACTGGGAAGAGACG	DNA fragment PCR
Vasa-DNA-S1	ACACGGCTCGTGACTGTGAGCTA	DNA fragment PCR
Vasa-DNA-A2	CGTCTCTTCCCAGTCATCCATATTC	DNA fragment PCR
Vasa-DNA-S2	GAATATGGATGACTGGGAAGAGACG	DNA fragment PCR
Vasa-DNA-A3	CATCCATTTATTGGTCCTTATCGTCA	DNA fragment PCR
Vasa-DNA-S3	CCAAAAATGTCAGCAGATCTGGATATG	DNA fragment PCR
Vasa-RT-A	AGGGCGAGACAGATGACCG	qPCR
Vasa-RT-S	CCCAGGCAGCATTGATGAGTA	qPCR
Actin-RT-A	GCTGTGCCCTGTA	qPCR
Actin-RT-S	GAGTAGCCACGCTCTGTC	qPCR
Vasa3U-A	GGGGTACCGCTTGGATTACATCCATTTATTGG	3' UTR clone
Vasa3U-S	CCGCTCGAGGAGGAAGAAGAATGGGAGTAG	3' UTR clone

Each relative expression result was calculated using the  $2^{-\bigtriangleup \bigtriangleup Ct}$  method.

# 2.6. Preparation of GFP-Csvasa 3'UTR mRNA and microinjection

The 3'UTR of Csvasa mRNA was amplified using primers (Vasa3U-A, Vasa3U-S)(Table 1) digested with XhoI and KpnI, and cloned into the corresponding site of GFP-Drnos 3'UTR (supplied by Doctor Lin [22]), and then designed as the GFP-Csvasa 3'UTR vector. The resultant vector was linearized by KpnI digestion for in vitro transcription. Chimeric mRNA containing the GFP gene fused to 3'UTR of the Csvasa gene was synthesized using a mMessage machine SP6 kit (Ambion, Austin, TX, USA) according to the manufacturer's instructions. Japanese medaka (Oryzias latipes) were maintained and artificially induced for spawning as previously described [22]. GFP-Csvasa 3'UTR mRNA (200 ng/µL, 0.05% phenol red) was injected into fertilized medaka eggs at the one-cell stage. Injected embryos were cultured in Holtfreter's solution at 28 °C. To visualize GFP expression, embryos at different stages were observed and photographed under a fluorescence microscope (Nikon, Japan).

### 2.7. Statistical analysis

Statistical analysis was performed using SPSS 17.0 software. Differences in *Csvasa* expression during different development stages or between different tissues were analyzed by one-way ANOVA followed by Duncan multiple comparison tests. Differences in *Csvasa* expression between males and females during the same and different developmental stages were analyzed by two-way ANOVA followed by Duncan multiple comparison tests. Differences between different sexes of the same stage were detected using a Student's *t*-test. The data were tested for normality and log-transformed to ensure homogeneity of variances before performing an ANOVA. A P-value of less than 0.05 was considered statistically significant.

### 3. Results

### 3.1. cDNA and genomic sequences of Csvasa gene

The expected 1097 bp fragment of *Csvasa* was obtained by RT-PCR, and the two remaining regions were produced by 3'- and 5'-RACE. Fragments were cloned and sequenced, overlapping segments were spliced *via* Vector NTI, and a fulllength 2445 bp *vasa* cDNA sequence (GenBank ID: KF819358) was isolated. In the present study, the *Csvasa* sequence consisted of a 5'UTR of 100 bp, an open reading frame of 1992 bp,

and a 3'UTR of 353 bp containing a polyadenylation signal. The sequence of the open reading frame encodes a 663amino acid protein. A total of 23 exons and 22 introns were identified by comparing the Csvasa cDNA sequence with the 6425 bp genomic DNA fragments (Fig. 1). Exon lengths varied between 21 bp and 452 bp, and all introns followed the GT-AG splicing rule, except the intron 8, which was found to have a GC-AG splice site. In accordance with the bioinformatics analysis, the predicted protein contained eight conserved domains of the DEAD-box protein family, including GYRKPTPIQ (Q motif), AQTGSGKT, GG, TPGR, DEAD, SAT, RGLD, and HRIGRTGR. In addition, the N-terminus comprised 10 arginine-glycine repeats, five arginine-glycine-glycine repeats, and the sequence EARKF. Well conserved tryptophan (W), glutamic acid (E), and aspartic acid (D) residues were also present in N-terminal and C-terminal regions as typical characteristics of Vasa protein (Fig. 2).

#### 3.2. Phylogenetic and alignment analysis

To support the homology of the putative amino acid sequences of Vasa and determine the phylogenetic relationship between tongue sole Vasa (CsVasa) and other species, a phylogenetic tree was constructed with fulllength Vasa sequences from 29 species belonging to six phyla, including Chordata, Nematoda, Cnidaria, Arthropoda, Echinodermata, and Mollusca. Phylogenetic analysis revealed that there were eight main branches, five of them corresponding to invertebrates. The remaining three branches belonged to vertebrates including teleost fish, amphibians, birds, and mammals. CsVasa belonged to homologues of the teleosts and clustered with Senegalese sole (Solea senegalensis) (Fig. 3A). The CsVasa sequence was aligned with other typical Vasa sequences available in the GenBank by the BLASTP program. The result revealed that the putative CsVasa shared 43.3% to 71.3% identity with typical Vasa of other organisms (Fig. 3B).

# 3.3. Csvasa expression in different tissues and during embryogenesis

Real-time quantitative PCR was performed to study *Csvasa* mRNA expression levels in different tissues including heart, liver, gill, skin, blood, kidney, intestine, brain, spleen, muscle, pituitary, and gonads of 1-year-old tongue sole. The sampled fish had ovaries at stage III of development and testes at stage IV. The results showed that *Csvasa* expression was mainly restricted to the gonads, with higher levels in the ovaries. Expression was also observed at a very low level in gill, liver, and heart, but was undetectable in other tissues (Fig. 4).



**Fig. 1.** Schematic representation of the *Csvasa* gene. Exons are represented by boxes (coding region by black box; UTR by blank box) and introns are denoted by black lines. Exon and intron lengths are indicated above and below the schematics, respectively. Scale bar = 200 bp.

### J. Huang et al. / Theriogenology xxx (2014) 1-12

1 GCAGAAACGCTCTGCACACGGCTCGTGACTGTGAGCTAGGACAGTGCGGCGAAATCAAGCTGAAGTTTACCTCAGCTTACACTTCACCAG 91 gtatgatttcgaattaattatatacatgtttaagctttagattattttattttattattgaaatgttgataccttaatcgtttgttgctcg 271 gaatgcaccttggtttttttatttcgttttaacttaatgtttgaaagatacatttggtgtttattggtatttgcagctcgttgattttct 361 gatetccattgtacgtcaatgtagttacgaagtgaggccaaaagtettgaccetggttaactacgtaaaaagttaggagttaaagttaat  $451\ tttttgaaataatgcactaaataaccaacgggtaggttcaaagattgagtgtattgcttttcaaacgtcagtgaaacgctgtaaattagga$  $541 \ {\tt gtcgttcgcaagagcacgaggctggatgtttaaattgcacaaaacttaatttacgtctcgcgatgtttccaatttcccggtatttgttg$ 631 cgcccttttgtgaagactacacaaaccagatccccataacggggttatacgatcctttttttaaaccacgttttttcttttcttttatcc 811 at att tcccatt tt at ctcaat cat at atg tcaa a at cat a ctt tt g a t a a g cag ctt tt catt g ctg ctt t at atg ctg g tt g cacat  $991\ {\tt ctacacctttcagagatgtcctggctgattatacaaaaaggggggaaatgggcagataaggagaccacatgcagaatagaaaagcagag$ MDDWEET 1 8 ENAEVV 1261 ttactatgctaattgtgtccaactgcagttctttctaactgtcaacacctgatcaatgacctttgattatagGAAAATGCAGAAGTGGTG 14 S V I S C A S T E G 1351 TCTGTAATCAGCTGTGCCTCAACAGAAGqtgattttactttccctttctqqtttctqtcatttctqqqaaaaatattttatqatacatca 24 SREHFWKDDSS 1441 tggtgtttactgctaaaattaaactatgtataactgtaagataacactgtttcagGTTCAAGGGAACACTTCTGGAAAGATGACAGTTCT 35 G K R  $1531 \ {\tt GGAAAGAgtaagaaaaaaagagtgaacctgtttgccctttgtgaaccctgcgtaagatttttgtcttaatttgttcaggcagtggaggtt$ 1621 ggcgaggcaggtggggaagatttacagattctcttttctcaagtaattcatttcaatctatgttgttgaacattttggaaattgctagat 1711 gacttgaaatgggacataqcaqtcqctagatggcattttatqtcctqqtttqgaaaagaaggqttqqtqttcactaqtataagacctqca  $1891 \ gagatgatgatgtgctttaatctaaaatctggcaaaactactatactaccattctctaaacaattgccacagttgttaatgcttctgcct$ 38 A D E K S N G T G T 2071 tggctgttgaggcacaaaactaactgggctgttttctctgtcattatgtgccttctgcagGAGCTGATGAGAAAAGTAACGGCACTGGAA 48 I T G T S S F N Q D 2251 tgccacacagtacatgtctggtttgttttccatatgttagaactagttgcaaactatcatatcataactatcgtgtcccttaatccgttt 58 WEFFKN 2341 aacqatqacacaatqaaqqcttaaatqtaaaaqtcattaccccactqaccctctqatttqtcacaaqcaqATTGGGAATTCTTCAAGAA 64 T E G Q Q G G I R G R G 2431 CACAGAAGGACAACAAGGTGGTATTAGAGGAAGAGgtgagaataatccgtcacctgcttcaagggacatttaaggcgaactatgtagctc 2521 cctttcctgaaatccgtactgagctcattgttggaccctgtcaataattcacacaatcggcacactcaccattaaagacacaattcaagg 2701 cttattcaaaatgtttcactctcaaatatgctgtgtggactccctgtgaaattgaaattattcagtgttattttttaagaattccctgtgt 76 <u>RGRG</u>FGRENR  $2791 \ agtaaatcctcaagattgtgtcattcattcttcaattcttatgtttatttgtcagGACGCGGCAGAGGATTTGGGAGAGAGAACAG$ 86 S D F N G  $2881 \ {\tt AAGTGATTTCAATGgtatgttcatggttcctacttccttaatgtacattcaaatagcctgctgcactgaacttattctactctttatttt$ 91 DSNGFCESE  $2971\ {\tt ttttcacaqGAGACAGCAATGGATTTTGTGAAAGCGgcaaqtgttatggactatctttatattaattattctgtttgatattgctaagt$ F<u>RG</u>VS<u>RGG</u>HGG<u>RGG</u>RGFR 100 3061 gataataacagtgctggatgatgtgaacagAGTTCAGAGGAGTAAGCCGAGGAGGTCATGGTGGCCGAGGAGGAAGAGGGAAGAGGCTTCA 120 R G  $\tt 3151 \ GACGTGgtatgaactaccatttttttttcactttgtgtaatgatatgtgttcataataggaatatgtcagaaataaccacatccctcatt$ TEGGS<u>RGG</u>FRED 122 3241 ggtgccagtcagtgatgttgaccattttgctgcagGCACTGAAGGAGGCTCCAGGGGAGGCTTCAGAGAAGgtaataagatataataaca 134 GEQRG<u>RGG</u>F<u>RGG</u>  $\tt 3421 \ \tt ACGGAGAACAAAGAGGTAGAGGAGGCTTTAGAGGAGGtacaaaaacactttttttgcccctgtgatccatataaaagtgacgatgataac$ Y<u>RG</u>KGENNFF 146 3511 cagtgatgtttcaggagacaagtagcttgtattaatccatggcatttgtgtaacggcagGTTACCGTGGAAAGGGCGAAAATAACTTTTT 156 Q E 158 K E N K E L K Q R D E G S A D K 174 PKVSYIPPALAEDE

3781 tgggattgttcctcagcttactctttcctcctgctctgacccacagAGCCAAAAGTCAGTTATATCCCTCCAGCCCTCGCTGAGGATGAA 188 N S V F A H Y E T G I N F D K Y D D I L V N V S G T N P P Q

6

## ARTICLE IN PRESS

J. Huang et al. / Theriogenology xxx (2014) 1-12

3871	AACTCAGTTTTTGCTCACTATGAGACTGGAATCAATTTTGACAAGTATGACGACATACTGGTGAACGTCAGTGGAACCAACC
218	A I M
3961	${\tt GCAATAATGgtatgatcaaatgcagtatgtctagactgaaattatgcataagagctacaaatctgcccccctgccccccctgacg}$
221	T F E E A A L C D S L S K N V S R S G Y V K
4051	ccactqcatqtttactqtqtttaqACATTTGAGGAGGCTGCCCTGTGCGACTCTTTATCCAAAAATGTCAGCAGATCTGGATATGTCAAG
243	PTPVOKHGIPIISAGRDLMACAOTGSGKT
4141	CCGACTCCTGTACAGAAGCACGGCATTCCTATCATTTCTGCTGGCAGAGATCTCATGGCCTGTGCACAGACGGGGTCTGGAAAAACGgtg
272	
1021	
4231	
277	PILHQLMIGGASSSFFSELQEPKAIIVA <u>PI</u>
4321	GCCCATCCTGCATCAGCTGATGATAGGCGGAGCATCATCATCTTTCTT
307	<u>reli</u> n qiflearkfay g
4411	CAGGGAGCTTATAAACCAGATTTTTCTGGAGGCCAGGAAATTTGCCTATGGgtataattttataacaaactgtggctatg <u>a</u> taaaaaaa
324	T V V R P V V V Y <mark>G G</mark> V S
4501	caaacaacattgaggtaataaaagagctgcttctgcttcttcttcttccccagAACTGTTGTGCGTCCAGTGGTTGTTACGGTGGAGTCA
337	I R H Q V S D L N K G C N V L C G <mark>T P G R</mark> L M D M I E R G K
4591	GTATTAGACACCAGGTTAGTGACCTTAACAAGGGATGTAATGTCCTTTGTGGAACACCAGGGCGGCTGATGGACATGATAGAGAGAG
367	V
4681	AGgtaaaqataaacaaaaggatttgcaactgcttgagtttgttgtcattataaagggtctgaatacggatgctgtgatcatcttagGT
368	G L S K V O Y F V L D E A D R M L D M G F E P D M R R L V A
4771	TGGGTTGAGCAAGGTGCAATACTTTGTTCTGGATGAGGCTGATCGAATGTTGGACATGGGCTTCGAGCCAGATATGCGCCCGCC
398	S P G M P S R F G R O T I M F <b>S A T</b> F P F F I O K
4861	
1001	
425	
4951	
431	DYLFLAVGIVGGACSDVEQTFIEVGKFSK
5041	AGACTATCTCTTTTGCCTGTCGGAATAGTGGGTGGAGCTTGCAGTGACGTAGAGCAGACGTTTATCGAAGTTGGCAAGTTCTCAAAGAG
461	E Q L L D L V K T C G
5131	GGAACAGCTTCTTGATCTTGTGAAGACATGTGgtacttttttattttagtttttttagttttttagcttattcctgccacagacag
472	N E R T I V F V E K K R Q A D F I A T F L C Q
5221	aatacagtctaatgttgcagGAAATGAGCGCACCATTGTGTTTGTGGAGAAGAGAGACAGGCGGATTTTATCGCCACATTCCTGTGTCA
495	E N V P T T S I H G
5311	GGAGAATGTTCCAACCACCAGCATTCATGGgtacgttacatgttcaacacttgtacagcaactactctggacttgtataggtgtaataat
505	D R E Q W Q R E Q A L A D F K Y G K C P V L V A
5401	attttattttgtactgcagGGACCGTGAGCAGTGGCAGCGAGGAGCAGGCCTTAGCAGACTTTAAATACGGCAAATGTCCAGTCCTGGTTG
529	TSVAARGIDTPDVOHVVNFDIPGSTDEVVH
5491	
0 1 9 1	CAACCTCTGTGGCAGCCCGCGGTCTGGATATTCCTGATGTTCAGCATGTGGGTCAACTTCGACCTCCCAGGCAGCATTGATGATGATGATGTTC
559	CAACCTCTGTGGCAGCCCGCGGTCTGGATATTCCTGATGTTCAGCATGTGGTCAACTTCGACCTCCCAGGCAGCATTGATGATGATGATGTTC R I G R T G R C G N V G R A V S F F D T D V D G H L S R P L
559 5581	CAACCTCTGTGGCAGCCGCGGTGTGGATATTCCTGATGTTCAGCATGTGGGTCAACTTCGACCTCCCAGGCAGCATTGATGAGTATGTTC R I G R T G R C G N V G R A V S F F D T D V D G H L S R P L ACCGCATTGGGAGAACTGGCCGCTGTGGGAACGTGGGTAGAGCGGTGTCTTTCTT
559 5581 589	CAACCTCTGTGGCAGCCGCGGGTGTGGATATTCCTGATGTTCAGCATGTGGGTCAACTTCGACCTCCCAGGCAGCATTGATGAGTAGGTAG
559 5581 589 5671	CAACCTCTGTGGCAGCCGGGTCTGGATATTCCTGATGTTCAGCATGTGGGTCAACTTCGACCTCCCAGGCAGCATTGATGAGTAGGTC R I G R T G R C G N V G R A V S F F D T D V D G H L S R P L ACCGCATTGGGAGAACTGGCCGCTGTGGGAACGTGGGTAGAGCGGGTGTCTTTCTT
559 5581 589 5671 595	$\begin{array}{c} \label{eq:constraint} \textbf{CAACCTCTGTGGCAGCCGGGGTGTGGATATTCCTGATGTCAGCATGTGGGCAACTTCGACCATCCCAGGCAGCATTGATGAGTATGTTC}\\ \hline \textbf{R} & \textbf{I} & \textbf{G} & \textbf{R} & \textbf{C} & \textbf{G} & \textbf{N} & \textbf{V} & \textbf{G} & \textbf{R} & \textbf{V} & \textbf{S} & \textbf{F} & \textbf{F} & \textbf{D} & \textbf{T} & \textbf{D} & \textbf{V} & \textbf{D} & \textbf{G} & \textbf{H} & \textbf{L} & \textbf{S} & \textbf{R} & \textbf{P} & \textbf{L} \\ ACCGCATTGGGAGAACTGGCCGCTGTGGGAACGTGGGTAGAGCGGGGTGTCTTTCTT$
559 5581 589 5671 595 5761	$ \begin{array}{c} CAACCTCTGTGGCAGCCGCGGTCTGGATATTCCTGATGTTCAGCATGTGGTCAACTTCGACCTCCCAGGCAGCATTGATGAGTATGTTC \\ \hline R  I  G  R  T  G  R  C  G  N  V  G  R  A  V  S  F  F  D  T  D  V  D  G  H  L  S  R  P  L \\ \mathsf{ACCGCATTGGGAGAACTGGCGCGCTGTGGGAACGTGGGTAGAGCGGGTGTCTTTCTT$
559 5581 589 5671 595 5761 620	$\begin{array}{c} \label{eq:construction} CAACCTCTGTGGGGCGCGCGGGTGGGATATTCCTGATGTTCAGCATGTGGTCAACTTCGACCTCCCAGGCAGCATGAGATGGGGATGGGTGGTTCTCTTGGGCGGCGCGCGC$
559 5581 589 5671 595 5761 620 5851	$\begin{array}{c} \label{eq:calcolor} CAACCTCTGTGGCAGCCCCCGGGTGTGGATATTCCTGATGTTCAGCATGTGGTCAACTTCGACCTCCCAGGCAGCATGAGTAGTGAGTAGTTCC \hline R I G R T G R C G N V G R A V S F F D T D V D G H L S R P L ACCGCATTGGGAGAACTGGCCGCTGTGGGGAACGTGGGTAGAGGGGGGTGTCTTTCTT$
559 5581 589 5671 595 5761 620 5851	$\begin{array}{c} \label{eq:constraint} CACCTCTGTGGGGCCCCCGGGTCTGGATATTCCTGATGTCAGCATGTGGTCACTTCGACCTCCCAGCCAG$
559 5581 589 5671 595 5761 620 5851 632	$\begin{array}{c} \label{eq:calcorrest} CAACCTCTGTGGCAGCCCCCGGGTCTGGATATTCCTGATGTCAGCATGTGGTCAACTTCGACCTCCCAGGCAGCATTGAGAGTAGGTCGTCTCTTGGGCAGCCCGCGGGTGGGGTAGGGGGGGG$
559 5581 589 5671 595 5761 620 5851 632 5941	$\begin{array}{c} \text{CAACCTCTGTGGCAGCCCGCGGTCTGGATATTCCTGATGTTCAGCATGTGGTCAACTTCGACCTCCCAGGCAGCATGAGAAGGTATGTTC\\ \hline \textbf{R} \ \textbf{I} \ \textbf{G} \ \textbf{R} \ \textbf{T} \ \textbf{G} \ \textbf{R} \ \textbf{T} \ \textbf{G} \ \textbf{R} \ \textbf{C} \ \textbf{G} \ \textbf{N} \ \textbf{V} \ \textbf{G} \ \textbf{R} \ \textbf{A} \ \textbf{V} \ \textbf{S} \ \textbf{F} \ \textbf{F} \ \textbf{D} \ \textbf{T} \ \textbf{D} \ \textbf{V} \ \textbf{D} \ \textbf{G} \ \textbf{H} \ \textbf{L} \ \textbf{S} \ \textbf{R} \ \textbf{P} \ \textbf{L} \\ ACCGCATTGGGAGAACTGGCCGCTGTGGGAAACGTGGGTAGAGCGGGTGTCTTTCTT$
559 5581 589 5671 595 5761 620 5851 632 5941 652	$\begin{array}{c c c c c c c c c c c c c c c c c c c $
559 5581 589 5671 595 5761 620 5851 632 5941 652 6031	$\begin{array}{c c} c ACCTCTGTGGCAGCCCGCGGTCTGGATATTCCTGATGTTCAGCATGTGGTCAACTTCGACCTCCCAGGCAGCATGAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG$
559 5581 589 5671 595 5761 620 5851 632 5941 652 6031 6121	$\begin{array}{c c} c AACCTCTGTGGCAGCCCCGGGTCTGGATATTCCTGATGTTCAGCATGTGGTCAACTTCGACCTCCCCAGCCAG$
559 5581 589 5671 595 5761 620 5851 632 5941 652 6031 6121 6211	$\begin{array}{c c} c ACCTCTGTGGCAGCCCCGGGTCTGGATATTCCTGATGTCAGCATGTGGTCAACTTCGACCCCCAGCAGCATGAGAGATGTGAGATGTCCAGCATTGGGGCAGCCGGGGTGGGGAACGTGGGTAGAGCGGGGGGGG$
559 5581 589 5671 595 5761 620 5851 632 5941 652 6031 6121 6211 6301	$\begin{array}{c c c c c c c c c c c c c c c c c c c $
559 5581 595 5671 595 5761 620 5851 632 5941 652 6031 6121 6211 6301 6391	$ \begin{array}{c} CAACCTCTGTGGCAGCCCGCGGTCTGGATATTCCTGATGTTCAGCATGTGGTCAACTTCGACCTCCCAGCCAG$



To investigate the expression differences of *Csvasa* mRNA in gonads of males, females, and neomales after temperature treatment, expression of *Csvasa* mRNA was detected by qPCR in fish 2 years old. The ovaries of 2-year-old sampled fish were at stage IV to V of development and

testes were at stage V. The number of sperms in the neomale testis was less than that in the normal testis. Lower expression levels of *Csvasa* mRNA were detected in the neomales, and relative higher levels were detected in males and females (Fig. 5).

**Fig. 2.** Genomic and deduced amino acid sequence (GenBank accession number:KJ664813 and KF819358) of *Cs*Vasa. Exons and introns are indicated in uppercase and lowercase, respectively. Shaded boxes represent the eight conserved motifs of DEAD-box protein family. Arginine-glycine (RG) and arginine-glycine-glycine (RGG) repeats in the N-terminus are marked by underline and double underline, respectively. The tryptophan (W), aspartic acid (D), and glutamic acid (E) residues in the C-terminus and N-terminus are shown inside boxes. The stop codon is indicated by an asterisk. The polyadenylation signal and poly-A tail are in bold.

To understand the expression patterns of *Csvasa* mRNA during embryonic development, expression levels were detected by qPCR. The results showed that *Csvasa* expression persisted throughout all embryonic stages, and expression levels gradually decreased from four-cell to hatching stage. Expression levels were high in embryos before the late gastrula stage and dropped abruptly from the late gastrula stage to the neurula stage. Expression levels decreased from the somites stage to the hatching stage, and much lower expression levels were detected at 1 dph (Fig. 6).

# 3.4. Sexually dimorphic expression of Csvasa during early development and sex differentiation

Sexually dimorphic expression of *Csvasa* was observed by qPCR from 4 to 150 dph. *Csvasa* expression levels gradually decreased from 4 to 16 dph in males and females, the lowest expression level in females was detected at 16 dph, and in males was detected at 46 dph. Expression levels increased continuously in both sexes until 150 dph. An interesting sexually dimorphic expression of *Csvasa* was observed during early development and sex differentiation, the expression levels in males were higher than that in females from 4 to 26 dph, and then females exhibited higher expression levels than males from 36 to 115 dph, but males showed higher expression levels again at 150 dph (Fig. 7).

### 3.5. Visualization of medaka PGCs by microinjection of GFP-Csvasa 3'UTR mRNA

In medaka fertilized eggs injected with GFP-*Csvasa* 3'UTR mRNA, GFP fluorescence in PGCs was observed over the whole blastodisc at stage 20 and was difficult to differentiate from that of somatic cells in the early segmentation period (Fig. 8A). Relatively strong GFP fluorescence in PGCs was observed (Fig. 8B, C) in presumptive gonad region aligned bilaterally along the trunk as the embryo developed, and the fluorescence signal gradually diminished after stage 29 (Fig. 8D). Primordial germ cells in medaka can be visualized by microinjection of GFP-*Csvasa* 3'UTR mRNA during embryogenesis.

### 4. Discussion

In the present study, a full-length 2445 bp cDNA and 6425 bp genomic DNA sequences of the *Csvasa* gene were isolated and analyzed. The length of Csvasa cDNA sequence in the N-terminus is shorter than another cDNA sequence (2602 bp) from GenBank (GenBank ID: HQ727561). The CsVasa sequence contained eight characteristic domains of the DEAD-box protein family, which are essential for ATPase-dependent RNA binding and unwinding activity [36,37]. Rich arginine-glycine and arginine-glycine-glycine repeats, G residues, and the EARKF sequence at the N-terminal region were observed. In the C-terminal region, six of the last eight amino acids were glutamic acid (E) residues, as observed in several single-stranded nucleic acid binding proteins [36]. These typical features indicated that the CsVasa belonged to a member of DEAD-box family. Phylogenetic analysis indicated that the CsVasa was closely aligned to the teleosts. Multiple sequence alignments showed that *Cs*Vasa had a considerable identity with other typical homologues. These results together provide further evidence that the *Cs*Vasa is a member of the Vasa homologues with an ATP-dependent RNA helicase function and suggests high conservation during evolution.

Tissue distribution analysis of *Csvasa* was performed by qPCR, and the gonadal stages of adult fish were determined as previously described [38,39]. The results showed that the Csvasa mRNA was mainly detected in gonads. Specific expression of the vasa gene in germ cells was verified to play a prominent role in germline development [40]. Except for the gonads, very low expression was observed in the gill, liver, and heart of tongue sole. Vasa mRNA expression has been reported in the heart and brain of rainbow trout (Oncorhynchus mykiss) [13] and the kidney and brain of an adult frog species [41]. This could be because Vasa is implicated in translational regulation of mRNAs involved in the specification and differentiation of tissue-specific cell types through its helicase activity [42]. It is possible that the role of vasa might be multipotent rather than confined to the germline. Tongue sole subject to high temperatures during early development and sex differentiation showed a high male phenotypic sex ratio (about 70%). Lower levels of Csvasa mRNA in neomales were detected and compared with normal fishes, and the results suggested that Csvasa expression might be restrained and the Csvasa gene might be implicated in germ cell development after sex reversal. We inferred that the Csvasa expression pattern might be more related to the genetic sex than to the physiological sex, and lower expression in neomales is also probably due to worse development of gonads under high temperature treatment. The exact reason will need to be further studied.

Expression patterns of vasa mRNA during embryogenesis have been studied using RT-PCR and whole-mount in situ hybridization for many teleosts, such as zebrafish [4], grass carp (Ctenopharyngodon idella) [8], Japanese flounder (Paralichthys olivaceus) [10], Atlantic cod (Gadus morhua) [15], turbot (Scophthalmus maximus) [16], and catfish (Clarias gariepinus) [43], and the consistent conclusion was that vasa mRNA was a maternally supplied transcript. For example, the expression levels in the grass carp during the early embryonic stages were high, even in fertilized eggs, and these high levels decreased gradually during cleavage and blastula stages and continued to decrease to lower levels until the gastrula stage. Four vasa transcripts have been isolated from Senegalese sole, Ssvasa1 and Ssvasa2 were maternally supplied, Ssvasa3 and Ssvasa4 were dependent on the *de novo* expression of the growing juveniles, which suggests that expression patterns of vasa mRNA could vary [11]. In the present study, the results suggest that the *Csvasa* transcript was also maternally supplied, as it was abundant in embryos before the late gastrula stage and decreased gradually from the late gastrula stage to the hatching stage, and expression levels from somites to the hatching stage were much lower than those of cleavage stages. Decreased expression levels of Csvasa might be from degradation of maternal vasa mRNA or a reduced PGC ratio caused by a rise in the total number of embryonic cells. Further studies are needed to determine the reasons.

J. Huang et al. / Theriogenology xxx (2014) 1-12



**Fig. 3.** Multiple alignment and phylogenetic tree of *Cs*Vasa and other DEAD-box family proteins. (A) Phylogenetic tree of *Cs*Vasa constructed by MEGA 5.1 using the neighbor-joining method. Bootstrap values of 1000 replicates (>80%) are indicated next to the branches. The sequences used for alignment are as follows: *Drosophila melanogaster (DmVasa,* AAF53438); *Homo sapiens (HsVasa,* AAF72705); *Mus musculus (MmVasa,* BAA03584); *Gallus gallus (GgVasa,* BAB12337); *Xenopus laevis (XlVasa,* AAC03114); *Danio rerio (DrVasa,* AAI29276); and *Oryzias latipes (OlVasa,* BAB61047). (B) Multiple alignment was generated with the AlignX program. Eight conserved motifs are framed in block. The identities of *Cs*Vasa to other Vasa proteins are indicated following the sequence alignment.

To understand the expression patterns of Csvasa in males and females, respectively, the sex linkage microsatellite marker and gonad-specific expression genes were used to distinguish sex and gonad parts. This method was used because sex could not be distinguished by the morphological characteristics and gonads could not be isolated during early development. In addition, the PGC proliferation and gonad sex differentiation have been previously described [44]. Expression of Csvasa mRNA gradually decreased from 4 to 16 dph in both sexes, except for the degradation of maternal vasa mRNA, a possible reason for the expression is that the amount of the maternally inherited vasa transcript became diluted with individual growth. Interestingly, the lowest expression level in females was observed at 16 dph, and males at 46 dph. We inferred that prolonged degradation of maternal Csvasa mRNA might be caused by higher expression in male larvae, and the *Csvasa* might play a more significant role in males before sex differentiation. Expression levels increased gradually in both sexes, and the primary gonads enlarged concomitant with the start of PGC division and mass proliferation at this stage. Similar expression trends have been reported in tilapia (Oreochromis niloticus) [45], zebrafish [46], and European sea bass (Dicentrarchus labrax) [9], suggesting a switch from maternal to de novo vasa expression. Expression levels in females were higher than those in males from 36 to 115 dph, which was probably because histological differentiation of the gonads in females was faster and earlier than that in males [47]. In European sea bass and catfish, it was reported that *vasa* expression levels in females were always higher than those in males during early development and sex differentiation [9,43]. It is, however, exceptional that the expression levels in male tongue sole were higher than the expression levels in females at 150 dph. The histological results suggest that histological differentiation in females happened earlier than in males, but cytological differentiation was faster in males than in females [47]. Although males were immature at 150 dph, a small population of sperm was observed by chance during sampling. This is consistent with the phenomenon that sexual maturity happens earlier in male than in female tongue sole.

*Vasa* mRNA expression levels during gonad differentiation have been reported for several fish species [9,43–46]. However, accurate quantification of *vasa* mRNA levels in males and females during early development and sex differentiation using the sex-linked microsatellite marker and gonad-specific expression genes has been carried out in very few fish species [48]. Results of the present study suggest that *vasa* expression presented a sexually dimorphic pattern during this stage and that *vasa* gene might play a differential role in the proliferation and differentiation of male and female PGCs. Therefore, this information on *Csvasa* expression during the early development and sex

ARTICLE IN PRESS

J. Huang et al. / Theriogenology xxx (2014) 1–12

Р	(1)	N 1 10 20 30 40 50 60 70 80	91
	(1) $(1)$	)MDDWEETENAEVVSVISCASTEGSREHFWKDDSSGKR	
DrVasa	(1)	)mdd <mark>we</mark> edq-spvv <mark>s</mark> cssgfglgsngsdggfksfytggagndksn <mark>segt</mark> egsswkmtgd <mark>s</mark> frgr	
DmVasa	(1)	) -MS9DDWDDEPIVDTRGARGGDWSDDEDTAKSFSGEAEGDGVGGSG	
Ggvasa HsVasa	(1)	) - MEEDWUTELEQEAAAASQGKSEEQAMANSGKPNSPSLKPS5 MGDEDWUTELEQEAAAASQGK	NRDA
MmVasa	(1)	) MGDEDWAALIN FINNSSIVFI ) MGDEDWAALIN FINNSSIVFI	DI
OIVasa	(1)	)MDDWEEEETAPSFAPVSSTDAAPQRSSWNGGARDSG	
XIVasa	(1)	)-MEENWDTEIETEKPTYVPNFSTLETENTDNYSAYSNNDINNQNYDSE	
Consensus	(1)	) DWEE S DSGSNSS	
100000	(92)	) <u>92 100 110 120 130 140 150 160 170</u>	182
CsVasa	(38)	)SSFNQDWEFFKNTEGQQGGIRGSSFNQDWEFFKNTEGQQGGIRG	R
Drvasa	(03)	)GRGGSRGRGGCSGFRSEIDENGSDGGWNGGESRGRGGFRGGFRGSDENDENGNDDGWRGGESR )	GOGG
GgVasa	(43)	RPSSPLSGFPGRPNSPFFGFSQNKGSLGANEGLNRSLFVQHDIGGYSGSRESVVRQNRE	DQPV
HsVasa	(71)	) GECNKRDNTSTM <mark>GG</mark> FGVGKSFGNR <mark>GF</mark> SNSRFED <mark>G</mark> DSS <mark>G</mark> FWRESSNDCEDNPTRNRGF <mark>S</mark> KRGGYRDGNNSEA-	SG <mark>P</mark> Y
MmVasa	(46)	) GESSKKENTSTT <mark>GG</mark> FGRGKGFGNR <mark>GF</mark> LNNKFEE <mark>G</mark> DSS <mark>G</mark> FWKESNNDCEDNQT <mark>R</mark> SRGF <mark>S</mark> KRGGCQDGNDSEA-	SG <mark>P</mark> F
ONasa	(37)	)NDGDSWNRSNRG <b>G</b> GSAG	RGGR
Consensus	(48)	)RSFGNRGETISERSRPSNFNRGSRFRRGRGRGFGTNRNDN SSERDVFGDDERDQRRGFGRGGTNGNEDG	P
,	()		272
(cv/aca	183) (75)	) 183 190 200 210 220 230 240 250 260	Z / 3
DrVasa(	136)	GGFGGGFRGGGFRGGREDTGRGGFGGFREDNENGDEGEGRGRGGGFRGGFRDGGBESGKRGFGRGGFR	GRNE
DmVasa(	107)́	) SRG <mark>GQG</mark> GSRGGQ <mark>GGFRGCEG</mark> GFRGRLYENEDGDERR <mark>G</mark> RLDREERGGERRGRLDREERGGER <mark>G</mark> ER	GDGG
GgVasa(	106)	) TRF <mark>GRG</mark> RSSGSRDFQERNSANDPGMQDQGFRRVPGIF <mark>G</mark> QSKCFNSEERNSPLRGSPFAP <mark>G</mark> GR <mark>G</mark> AV	GGPA
HsVasa(	146)	) REGERESFEGCREGEFGLESPNNDLDPDECMQRTGELFGSRRPVLSGTGNGDTSQSRSGSCSEREGYK	GLNE
OM/asa	121) (59)	) GGR <mark>GRGFGRSDODELNGGGGDSENGFRGRGRGCRGGFRSGCEPPCRCCCVP</mark>	GRDE
XIVasa	123)	) AFR <mark>GRG</mark> GFRNENEQRRGF <mark>GERGGF</mark> RSENGQRNFDNRGDFGNSGEEEDRPRSYGRGGFNNSDTGGRGRRGGRGGGS-OYGGYK	GRNE
Consensus(	183)	R GRG R G G G G G G G G G G G G G G G G	G NE
0	274)	274 280 290 300 310 320 330 340 350	364
CsVasa(	152)	) NNFFQEKENKELKQRDEGSADK <mark>PKVS<mark>XIPP</mark>ALA<mark>EDENSVFAH-YETGINFDKY</mark>DDIL<mark>V</mark>NVS<mark>C</mark>TNP<mark>PQAIMTFEEA</mark>ALCDSLSKN</mark>	V <mark>S</mark> RS
DrVasa(	209)	) E-VFSKVTTADKLDQEGSENA <mark>GPKVVYVPPPPPEEESSIFSH-YATGINFDKYDDILVDVSG</mark> SNP <mark>PKAIMTFEEAGLC</mark> DSLSKN	VSKS
DmVasa(	175)	) FARRENEDDINNNNN-IVEDVERKREFYI PEEPSNDATEI FSSGTASGIHFSKYNNI PVKVTGSDVPPI OHFTSADLRDIIDN	VNKS
Gyvasa HsVasa	217)	J GVINGKSE-EIDSG KGENVIIVFEFFEDGSIFACIGSGIRFDKIDELVEGDPFAFLAFEDARAQIDKAN I EVITGSEKNEWKSEFECGFSSTOCEKVIVT DEDEEDEDEETEL-VOTCINERKVOTIDELVEGDEDEFAFLAFEDARAQIDKAN	TAKA
MmVasa(	190)	) EVVTGSGKNSWKSETEGGESSDSQGPKVTYIPPPPPEDEDSIFAH-YQTGINFDKYDTILVEVSGHDAPPAILTFEEANLCQTLNNN	IAKA
OIVasa(	114)	) D-VFAAGDGRGAENSDAADPERPKVT <mark>YIPP</mark> SL <mark>PEDEDSIFSH-</mark> YKM <mark>GINFDKY</mark> DDIL <mark>V</mark> DVS <mark>C</mark> TNL <mark>PAAIMTFEEA</mark> KLCESLENN	ISRS
XIV asa(	208)	) EVGVESGKSQEEGNEKDEKPKKVT <mark>YIPP</mark> PPPDGEDNIF <mark>RQ-YQSGINFDKY</mark> DEIL <mark>V</mark> DVT <mark>G</mark> KDV <mark>P</mark> PAILTFEEANLCETLRR <mark>N</mark>	VARA
Consensus	2/4)	) EV G D GPKVTYIPPPPPEDEDSIFAH YQTGINFDKYD ILV VSG D P AILTFEEANLC TL N	SKS
()	365)	365 370 380 ATPase Motif A 400 410 420 430 440	455
CsVasa(	239)	) <mark>GYVKPTPVQKHGIPIISAGRDLMAQAQTGSGKT</mark> AAFLLPILHQLMIGGASSSFFSELQE <mark>PKAIIVAPTRELINQI</mark> FLE <mark>ARKF</mark> AYGTV	VRPV
DrVasa(	295)	) GYVKPTPVOKHGIPIISAGRDLMAGAQTGSGKTAAFLLPILQRFMTDGVAASKFSEIQEPEAIIVAPTRELINQIYLEARKFAYGTC	VRPV
DmVasa(	264)	) GYKIPTPICKCSIPVISSGRDLMAGAOTCSGKTAAFLEFISKLEDPHELELGRPOVVIVSPTRELAIOIFMEARKFAFESY	LKIG
Gyvasa(	2337		INFV
HSV asa.	3071	I GITNLIEVUNISIELLLAGKULMAUAUIGSGN IAAFLLEFILAHMMHUGITASKENELUEEEUILVAFIKELVNULILEAKNESFGIU	VKAV
MmVasa(	307) 280)	) GITKLTPVQRISIFITLAGROLMAQAQTSGKTAAFLDFILAHMRHDGITASRFKELQEPECITVAFTKELNQITLEAKKESFGTC ) GYTKLTPVQKYSIPIVLAGROLMAQAQTGSGKTAAFLLPILAHMRDGITASRFKELQEPECIIVAPTKELNQIYLEAKKESFGTC	VRAV
MmVasa( ONasa(	307) 280) 200)	OTINIE VVRTSTETTI LAGRUMANAQUSSCAMARILETIANMANDGIASKRALDEBECTIVARIKELMUUTULAKKESPETC GYTKLEVGKYSIETVLAGRUMAQUSSCKTARILETIANMANDGIASKRALDEBECTIVARIKELMQYLEAKKESFETC GYVKPTPVQKYSLETIISAGRUMAQUSSCKTARILETIANMANDGIASKRASEIGEBEAVIVARIKELMQYLEAKKESFETC	VRAV VRAV VRPV
MmVasa( MmVasa( ONasa( XIVasa(	307) 280) 200) 293)	OTINLIPVENTSILILAGRDINA AQUOSCA MARILELELANMANG INASEKELGE BOLIVARI HELU VULLA KASISI O OTINLIPVENTSILILAGRDINA AQUOSCA MARILETLANMANG INASEKELGE BOLIVARI HELU VULLA KASISI O OTINLIPVENTSILILAGRDINA AQUOSCA MARILETLANMANG INASEKELGE BOLIVARI HELU VULLA KASISI O OTIVLIPVENTSILILAGRDINA AQUOSCA MARILETLANMANG INASEKELGE BOLIVARI HELU VULLA KASISI O OTIVLIPVENTSILILAGRDINA AQUOSCA MARILETLANVAN GI VASISI OLOBEAN VULLA MARIA VULLA KASISI OLOBEAN OTIVLIPVENTSI DI MAGRDINA AQUOSCA MARILETLATI VULLA VULLA VULLA KASISI OLOBEAN VULLA VUL	VRAV VRAV VRPV VRPV
MmVasa( MmVasa( OIVasa( XIVasa( Consensus(	280) 280) 200) 293) 365)	GYTKLEYQKISIEIILAGRDIMAQUGSGKIAAELDEILAHMMHDGITASFKELDEEGUTVAFIKELUUUU GYTKLEYQKYSIEIVLAGRDIMAQUGSGKIAAELLETLAHMMHDGITASFKELDEEGUTVAFIKELNOILLARKESFGTC GYVKPTPVQKYGLEIISAGRDIMAQUGSGKIAAELLEILDILAUMADGVAASRESEIGEEAVTVAFTKELNOILGARKESFGTC GYVKLEYQKHSIEIIMAGRDIMAQAQTGSGKIAAELLEILSYMMNEGITASGYLQLOEESAIIIA <u>PTKEL</u> NOIYLDAKKESYGTC GYVK TPVQKHSIEIIMAGRDIMAQAQTGSGKTAAELLEILSYMMNEGITASGYLQLOEESAIIIA <u>PTKEL</u> NOIYLDAKKESYGTC	VRAV VRAV VRPV VRPV VRPV
MmVasa( ONVasa( XIVasa( Consensus(	307) 280) 200) 293) 365) 456)	OTIXLIPV0RISIEILDAGRDIMAGOTOSCHARELDELDAMMENDEIDASRFSLDE BCUTWAFRELMOTULARRESFELDE OTIXLIPV0RYSIEIVLAGRDIMAGOTOSCHARELDEILAMMENDEIDASRFSLDE BCUTWAFRELMOTULARRESFELDE GYVKPTPV0RYGLPIISAGRDIMAGOTOSCHARELDELDEILAMMENGITASQYLOLOB PSATIIAPTRELMOTULARRESFGTC GYVKLIPV0KHSIPIISAGRDIMAGOTGSCKTAAFLLPILSYMMEGITASQYLOLOB PSATIIAPTRELMOTYLOARRESYGTC GYVK TPV0KHSIPIISAGRDIMAGOTGSCKTAAFLLPILSYMMEGITASQYLOLOB PSATIIAPTRELMOTYLOARRESYGTC GYVK TPV0KHSIPIISAGRDIMAGOTGSCKTAAFLLPILSYMMEGITASGYLOLOB PSATIIAPTRELMOTYLOARRESYGTC GYVK TPV0KHSIPIISAGRDIMAGOTGSCKTAAFLLPILSYMMEGITASGYLOLOB PSATIIAPTRELMOTYLOARRESYGTC GYVK TPV0KHSIPIISAGRDIMAGOTGSCKTAAFLLPILSYMMEGITASGYLOLOB PSATIIAPTRELMOTYLOARRESYGTC GYVK TPV0KHSIPIISAGRDIMAGOTGSCKTAAFLLPILSYMMEGITASGYLOLOB PSATIIAPTRELMOTYLOARRESYGTC ASSATING AND ADAGATAGOTGSCKTAAFLLPILSYMEGITASGYLOLOB PSATIIAPTRELMOTYLOARRESYGTC GYVK TPV0KHSIPIISAGRDIMAGOTGSCKTAAFLLPILSYMEGITASGYLOLOB SIGNAFAGATAGOTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	VRAV VRAV VRPV VRPV VRPV 546
HSV434 MmVasa( OlVasa( XIVasa( Consensus( CsVasa( CsVasa(	307) 280) 200) 293) 365) 456) 330)	OTTALEPVRISTITUS Indextransmitus Ottot Indextransmitus Ottot Indextransmitus Ottot Indextransmitus <t< th=""><th>VRAV VRAV VRPV VRPV VRPV 546 PEEI</th></t<>	VRAV VRAV VRPV VRPV VRPV 546 PEEI
HSV434( MmVasa() OlVasa() XIVasa() Consensus() CsVasa() DrVasa() DrVasa()	307) 280) 200) 293) 365) 456) 330) 386) 351)	OTTALLEVENTS1 Indexeduated to some that is be is an indexeduated to some the index the indexeduated to some the indexeduated	VRAV VRAV VRPV VRPV VRPV 546 PEEI PEDI PEEI
nsvasa MmVasa( OVasa( XVasa( XVasa( Consensus() CsVasa( DrVasa( GgVasa()	307) 280) 200) 293) 365) 456) 330) 386) 351) 345)	OTYKLIPVOKISIETULAGKUDANAGOTOSOKIAAFLDEILAHMENDOITASEKELDE BOTI VARTKEL NOTLEAKKESTGIO   OYYKLIPVOKYSIETULAGKUDAGUTOSOKIAAFLDEILAHMENDOITASEKELDE BOTI VARTKEL NOTLEAKKESTGIO   OYYKLIPVOKYSIETULAGKUDAGUTOSOKIAAFLDEILAHMENDOITASEKELDE BOTI VARTKEL NOTLOBAKKESYGTO   OYYKLIPVOKHSIPTIAGRDIMAGOTOSOKIAAFLDEILAHMENDOITASEKELDE BOTI VARTKEL NOTLOBAKKESYGTO   OYYKLIPVOKHSIPTIAGRDIMAGOTOSOKIAAFLDEILAPILAMENGTASEKELDE BOTI VARTKEL NOTYDEAKKESYGTO   OYYKLIPVOKHSIPTIAGRDIMAGOTOSOKIAAFLDEILAPILAMENGTASEKELDE BOTI VARTKEL NOTYDEAKKESYGTO   OYYKLIPVOKHSIPTIAGRDIMAGOTOSOKIAAFLDEILAPILAMI MAGUTASEKSYGTO   OYYKLIPVOKHSIPTIAGRDIMAGOTOSOKIAAFLDEILMI MM DG TAS FELDEPENTIAPTKELINOTYDEAKKESYGTO   456 470 480 490 ATPase Motif B 510 520 RNA unwinding Motif   VVYGUNTGYTIREVLKOCULCATEGRLHDLIGRGKIGISKVRYLVIDEARMLDMGFEPEMKRLVASPGMPSKEEGOTIMISATY 1000000000000000000000000000000000000	VRAV VRAV VRPV VRPV 546 PEEI PEEI PEEI PEEI
MmVasa( ONasa( ONasa( XIVasa( Consensus( Covasa( DrVasa( GgVasa( HsVasa(	307) 280) 200) 293) 365) 456) 336) 336) 386) 351) 345) 398)	0 STKLEPVCKYSIE TULAGRDUMA QUGSCHAAFLDE LLAHMENDGITASFEKELGE FOLT VAFTEL INGTI LARKESFGTC   0 SYKLEPVCKYSIE TULAGRDUMA QUGSCHAAFLDE ILAHMENDGITASFEKELGE FOLT VAFTEL INGTI LARKESFGTC   0 SYKLEPVCKYSIE TULAGRDUMA QUGSCHAAFLDE ILSTMINGGITASFEKELGE FOLT VAFTEL INGTI VAFTEL	VRAV VRAV VRPV VRPV 546 PEEI PEDI PEEI PEEV PEEI
MWVasa MWVasa OVasa XIVasa Consensus Covsensus CsVasa DrVasa GgVasa HsVasa MWVasa	307) 280) 200) 293) 365) 456) 336) 386) 386) 386) 351) 345) 398) 371)	OTTALEPVENTSTILLEGROUNDAGUGSCHARELDELEANDERGINSTRELDEECTIVETERELNGTLEGARESTGUG   OTTALEPVENTSTILLEGROUNDAGUGSCHARELDELETLANDERGINSTRELDEECTIVETERELNGTLEGARESTGUG   OTTALEPVENTSTILLEGROUNDAGUGSCHARELDELETLANDERGINSTRELDEECTIVETERELNGTLEGARESTGUG   OTVALEPVENTSTILLEGROUNDAGUGSCHARELDELETLENDENGTASERSELGEECTIVETERELNGTLEGARESTGUG   OTVALEPVENTSTILLEGROUNDAGUGSCHARELDETLESUMINGTASERSEGIDEEDATITAPTRELNGTLEGARESTGUG   OTVALEPVENTSTILLEGROUNDAGUGSCHARELDETLESUMINGTASERSEGIDEEDATITAPTRELNGTLEGARESTGUG   456 470 480 490 ATPase Motif B 510 520 RNA unwinding Motif   VVFGUNTGVERGVENGUESCONULCATEGRUNDUEGRESUMENCUSSERVENVEVEDEANLANDEGEEDEMERLASEGERSEGROTIMISATE 520 RNA unwinding Motif   VVFGUNTGVERGVENGUEGGERGUEGTEGENOMUERGREGESUMENTERGEDENGUEGENERGUEGENERGUEGUEGENERGUEGUEGENERGUEGUEGENERGUEGUEGENERGUEGUEGENERGUEGUEGUEGENERGUEGUEGUEGENERGUEGUEGUEGUEGUEGUEGUEGUEGUEGUEGUEGUEGUEG	VRAV VRAV VRPV VRPV VRPV VRPV PEEI PEEI PEEI PEEI PEEI
nsvasa MmVasa ONasa XIVasa Consensus Cosvasa DrVasa GgVasa HsVasa MmVasa ONasa VVasa	307) 280) 200) 293) 365) 456) 330) 386) 351) 345) 345) 345) 398) 371) 291) 384)	OTTALLEVENTST Indexturation of tober have tober handled in Approxide to the result of the result	VRAV VRAV VRPV VRPV VRPV PEEI PEEI PEEI PEEI PEEI PEEI PEEI
MmVasa MmVasa OVasa Consensus ( CsVasa DrVasa DmVasa GgVasa HsVasa ONvasa OVasa Consensus Consensus	307) 280) 200) 293) 365) 456) 330) 338) 3351) 345) 398) 3371) 291) 384) 456)	OTTALLEVENTST Indextuaring to some mark is between the indext as the base of the other is by the indext as the	VRAV VRAV VRPV VRPV 546 PEEI PEEI PEEI PEEI PEEI PEEI PEEI PEE
MMVasa MMVasa OVasa Consensus Consensus DrVasa DrVasa GgVasa HsVasa ONvasa XVasa Consensus	307) 280) 200) 293) 365) 456) 338) 338) 3351) 345) 371) 291) 384) 456) -	OFTINLIEVENISTI ILLEGRUDAUGUSUS MARILELILAIMMENDI LASKERELDE BOLT VARTIKEL (NOT LEAKKERFGTC)   OYTKLIEVENISTI ILLEGRUDAUGUSUS MARILELILEILAIMMENDI LASKERELDE BOLT VARTKEL NOT ULEAKKERFGTC)   OYTKLIEVENISTI ILLEGRUDAUGUSUS MARILELI LAIMMENDI LASKERELDE BOLT VARTKEL NOT VELAKKERFGTC)   OYTKLIEVENISTI ILLEGRUDAUGUSUS MARILETILAIMMENDI LASKERELDE BOLT VARTKEL NOT VELAKKERFGTC)   OYTKLIEVENISTI ILLEGRUDAUGUSUS MARILETILETILAMMENDI LASKERELDE BOLT VARTKEL NOT VELAKKERFGTC)   OYTKLIEVENISTI ILLEGRUDAUGUSUS MARILETILETILAIMMENDI LASKERELDE BOLT VARTKEL NOT VELAKKERFGTC)   OYTKLIEVENISTI ILLEGRUDAUGUSUS MARILETILETILAIMMENDI LASKERE VARTKERVENISTI ILLETILAIMENDI LASKERE VARTUNAUKUSUS VARTUNAUKUS PERESKOTIMISATI   0456 470 490 ATPase Motif B 510 520 RNA unwinding Motif   VVVDGVISHEVLUGAUGUSUS VALOUTERIKUUS PERESKOTIMISATY VVDGVISHEVLUGAUKUS PERESKOTIMISATY VVDGVISHTEVLKORVLGATERGULDEVDTETIFEDTERVLVDEAURULMOGESEDMERILVASPERSEROTIMISATY   VVDGVISHEVLOUCUS TERCHUDI ILEKKIGLKOVKYLVUDEAURULDMEGEDMERKLISVENSKEROTIMISATY VVDGVISHTEVLKOCULLETIRGKUDI ILEKKIGLKOVKYLVUDEAURULDMEGESEDMERKLISVENSKEROTIMISATY   VVDGVISHTEVLKOCULLETIRGKUDI ILEKKIGLKOVKYLVUDEAURULDMEGESEMERKLISVENSKEROTIMISATY VVDGVISHTEVLKOGUNULGITIGKKULSVENKYLVUDEAURULDMEGESEMERKLISVENSKEROTIMISATY   VVDGVISHTEVLKOCULLETIRGKUDI ILEKKIGLKOVKYLVUDEAURULDMEGESEMERKLISSEROVSKEROTILLAITIGKKULSKERULVUDEAURULDMEGESEMERKLISSEROVSKEROTILLAITIGKKULSKERULVUDEAURULDMEGESEMERKLISSEROVSKEROTILISATY   VVDGVISHOVISU OGUSULLETIRGK	VRAV VRAV VRPV VRPV 546 PEEI PEEI PEEI PEEI PEEI PEEI PEEI
MmVasa( ONvasa( XIVasa( Consensus( DrVasa( DrVasa( DrVasa( GgVasa( HsVasa( ONv	307) 2280) 2293) 2293) 3365) 456) 3330) 3386) 3351) 3345) 3398) 3371) 291) 384) 456) 547)) 547))	OTTALEPVENTSTILLEGROUPAGUESKIAAFLEDILETLAHMENGTASSKALLEBECTTVATEREL NOTLEARKESFEGT   OTTALEPVENTSTILLEGROUPAGUESKIAAFLEDILETLAHMENGTASSKALLEBECTTVATEREL NOTVEARKESFEGT   OTVALEPVENTSTILLEGROUPAGUESKIAAFLEDILETLSVMINGTASSKALLEBECTTVATEREL NOTVEARKESFEGT   OTVALEPVENTSTILLEGROUPAGUESKIAAFLEDILETLSVMINGTASSKALLEBECTTVATEREL NOTVEARKESFEGT   OTVALEPVENTSTILLEGROUPAGUESKIAAFLEDILSVMINGTASSYLOLEBEAT TATELNOTVEAKKES GTC   OTVALEPVENTSTILLEGROUPAGUESKIAAFLEDILSVMINGTASYLOLEBEAT TATELNOTVEAKKES GTC   1456 470 480 490 ATPase Motif B 520 RNA unwinding Motif   VVMSGINTGVSLSKCNULGTEGTINGHUNGEREVGLSKVQYFVDEATINLOMGFEDEMKLIVASPGEPSREGROTIM SATF 520 RNA unwinding Motif   VVMSGINTGVSLSKCNULGTEGTINGHUNGTEREVGLSKVQVFVDEATINLOMGFEDEMKLIVASPGEPSREGROTIM SATF 531 520 RNA unwinding Motif   VVMSGINTGVSLSKCNULGTEGTEGTUNGTEREKLIGLSKVRVVDVDEAGMMLDMGFEDEMKLISSPGEPSREGROTIM SATF 531 532 RNA unwinding Motif   VVMSGINTGVSLSKCNULGTEGTEGTUNGTEKKISLOVSVVDVDEAGMMLDMGFEDEMKLISSPGEPSREGROTIM SATF 531 532 RNA unwinding Motif   VVMSGINTEGVSLSKCNULGTEGTEGENULTITEKKISLOVSVVDVDEAGMMLDMGFEDEMKLISSPGEPSKERGTIM SATF 531 531 531 531 531 531 531 531 531 531 531 531 531 531 531 531	VRAV VRAV VRPV VRPV VRPV 546 PEEI PEEI PEEI PEEI PEEI PEEI PEEI 637
MmVase( ON/ase( XIVase) Consensus( Consensus( DrVase) DrVase( GgVase( HsVase) MmVase( ON/ase( Consensus( Consensus( CSVase( CVase( DrVase(	307) 280) 200) 293) 365) 3365) 3386) 3386) 3351) 3383) 371) 291) 384) 456) 547) 421) 477)	OF TALLEY WRITEL TARGEDUMA COUSSIN MALE DE LAMMENDOL TASPER SLOP BOLL WAT HED (WOLLD ARKEFYSTOL GYTKLEY VOKYSTE TULAGEDUMA CAOTOSCKI AAFLLET LAMMENDOL TASPER SLOP BOLT WAT HED (WOLLD ARKEFYSTOL GYVK TEY VOKHSTELISAGRDIMA CAOTOSCKI AAFLLET LOULAD OVASRESSIDE BOLT WAT HED (NOT YCH ARKEFYSTOL GYVK TEY VOKHSTELISAGRDIMA CAOTOSCKI AAFLLET LEVING UT AD ANALYSTOL DE PEAT TA PTREL NOT YCH ARKEFYSTOL GYVK TEY VOKHSTELISAGRDIMA CAOTOSCKI AAFLLET LEVING UT AD ARKEFYSTOL GYVK TEY VOKHSTELISAGRDIMA CAOTOSCKI AAFLLET LEVING UT AD ANALYSTOL DE PEAT TA PTREL NOT YCH ARKEFYSTOL GYVK TEY VOKHSTELISAGRDIMA CAOTOSCKI AAFLLET LEVING UT AD ANALYSTOL DE PEAT TA PTREL NOT YCH ARKEFYSTOL GYVK TEY VOKHSTELISAGRDIMA CAOTOSCKI AAFLLET LEVING UT AD ANALYSTOL DE PEAT TA PTREL NOT YCH ARKEFYSTOL 350 470 480 490 ATPASE MOTIF B 510 520 RNA unwinding Motif VVYGGUNTGYT REVEK GOVLCA TEGRIM MERGKVGLSKVQYFV DEATMLIMGFE PEMRKLVAS PGM SREGROTIM SAT F VVYGGUNTGYT REVEK GOVLCA TEGRILD FURGKI SLVEVKYLV DEATMLIMGFE DEMKKLISSFER VOKTON SAT VVYGGUNG SREGONIC LOT TEGRILD FURGKI SLVEVKYLV DEATMLIMGFE GEMKKLISSFER VING TA SEGNE SKER OTT MISAT VVYGGUNG SREGONIC LOT TEGRILD FURGKI SLVEVKYLV DEATMLIMGFE GEMKKLISSFER VING SCH VOKTON VVYGGUNG SREGONIC TA TEGRILD TICKEK IGLKVKYLV DEATMLIMGFE GEMKKLISSFER OFT MISAT VVYGGUNG SKER VOKTOCH TEGRILD TICKEK IGLKVKYLV DEATMLIMGFE GEMKKLISSFER OFT MISAT VVYGGUNG SCH ILCA TEGRILD TICKEK IGLKVKYLV DEATMLIMGFE GEMKKLISSFER OFT MISAT VVYGGUNG VOCONILCA TEGRILD TICKEK IGLKVKYLV DEATMLIMGFE GEMKKLISSFER OFT MISAT VVYGGUNG SCH ILCA TEGRILD TICKEK IGLKVKYLV DEATMLIMGFE FEITHEV OFT SKER OTT MISAT VVYGGUNG SCH ILCA TEGRILD TICKEK IGLKVKYLV DEATMLIMGFE FEITHER VERKER TEKTON VVYGGUNG SCH ILCA TEGRILD TICKEK IGLSK VYLVIDEATMLIMGFE POMKLISSFER OTT MISAT VVYGGUNG SCH ILCA TEGRILD TICKEK IGLSK VYLVIDEATMLIMGFE POMKRING SCH SKER OTT MISAT VVYGGUNG SCH SVC VOCON TICKYCGUNG SCH TET TUG TEGNE S47 560 570 580 590 600 610 620 0KYAADE VAND TELAVGVYGGAS SVC TET VOYD VYKSK CLUCHTEN TOT TEGNER VVYGGUNG TAR THE TOWK TE	VRAV VRPV VRPV 546 PEEI PEDI PEEI PEEI PEEI PEEI PEEI 637 QWQR QREP
nsvase, MmVasc, OVasc, XIVasc, Consensu; CsVasc, GyVasc, GyVasc, GyVasc, MmVasc, OVasc, XIVasc, Consensu; Consensu; MmVasc, DrVasc, DrVasc, DrVasc, DrVasc, DrVasc, DrVasc, DrVasc,	307) 280) 200) 293) 365) 456) 330) 386) 3351) 345) 398) 371) 291) 384) 456) 547) 4421) 4477) 440)	1 1	VRAV VRAV VRPV VRPV VRPV VRPV VRPV S46 PEEI PEEI PEEI PEEI 637 QWQR R 637 QWQR R QRER QSQR
nsvase, MmVasc, VNasc, XNasa( Consensus) ( CsVasc, GgVasc, GgVasc, Consensus) ( Consensus) ( Consensus) ( Consensus) ( Consensus) ( CsVasc, Consensus) ( CsVasc, Consensus) ( CsVasc, Consensus)	307) 2280) 2293) 2293) 3365) 4566) 3386) 3371) 3383) 3371) 3384) 4569) 5477) 4421) 4477) 4440) 4440)	1 1	VRAV VRAV VVRPV VVRPV VVRPV VVRPV VVRPV VRPV
nsvidse, MmVasz, Olvasz, XiVaszi, Consensus, ( CsVasz, DrVasz, Olvasz, ViVasz, Consensus	307) 280) 2200) 2293) 365) 456) 330) 386) 3351) 345) 398) 371) 291) 384) 456) 547) 4421) 447) 440) 4480) 4480)	OF THE LEVENTS IT UNDERDUNATION OF DEPENDENCE OF THE DEPENDENCE OF TARGER SUCCESSOR OF THE DEPENDENCE OF TH	VRAV VRAV VRPV VRPV VRPV VRPV VRPV VRPV
nsvidse, MmVasz, OVasz, XiVasz, Consensus, DrVasz, DrVasz, DrVasz, MmVasz, Consensus, Consensus, Consensus, Consensus, MmVasz,	307) 280) 293) 293) 293) 365) 456) 3386) 3385) 3385) 3385) 3371) 291) 3384) 456) 5471) 445) 5472) 440) 4436) 4489) 382)	10 TYPE LEVENTS 10	VRAV VRAV VVRPV VVRPV VVRPV VVRPV S46 PEEI PEEI PEEI PEEI PEEI 637 QWQR QRER QSQR QRER QQRER QQRER QQRER
nsvidse, MmVasc, OVasc, XiVasc, Consensus; Consensus; DrVasc, OVasc, Nvasc, DrVasc, DrVasc, DrVasc, DrVasc, DrVasc, MmVasc, DrVasc, Nva	307) 280) 293) 293) 293) 365) 456) 3386) 3383) 3351) 345) 3398) 3371) 291) 384) 456) 547) 440) 4477) 440) 440) 440) 4489) 4462) 382)	10 THE LEVENS TO TURNED TO STORE TABLE DE LAMMENDO I MASTRE LOP BOLL WAT HERE IN UT LEARNES FOR CONSTRUCTIVE TREE IN TURNED TO THE ARKES FOR CONSTRUCTIVE TO AND THE DATA TO THE ARKES FOR CONSTRUCTIVE TO AND THE DATA TO THE ARKES FOR CONSTRUCTIVE TO AND THE DATA TO THE ARKES FOR CONSTRUCTIVE TO AND THE DATA TO THE ARKES FOR CONSTRUCTIVE TO AND THE DATA TO THE ARKES FOR CONSTRUCTIVE TO AND THE ARKES FOR CONSTRUCTIVE AND AND THE ARKES FOR CONSTRUCTIVE AND AND THE ARKES FOR CONSTRUCTIVE AND THE ARKES FOR CONSTRUCTIVE AND	VRAV VRAV VVRPV VVRPV VVRPV VVRPV S46 PEEI PEEI PEEI PEEI PEEI 637 QWQR QRER QRER QRER QRER QRER QRER QRER
nsvase, OVase, OVase, XIVasa, Consensu; CSVasa, GyVasa, GyVasa, OVasa, XIVasa, Consensu; Consensu; GyVasa, GyVasa, GyVasa, GyVasa, GyVasa, GyVasa, Consensu;	307) 2280) 2200) 2293) 3365) 456) 3386) 3386) 3385) 3385) 33980) 33980) 33980) 33980) 33980) 33980) 33980) 33980) 339	10 TINLERVENTSTEVLAGENDRAGUTOSON MARTEDE LIAMMENDOLTASSER SELDE BOLT WARTED INDIVIDENTED INDIVIDUATES AND SENSION OF THE SENS	VRAV VRAV VVRPV VVRPV VVRPV VVRPV VVRPV State PEEI PEEI PEEI PEEI G37 GWQR QRER QRER QRER QRER QRER QRER QRER Q
nsvidse, MmVasc, OVasc, XIVasc, Consensus, CsVasc, DrVasc, GgVasc, GgVasc, Consensus, Consensus, Consensus, Consensus, Consensus, Carsensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, XIVasc, Consensus, Consensus, XIVasc, Consensus, XIVasc, Consensus, XIVasc, Consensus, XIVasc, Consensus, XIVasc, Consensus, XIVasc, Consensus, XIVasc, Consensus, XIVasc, Consensus, XIVasc, Consensus, XIVasc, XXIVasc, XXIVasc, XXIVasc, XXIVasc, XXIVasc, XXIVasc, XXIVasc, XXIVasc, XXIVasc, XXIVasc, XXIVasc, XXIVasc, XXIVasc, XXIVasc, XXIVasc, XXIVasc, XXIVASC, X	3007) 2280) 2200) 2293) 3365) 456) 3380) 3383) 345) 3371) 3345) 3371) 3345) 3371) 3345) 3371) 3384) 456) 5477) 436) 440) 440) 440) 4420) 4382) 4455) 5547) 6381)	10 11 14 <td< th=""><th>VRAV VRAV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VRAV VRA</th></td<>	VRAV VRAV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VVRPV VRAV VRA
nsvidse, MmVasc, OVasc, XtVasc, Consensus, DrVasc, DrVasc, DrVasc, MmVasc, GgVasc, Consensus, Consensus, Consensus, Consensus, MMVasc, DrVasc, XtVasc, Consensus,	307)) 2280) 2280) 2293) 365) 456) 336) 3351) 3351) 3351) 3371) 3371) 3384) 3453) 3371) 3384) 4452) 547) 4489) 4462) 3382) 4489) 4462) 3382) 4455) 547) 638) 5547) 638) 5512)	1 1	VRAV VRAV VRAV VRPV VRPV VRPV VRPV VRPV
nsvidse, MmVasz, OVasz, Zonsensut; Consensut; DrVasz, DrVasz, MmVasz, Consensut; ( CSVasz, Consensut; ( CSVasz, DrVasz, DrVasz, Consensut; ( CSVasz, CSVasz, Consensut; ( CSVasz,	307)) 2280) 2280) 2293) 365) 456) 336) 3351) 3383) 3371) 3384) 456) 547) 4384) 4450) 384) 4450) 384) 4450) 382) 4475) 5547) 638) 5547) 638) 5547) 638) 5527	OTTALEPVENTSTITUTESKUDANGUSSENARTUDTISMETELDTIAMMENDETASSERUDE BOLT WATRED INGTIGENETSTO   OTTALEPVENTSTITUTESKUDANGUSSENARTUDTISMENELDTIASMENELDE BOLT WATRED INGTIGENETSTO   OTVALEPVENTSTITUTESKUDANGUSSENARTUDTISMENELDTIASMENELDE BOLT WATRED INGTIGENERSFERC   OTVALEPVENTSTITUTESKUDANGUSSENARTUDTISMENELDTISMENENDE PARTITAPTRELINGTUDERKESSTOC   OTVALEPVENTSTITUTESKUDANGUSSENARTUDTISMENENS   1456 470 480 490 ATPase Motif B 510 520 RNA unwinding Motif   1970 VERSULATION AND AND AND AND AND AND AND AND AND AN	VRAV VRAV VRPV VRPV VRPV VRPV VRPV VRPV
nsvidse, MmVasz OVasz, Zonsensus; Consensus; DrVasz OVasz, GgVasz, HsVasz OVasz, OVasz, DrVasz DrVasz DrVasz OVasz, SVasz OVasz, Consensus; Con	307)) 2280) 2280) 2280) 2293) 3365) 3365) 3386) 3383) 3383) 33930) 33930) 33930) 33930) 33930) 33930) 33930) 33930) 3	1 STRLEPVENTSTETULAGEDUNALOUSSEN MALELDELLAMMENGETASSERELDE BOLT WARTED. INVELTED INTERED. INVELSENTED.   1 STRLEPVENTSTETULAGEDUNALOUSSEN MALELDELLAMMENGETASSERELDE BOLT WARTED. INVELSENTED.   1 SYNCHSTETULAGEDUNALOUSSEN MALELDELLAMMENGETASSERELDE BOLT WARTED. INVELSENTED.   1 SYNCHSTETULAGEDUNALOUTSSEKTAAFLLETLSTAMMENGETASSERELDE BOLT WARTED. INVELSENTED.   1 SYNCHSTETULAGEDUNALOUTSSEKTAAFLLETUN WARTED TASSERELDE BOLT WARTED. INVELSENTED.   1 SSNCHSTETULAGEDUNALOUTSSEKTAAFLLETUN WARTED TASSEREGETUN STRLEPTED.   1 SSNCHSTETULAGEDUNALOUTSGEKTAAFLLETUN WARTEDE TASSEREGUNALOUTSEREGENTAL SSNCHSTUNGENTED.   1 SSNCHSTETULAGENULAGETEGEKTAMENTERGEVGLSKVOLFVIDEALMILDMGFEDMIRLUASSEGNSEGGOTIM.   1 SSNCHSTETULAGENULAGETEGEKUNDETETTETEDTERTEVVLOEDALMILDMGFEDMIRLUASSEGNSEGGOTIM.   1 VINCOTSENTIONECTINGENULATIONENTERGEVGLSKVOLVUDEALMILDMGFEDMIRLUASSEGNSEGOTIM.   1 VINCOTSENTIONECTINGENULATIONENTIERGEVGLSKVOLVUDEALMILDMGFEDMIRLUASSEGNSEGOTIM.   1 VINCOTSENTIONECTINGENULATIONENTIERGEVGLSKVOLVUDEALMILDMGFEDMIRLUASSEGNSEGOTIM.   1 VINCOTOTENSIROIMOCOTILCATEGRILDMITIEKKISIVEVKILVUDEALMILDMGFEDMIRLUSSEGNSEGOTIM.   1 VINCOTOTENSIROIMOCOTILCATEGRILDMITIEKKISIVEVKILVUDEALMILDMGFEDMIRLUSSEGNERKLISSEROTIM.   1 VINCOTOTENSIROIMOCOTILCATEGRILDMITIEKKISIVEVKILVUDEALMILDMGFEDMIRLUSSEGOTSEGOTIM.   1 VINCOTOTENSIROIMOCOTILCATEGRILDMITIEKKISIVEVKILVUDEALMILDMERGEDMIRLUSSEGOTILSSETTUNGSEGONGENTENTSTETUTENTERGENTENTED.   1 VINCOTOFENSTROIMOCOTILCATEGRILDMITIEKKISISVEVKILVUDEALMILDMERGEDMIRLUSSEGONGENTENTSTERGENT	VRAV VRAV VRPV VRPV 546 PECI PECI PECI PECI PECI PECI PECI 637 QWOR QRER QRER QRER QRER QRER QRER QYOR R QRER QYOR R QYOR R QYOR PECI S VPPW VPV VPV VPV VPV VPV VPV VPV VPV VP
nsvidse, MmVasc, OVasc, XIVasc, Consensu; CSVasc, DrVasc, GyVasc, GyVasc, Consensu; Co	307)) 2280) 2280) 2280) 2293) 3365) 3365) 3386) 3371) 3398) 3371) 3291) 3384) 3371) 2291) 3384) 3371) 2291) 3384) 4450) 5470) 4420) 4420) 4420) 3820) 4475) 5547) 5547) 5547) 55580) 5527) 55280)	10 TINLERVENTSTETULAGENDINAGUTOSUNAALUDI LIAHMENDOLIASSENELUDE BCUTUVETRELINOTYLEARKESFORD   10 YTKLEFVOKYSTETULAGENDINAGUTOSUKAAFLUPTLOQLMADEVASSESTIDE BCUTUVETRELINOTYLEARKESFORD   10 YVKLEFVOKKSTETULAGENDINAGUTOSUKAAFLUPTLOQLMADEVASSESTIDE BCUTUVETRELINOTYLEARKESFORD   11 YVKLEFVOKKSTETULAGENDINAGUTOSUKAAFLUPTLOQLMADEVASSESTIDE BCUTUVETRELINOTYLEARKESFORD   12 SVKLEFVOKKSTETULAGENDINAGUTOSUKAAFLUPTLSYMMEGITASOYLOLE BEATITAPTRELINOTYLEARKESFORD   1456 470 480 490 ATPase Motif B 510 520 RNA unwinding Motif   1456 470 480 490 ATPase Motif B 510 520 RNA unwinding Motif   1547 70 480 490 ATPase Motif B 510 520 RNA unwinding Motif   1547 1540 490 ATPase Motif B 510 520 RNA unwinding Motif   1547 1540 1540 1540 1540 1540 1540 1540   1547 1540 1540 1540 1540 1540 1540 1540 1540 1540 1540 1540 1540 1540 1540 1540 1540 1540 1540	VRAV VRAV VRPV VRPV 546 PECI PECI PECI PECI PECI PECI PECI 637 QWQR QRCR QRCR QRCR QRCR QRCR QRCR QRCR
nsvase, MmVasc, OVasc, XtVasc, Consensus, DvVasc, DvVasc, DvVasc, DvVasc, MmVasc, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, MmVasc, Consensus, MmVasc, MmVasc, Status, DmVasc, Consensus, MmVasc, Status, DmVasc, Consensus, Consensus, Consensus, MmVasc, Status, Status, MmVasc, Status, MmVasc, Status, Status, Status, Consensus, Consensus, Consensus, MmVasc, Status, Status, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, MmVasc, Status, Status, Consens, Consensus, Consensus, Consensus, Consensus,	307)) 2280) 2280) 2293) 3365) 3365) 3386) 3351) 3345) 3371) 291) 3383) 371) 291) 3383) 371) 291) 3456) 547) 4436) 4436) 4436) 4436) 4475) 547) 638) 5547) 5547) 55580) 5527) 5580)	1 1	VRAV VRAV VRAV VRPV VRPV VRPV VRPV State State VRPE VRPE VRE VRE VRE VRE VRE VPF VPF VPF VPF VPF VPF VPF VPF
nsvidse, MmVasz, OVasz, Zonsensut; Consensut; DrVasz, DrVasz, MmVasz, Consensut; ( CSVasz, Consensut; ( CSVasz, DrVasz, DrVasz, DrVasz, DrVasz, DrVasz, Consensut; ( CSVasz, Consensut; ( CSVasz, DrVasz, DrVasz, Consensut; ( CSVasz, Consensut; ( CSVasz, Consensut; ( CSVasz, DrVasz, Consensut; ( CSVasz, DrVasz, Consensut; ( CSVasz, CONsensut; ( CSVasz, CSVasz, ( CSVasz, CSVasz, ( CSVasz, CSVasz, ( CSVasz,	307)) 2280) 2280) 2293) 3365) 3365) 3365) 3386) 3371) 291) 3383) 3371) 291) 3383) 3371) 291) 3384) 3371) 291) 3384) 4450) 547) 4436) 4475) 547) 638) 5527) 5580) 5527) 5580) 5527)	1 1	VRAV VRAV VRPV VRPV VRPV VRPV VRPV VRPV
nsvidse, MmVasz, OVasz, ViVasz, Consensus, Consensus, GgVasz, HsVasz, MmVasz, OVasz, Zvasz, DrVasz, DrVasz, DrVasz, DrVasz, DrVasz, DrVasz, DrVasz, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, GgVasz, MmVasz, OVasz, Nvasz, Nvasz, Ovasz, Nvasz, Consensus,	307) 2280) 2280) 2280) 2280) 2293) 3365) 456) 3330) 3351) 3351) 3351) 3351) 3351) 3351) 3351) 3351) 3351) 3371) 291) 4450) 547) 4430) 4436) 547) 4436) 547) 4436) 5547) 55530) 55530) 55530 55530 55530 55530 55530 55530 55530 55530 55530 55530 55530 555530 555530 555530 555530 555530 555530 555530 555530 5555530 5555530 555530 555530 5555570 5555570 555570 5	10 THE LEVENS 11 LAGRODAN AQUISSEN MALE DE LAMMENGE LASSENELUE EQUIVALE THE LINGT LE ARKESFEGTE   10 YTK LEVENS 12 LAGRODAN AQUISSEN MALE DE LAMMENGE LASSENELUE EQUIT VALEEL INDIVERANCES FEGTE   10 YVK PTPVQKYSIE IVLAGRODAN AQUISSEN MALE DE LAIMMENGE LASSENELUE EQUIVALEEL INDIVERANCES FEGTE   10 YVK PTPVQKHSIPIISAGRDIMA AQUISSEN MALE DE LAIMMENGE LASSENELUE EQUIVALEEL INDIVERANCES FEGTE   155 470 480 490 ATPase Motif B 510 520 RNA unvinding Motif   1456 470 480 490 ATPase Motif B 510 520 RNA unvinding Motif   156 470 480 490 ATPase Motif B 510 520 RNA unvinding Motif   156 470 480 490 ATPase Motif B 510 520 RNA unvinding Motif   157 157 168 150 520 RNA unvinding Motif   150 157 150 150 150 150 150   150 150 150 150 150 150 150   150 150 150 150 150 150 150 150 150 <th>VRAV VRAV VRAV VRPV VRPV VRPV VRPV VRPV</th>	VRAV VRAV VRAV VRPV VRPV VRPV VRPV VRPV
nsvidse, MmVasc, OVasc, Consensus; Consensus; Consensus; GyVasc, GyVasc, GyVasc, GyVasc, GyVasc, Consensus; Co	307) 2280) 2280) 2280) 2280) 2280) 2280) 2293) 3365) 3365) 3371) 3381) 3371) 3393) 3371) 231) 3393) 3371) 231) 3393) 3371) 231) 3393) 3371) 231) 3393) 3371) 231) 3393) 3371) 231) 3393) 3371) 231) 3393) 3371) 231) 3393) 3371) 231) 3393) 3371) 231) 3393) 3371) 231) 3393) 3371) 231) 3393) 3371) 231) 3393) 3371) 231) 3393) 3371) 231) 3372) 231) 3372) 231) 3372) 231) 3372) 231) 3372) 231) 3372) 231) 3372) 231) 3372) 231) 3372) 231) 3372) 231) 3372) 231) 231) 3372) 231)	10 11 11 10 <td< th=""><th>VRAV VRAV VRAV VVRPV VVRPV 546 PEEI PEEI PEEI PEEI PEEI 637 QRER QRER QRER QRER QRER QRER QRER QRE</th></td<>	VRAV VRAV VRAV VVRPV VVRPV 546 PEEI PEEI PEEI PEEI PEEI 637 QRER QRER QRER QRER QRER QRER QRER QRE
nsvidse, MmVasc, OVasc, XNasc, Consensus, DrVasc, DrVasc, DrVasc, MmVasc, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, Consensus, XNasc, SNAsc, MmVasc, OVasc, SNAsc, MmVasc, Consensus, Con	300/) 2280) 2280) 2280) 2293) 3365) 456) 3386) 3386) 3351) 3383) 3371) 2391) 3384) 4456) 5371) 2391) 3384) 4456) 5371) 5472) 4473) 5580) 5529) 5229] 5229] 5	1 1	VRAV VRAV VRPV VRPV VRPV VRPV VRPV VRPV
nsvidse, MmVasz, OVasz, Consensut; Consensut; GdyVasz, DrWasz, MmVasz, Consensut; Consen	307) 2280) 2291) 23391) 23391) 23391) 23391) 23391) 23391) 23391) 23391) 23391) 23391) 23391) 23391) 23391) 23391) 23391) 23391) 2340) 4450) 4477) 2547) 2547) 2529) 2	1 1	VRAV VRAV VRPV VVRPV VVRPV S46 PEEI PEEI PEEI PEEI 637 QWOR QRER QRER QRER QRER QRER QRER QRER QR
nsvidse, MmVasc, OVasc, Consensus; Consensus; GgVasc, GgVasc, MmVasc, OVasc, MmVasc, Consensus; Consens; Consens; Consens; Consens; Consens; Consens; Consens; Consen	300/j) 2080) 2280) 2280) 2280) 2280) 2280) 2293) 3365) 3365) 3371) 2291) 3384) 4456) 547) 4456) 547) 4477) 4489) 4477) 4489) 4477) 4489) 4477) 55470 55470 555	STYKLEP VOKSTE FULASROLMA AQUISSIS MARELLET LAMMING IN ASKREDUE ECJ TVAPTREL INCTVOENKESFERGE   STYKLEP VOKSTE FULASROLMA AQUISSIS MARELLET LAMMING IT ASSREDUE ECJ TVAPTREL INCTVOENKESFERGE   STYKLEP VOKSTE FULASROLMA AQUISSIS MARELLET LAMMING IT ASSREDUE ECJ TVAPTREL INCTVOENKESFESTO   STYKLEP VOKSTE FULASROLMA AQUISSIS MARELLET LAMMING IT ASSREDUE ECJ TVAPTREL INCTVOENKESFESTO   STYKLEP VOKSTE FULASROLMA AQUISSIS MARELLET LAMMING ITASSREDUE ECJ TVAPTREL INCTVOENKESFESTO   STYKLEP VOKSTE FULASROLMA AQUISSIS MARILLET LAMMING ITASSREDUE ECJ TVAPTREL INCTVOENKESFESTO   STYKLEP VOKSTE FULASROLMA AQUISSIS MARILLET LAMMING ITASSREDUE ECJ TVAPTREL INCTVOENKESFESTO   STYKLEP VOKSTE FULASROLMA AQUISSIS MARILLET LAMMING INTASSPESTOR ENDITASSREDIE   STYKLEP VOKSTE FULASROLMA AQUISSIS MARILLET LAMMING INTASSRED ETAMINASIS   STYNELT VOKSTE FULASROLMA AQUISSIS MARILLET LAMMING ADUISSIS MARINE AD	VRAV VRAV VRAV VRPV VRPV S46 PEE1 PEE1 PEE1 PEE1 PEE1 PEE1 PEE1 PEE
nsvidse, MmVasc, OVasc, VXvasc, Consensus, GyVasc, GyVasc, GyVasc, GyVasc, Consensus, Co	307) 2280) 2280) 2280) 2280) 2280) 2280) 2280) 2293) 3365) 3365) 3371) 2291) 3384) 4456) 547) 4420) 4436) 4477) 4436) 4477) 4436) 4477) 44477) 44477) 44477) 44477) 44477) 44477) 44477) 44477) 44477) 44477) 44477) 44477) 44477) 44489) 44477) 55470 55470 555	1 1	VRAV VRAV VRPV VRPV VRPV VRPV VRPV VRPV
nsvidse, MmVasc, OVasc, XNasc, Consensus; DrVasc, DrVasc, DrVasc, MmVasc, GgVasc, Consensus; Consens; Consensus; Consens; Consens; Consens; Consens; Consens; Consens; Consens;	300/) 2080) 2291) 2335) 23	1 1	VRAV VRAV VRPV VRPV VRPV VRPV VRPV VRPV
۱۹۷۷ ۵۵۹, ۱۹۷۷ ۵۹۹, ۱۹۷۷ ۵۹۹, ۱۹۷ ۵۹۹, ۱۹۷۷ ۵۹۹, ۱۹۷۹, ۱۹۹۹, ۱۹۷۹, ۱۹۹۹, ۱۹۷۹, ۱۹۹۹, ۱۹۹۹, ۱۹۹, ۱۹۹,	300/) 2080) 2291) 2291) 2568) 25529) 255310 255310 255310 255310	1 1	VRAV VRAV VRPV VRPV VRPV VRPV VRPV VRPV
nsvidse, MmVasz, OVasz, Consensus, Consensus, Gyvasz, Gyvasz, MmVasz, Consensus, Consens	300/) 2280) 2280) 2280) 2293) 3365) 456) 3383) 3383) 3383) 3383) 3383) 3384) 456) 547) 4473) 4475) 4473) 4475) 4473) 4475) 4473) 4475) 5547) 5547) 5547) 5547) 5547) 5547) 55527) 5568) 5527) 5553) 55527) 5566) 6638) 6639) 6639) 6639) 6644) 6657)	1 1	VRAV VRAV VRPV VRPV VRPV VRPV VRPV VRPV

Fig. 3. (continued).

### 10

### ARTICLE IN PRESS

J. Huang et al. / Theriogenology xxx (2014) 1-12



**Fig. 4.** The relative expression of *Csvasa* in various adult tissues. The data represent the mean normalized gene expression levels against the levels of  $\beta$ -actin mRNA  $\pm$  the standard error of mean (SEM) from three separate individuals performed in triplicate (three biological replicates for gonads and six replicates for other tissues). Values with different superscripts differed with statistical significance (P < 0.05), calculated by one-way ANOVA followed by Duncan's multiple range tests using SPSS 17.0.

differentiation collected from this study may aid in the understanding of sex determination in half-smooth tongue sole.

Recently, PGC visualization and labeling in fish embryo has attracted considerable attention. However, a marker to identify PGCs of economically important flatfish species (e.g., tongue sole) is still lacking. Analysis of *vasa* transcriptional regulatory region contributes to the study of germline cells. The 3'UTR of *vasa* has an important function in the stabilization of mRNA; this being that it can eliminate mRNA from somatic cell lines by microRNA activity and stabilize mRNA in PGCs of several fish species. Embryonic injection of mRNA containing 3'UTR and the GFP gene have been used to



Fig. 5. The relative expression of *Csvasa* mRNA in the gonads of neomales, normal males, and females. The data represent the mean  $\pm$  SEM from three separate individuals (n = 3). Values with different superscripts differed at P < 0.05.



**Fig. 6.** The expression profile of *Csvasa* at different developmental stages of the embryos. Each sample consisted of a pool of 30 embryos, a pool of 30 larvae (1 dph). The data represent mean  $\pm$  SEM from three separate samples (n = 3). Values with different superscripts differed at P < 0.05.

visualize and label PGCs in zebrafish [49], rainbow trout [20], medaka [50], and Atlantic salmon [17]. In the present study, GFP-*Csvasa* 3'UTR mRNA was injected into fertilized embryos of medaka. The results indicated that the GFP signal intensity gradually decreased in somatic cells because of the degradation of injected chimeric RNA; however, the intensity became clearer and larger in the presumptive gonad region. This indicated that GFP-*Csvasa* 3'UTR mRNA was specifically stabilized and expressed in PGCs of medaka during embryogenesis. These data confirmed the function of *Csvasa* 



**Fig. 7.** The expression profile of *Csvasa* at different developmental stages of the gonads. The data represent mean  $\pm$  SEM from three separate samples (n = 3). Each sample consisted of a pool of six body trunks (4–66 dph) and a pool of three gonads (80–150 dph). Values with different superscripts differed at P < 0.05, calculated by two-way ANOVA followed by Duncan's multiple range tests using SPSS 17.0. Asterisks show statistical significance between sexes (\* P < 0.05; \*\* P < 0.01) at each sampling time after a Student's *t*-test.

J. Huang et al. / Theriogenology xxx (2014) 1-12



Fig. 8. Green fluorescence protein–labeled PGCs observed by microinjection of GFP-Csvasa 3'UTR mRNA in medaka embryos. PGCs are denoted by white arrows. (A–D) embryos of stage 20, stage 22, stage 25, and stage 29 were observed under light and fluorescence microscopy.

3'UTR, and will be helpful in the study of PGC visualization, cryopreservation, and transplantation for tongue sole.

In conclusion, the *Csvasa* gene was cloned and characterized in the present study. *Csvasa* mRNA was a maternally supplied transcript during the embryonic stage, and it was mainly observed in gonads of adult fish. Based on the accurate distinction of sex and gonads, typical sexually dimorphic expression of *Csvasa* was detected by qPCR during early development and sex differentiation. The PGC labeling function of *Csvasa* 3'UTR was verified in medaka embryos. Our results confirmed that the *Csvasa* gene may possibly be appropriate as a germ cell marker. In future studies, the *Csvasa* gene will be used as a molecular marker to illuminate the processes of PGC formation, migration and differentiation, and investigate PGC biotechnology for tongue sole reproduction.

### Acknowledgments

This study was supported by grants from State 863 High-Technology R&D Project of China (2012AA092203, 2012AA10A403-2), National Nature Science Foundation of China (31130057), and Special Scientific Research Funds for Central Non-profit Institutes, Chinese Academy of Fishery Sciences (2013A0402).

### **Competing interests**

The authors declare that there are no conflicts of interest.

### References

- [1] Wylie C. Germ cells. Cell 1999;96:165-74.
- [2] Siegfried KR, Nüsslein-Volhard C. Germ line control of female sex determination in zebrafish. Dev Biol 2008;324:277–87.

- [3] Xu HY, Li MY, Gui JF, Hong YH. Fish germ cells. Sci China Life Sci 2010;53:435–46.
- [4] Yoon C, Kawakami K, Hopkins N. Zebrafish vasa homologue RNA is localized to the cleavage planes of 2- and 4-cell-stage embryos and is expressed in the primordial germ cells. Development 1997;124: 3157–65.
- [5] Schupbach T, Wieschaus E. Germline autonomy of maternal-effect mutations altering the embryonic body pattern of *Drosophila*. Dev Biol 1986;113:443–8.
- [6] Ikenishi K. Germ plasm in Caenorhabditis elegans, Drosophila and Xenopus. Dev Growth Differ 1998;40:1–10.
- [7] Noce T, Okamoto-Ito S, Tsunekawa N. Vasa homolog genes in mammalian germ cell development. Cell Struct Funct 2001;26:131–6.
- [8] Li CJ, Liu L, Chen XH, Zhang T, Gan F, Cheng BL. Identification of a vasa homologue gene in grass carp and its expression pattern in tissues and during embryogenesis. Comp Biochem Physiol B Biochem Mol Biol 2010;157:159–66.
- [9] Blázquez M, González A, Mylonas CC, Piferrer F. Cloning and sequence analysis of a vasa homolog of the European sea bass (*Dicentrarchus labrax*): tissue distribution and mRNA expression levels during early development and sex differentiation. Gen Comp Endocrinol 2011;170:322–33.
- [10] Wu X, Wang Z, Jiang J, Gao J, Wang J, Zhou X, et al. Cloning, expression promoter analysis of vasa gene in Japanese flounder (*Paralichthys olivaceus*). Comp Biochem Physiol B Biochem Mol Biol 2014;167:41–50.
- [11] Pacchiarini T, Cross I, Leite RB, Gavaia P, Ortiz-Delgado JB, Pousão-Ferreira P, et al. Solea senegalensis vasa transcripts: molecular characterisation, tissue distribution and developmental expression profiles. Reprod Fertil Dev 2013;25:646–60.
- [12] Shinomiya A, Tanaka M, Kobayashi T, Nagahama Y, Hamaguchi S. The vasa-like gene, olvas, identifies the migration path of primordial germ cells during embryonic body formation stage in the medaka, Oryzias latipes. Dev Growth Differ 2000;42:317–26.
- [13] Yoshizaki G, Sakatani S, Tominaga H, Takeuchi T. Cloning and characterization of a *vasa*-like gene in rainbow trout and its expression in the germ cell lineage. Mol Reprod Dev 2000;55:364–71.
- [14] Kobayashi T, Kajiura-Kobayashi H, Nagahama Y. Differential expression of vasa homologue gene in the germ cells during oogenesis and spermatogenesis in a teleost fish, tilapia, Oreochromis niloticus. Mech Dev 2000;99:139–42.
- [15] Presslauer C, Nagasawa K, Fernandes JM, Babiak I. Expression of vasa and nanos3 during primordial germ cell formation and migration in Atlantic cod (*Gadus morhuaL*.). Theriogenology 2012; 78:1262–77.
- [16] Lin F, Xu S, Ma D, Xiao Z, Zhao C, Xiao Y, et al. Germ line specific expression of a vasa homologue gene in turbot (Scophthalmus)

J. Huang et al. / Theriogenology xxx (2014) 1-12

*maximus*): evidence for *vasa* localization at cleavage furrows in euteleostei. Mol Reprod Dev 2012;79:803–13.

- [17] Nagasawa K, Fernandes JM, Yoshizaki G, Miwa M, Babiak I. Identification and migration of primordial germ cells in Atlantic salmon, *Salmo salar*: characterization of *vasa*, *dead end*, and *lymphocyte antigen* 75 genes. Mol Reprod Dev 2013;80:118–31.
- [18] Knaut H, Steinbeisser H, Schwarz H, Nüsslein-Volhard C. An evolutionary conserved region in the vasa 3'UTR targets RNA translation to the germ cells in the zebrafish. Curr Biol 2002;12:454–66.
- [19] Tanaka M, Kinoshita M, Kobayashi D, Nagahama Y. Establishment of medaka (*Oryzias latipes*) transgenic lines with the expression of green fluorescent protein fluorescence exclusively in germ cells: a useful model to monitor germ cells in a live vertebrate. Proc Natl Acad Sci USA 2001;98:2544–9.
- [20] Yoshizaki G, Tago Y, Takeuchi Y, Sawatari E, Kobayashi T, Takeuchi T. Green fluorescent protein labeling of primordial germ cells using a nontransgenic method and its application for germ cell transplantation in Salmonidae. Biol Reprod 2005;73:88–93.
- [21] Saito T, Fujimoto T, Maegawa S, Inoue K, Tanaka M, Arai K, et al. Visualization of primordial germ cells in vivo using GFP-nos1 3'UTR mRNA. Int J Dev Biol 2006;50:691–9.
- [22] Lin F, Liu Q, Li M, Li Z, Hong N, Li J, et al. Transient and stable GFP expression in germ cells by the vasa regulatory sequences from the red seabream (*Pagrus major*). Int J Biol Sci 2012;8:882–90.
- [23] Kobayashi T, Takeuchi Y, Takeuchi T, Yoshizaki G. Generation of viable fish from cryopreserved primordial germ cells. Mol Reprod Dev 2007;74:207–13.
- [24] Lacerda SM, Costa GM, Campos-Junior PH, Segatelli TM, Yazawa R, Takeuchi Y, et al. Germ cell transplantation as a potential biotechnological approach to fish reproduction. Fish Physiol Biochem 2013; 39:3–11.
- [25] Ji XS, Chen SL, Jiang YL, Xu TJ, Yang JF, Tian YS. Growth differences and differential expression analysis of pituitary adenylate cyclase activating polypeptide (PACAP) and growth hormone-releasing hormone (GHRH) between the sexes in half-smooth tongue sole *Cynoglossus semilaevis*. Gen Comp Endocrinol 2011;170:99–109.
- [26] Chen SL, Li J, Deng SP, Tian YS, Wang QY, Zhuang ZM, et al. Isolation of female-specific AFLP markers and molecular identification of genetic sex in half-smooth tongue sole (*Cynoglossus semilaevis*). Mar Biotechnol (NY) 2007;9:273–80.
- [27] Deng SP, Chen SL, Xu JY, Liu BW. Molecular cloning, characterization and expression analysis of gonadal P450 aromatase in the halfsmooth tongue-sole, *Cynoglossus semilaevis*. Aquaculture 2009; 287:211–8.
- [28] Chen SL, Ji XS, Shao CW, Li WL, Yang JF, Liang Z, et al. Induction of mito-gynogenetic diploids and identification of WW super-female using sex-specific SSR markers in half-smooth tongue sole (*Cynoglossus semilaevis*). Mar Biotechnol (NY) 2012;14:120–8.
- [29] Kurokawa H, Saito D, Nakamura S, Katoh-Fukui Y, Ohta K, Baba T, et al. Germ cells are essential for sexual dimorphism in the medaka gonad. Proc Natl Acad Sci USA 2007;104:16958–63.
- [30] Lewis ZR, McClellan MC, Postlethwait JH, Cresko WA, Kaplan RH. Female-specific increase in primordial germ cells marks sex differentiation in threespine stickleback (*Gasterosteus aculeatus*). J Morphol 2008;269:909–21.
- [31] Fujimoto T, Nishimura T, Goto-Kazeto R, Goto-Kazeto R, Kawakami Y, Yamaha E, Arai K. Sexual dimorphism of gonadal structure and gene expression in germ cell-deficient loach, a teleostfish. Proc Natl Acad Sci USA 2010;107:17211–6.
- [32] Goto R, Saito T, Takeda T, Fujimoto T, Takagi M, Arai K, et al. Germ cells are not the primary factor for sexual fate determination in goldfish. Dev Biol 2012;370:98–109.

- [33] Chen SL, Zhang GJ, Shao CW, Huang Q, Liu G, Zhang P, et al. Wholegenome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. Nat Genet 2014;46:253–60.
- [34] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011;28:2731–9.
- [35] Li ZJ, Yang LJ, Wang J, Shi W, Pawar RA, Liu Y. β-Actin is a useful internal control for tissue-specific gene expression studies using quantitative real-time PCR in the half-smooth tongue sole *Cynoglossus semilaevis* challenged with LPS or *Vibrio anguillarum*. Fish Shellfish Immunol 2010; 29:89–93.
- [36] Hay B, Jan LY, Jan YN. A protein component of *Drosophila* polar granules is encoded by *vasa* and has extensive sequence similarity to ATP-dependent helicase. Cell 1988;55:577–87.
- [37] Shibata N, Umesono Y, Orii H, Sakurai T, Watanabe K, Agata K. Expression of vasa(vas)-related genes in germline cells and totipotent somatic stem cells of planarians. Dev Biol 1999;206:73–87.
- [38] Chen CF, Wen HS, Chen XY, Zhang JR, Jin GX, Shi B, et al. Studies on ovarian development and spawn type of cultured half-smooth tongue sole, *Cynoglossus semilaevis*. Mar Sci 2010;34:29–34.
- [39] Song HX, Weng YZ, Liu ZG, Fang YQ, A histological study on spermatogenesis in *Cynoglossus semilaevis*. J Oceanogr Taiwan Strait 2009;28:19–24.
- [40] Braat AK, Speksnijder JE, Zivkovic D. Germ line development in fishes. Int J Dev Biol 1999;43:745–60.
- [41] Jia R, Nie LW, Wang N, Wang JJ. Molecular cloning and expression patterns of the *vasa* gene from *Rana nigromaculata* (Amphibia: Anura). Zoologia 2009;26:316–22.
- [42] Ikenishi K, Tanaka TS. Spatio-temporal expression of *Xenopus vasa* homolog, *XVLG1*, in oocytes and embryos: the presence of *XVLG1* RNA in somatic cells as well as germline cells. Dev Growth Differ 2000;42:95–103.
- [43] Raghuveer K, Senthilkumaran B. Cloning and differential expression pattern of vasa in the developing and recrudescing gonads of catfish, *Clarias gariepinus*. Comp Biochem Physiol A Mol Integr Physiol 2010;157:79–85.
- [44] Liang Z, Chen SL, Zhang J, Song WT, Liu SS. Gonadal development process observation of half-smooth tongue sole in rearing population. J South Agric 2012;43:2074–8.
- [45] Kobayashi T, Kajiura-Kobayashi H, Nagahama Y. Two isoforms of vasa homologs in a teleost fish: their differential expression during germ cell differentiation. Mech Dev 2002;111:167–71.
- [46] Krøvel AV, Olsen LC. Sexual dimorphic expression pattern of a splice variant of zebrafish vasa during gonadal development. Dev Biol 2004;271:190–7.
- [47] Ma X, Liu X, Wen H, Xu Y, Zhang L. Histological observation on gonadal sex differentiation in *Cynoglossus semilaevis*. Mar Fisheries Res 2006;27:55–60.
- [48] Nagler J, Cavileer T, Hunter S, Drew R, Okutsu T, Sakamoto T, et al. Non-sex specific genes associated with the secondary mitotic period of primordial germ cell proliferation in the gonads of embryonic rainbow trout (Oncorhynchus mykiss). Mo Reprod Dev 2012;78:181–7.
- [49] Koprunner M, Thisse C, Thisse B, Raz E. A zebrafish nanos-related gene is essential for the development of primordial germ cells. Genes Dev 2001;15:2877–85.
- [50] Kurokawa H, Aoki Y, Nakamura S, Ebe Y, Kobayashi D, Tanaka M. Time-lapse analysis reveals different modes of primordial germ cell migration in the medaka *Oryzias latipes*. Dev Growth Differ 2006;48: 209–21.