

Electronic Supplementary Material S1: Similarity of *Sox2* gene between *O. niloticus* and several vertebrates based on Blast analysis at NCBI.

Accession number	Nucleotide sequence information	Match score
EF431921.1	<i>Oreochromis aureus Sox2 (Sox2) gene, partial cds</i>	100%
FJ432695.1	<i>Epinephelus coioides</i> cell-line GBC4 <i>Sox2</i> mRNA, partial cds	94%
GU001787.1	<i>Trachinotus blochii Sox2</i> mRNA, partial cds	94%
FJ895588.1	<i>Oryzias latipes</i> SRY-box containing gene 2 (<i>Sox2</i>) mRNA, partial cds	93%
AY277952.1	<i>Takifugu rubripes</i> transcription factor <i>Sox2 (Sox2) gene, complete cds</i>	93%
EF431922.1	<i>Oreochromis karongae Sox2 (Sox2) gene, partial cds</i>	99%
EF431923.1	<i>Oreochromis mossambicus Sox2 (Sox2) gene, partial cds</i>	99%
DQ900581.1	<i>Gadus morhua Sox2</i> mRNA, partial cds	89%
AB242329.1	<i>Danio rerio Sox2</i> mRNA for transcription factor <i>SoxSox2</i> , complete cds	82%
NM_001173447.1	<i>Felis catus</i> sex determining region Y-box 2 (<i>SoxSox2</i>) mRNA, complete cds	83%
XM_545216.2	PREDICTED: <i>Canis familiaris</i> similar to Transcription factor <i>SoxSox-2 (LOC488092)</i> , mRNA	82%
XM_002807565.1	PREDICTED: <i>Callithrix jacchus</i> transcription factor <i>SoxSox-2-like (LOC100407856)</i> , mRNA	82%
XM_002716451.1	PREDICTED: <i>Oryctolagus cuniculus</i> SRY (sex determining region Y)-box	82%
EU503117.1	<i>Sus scrofa</i> sex determining region Y-box 2 (<i>Sox2</i>) mRNA, complete cds	82%
NG_009080.1	<i>Homo sapiens</i> SRY (sex determining region Y)-box 2 (<i>SoxSox2</i>), RefSeqGene on chromosome 3	81%
XM_002814321.1	PREDICTED: <i>Pongo abelii</i> transcription factor <i>SoxSox-2-like (LOC100460719)</i> , mRNA	82%
NM_001105463.1	<i>Bos taurus</i> SRY (sex determining region Y)-box 2 (<i>SoxSox2</i>), mRNA	81%
XM_001506934.1	PREDICTED: <i>Ornithorhynchus anatinus</i> similar to <i>Sox2</i> transcription factor (<i>LOC100075503</i>), mRNA	81%
NM_001109181.1	<i>Rattus norvegicus</i> SRY (sex determining region Y)-box 2 (<i>Sox2</i>), mRNA	81%
NM_001142940.1	<i>Macaca mulatta</i> SRY (sex determining region Y)-box 2 (<i>SoxSox2</i>), mRNA	81%
XM_516895.2	PREDICTED: <i>Pan troglodytes</i> similar to <i>Sox-2 (LOC460869)</i> , mRNA	83%
NM_011443.3	<i>Mus musculus</i> SRY-box containing gene 2 (<i>Sox2</i>), mRNA	81%
NM_001143799.1	<i>Equus caballus</i> SRY (sex determining region Y)-box 2 (<i>SoxSox2</i>), mRNA	80%
D50603.1	<i>Gallus gallus</i> mRNA for <i>SoxSox-2</i> , complete cds	79%
XM_001368783.1	PREDICTED: <i>Monodelphis domestica</i> similar to <i>Sox2</i> transcription factor (<i>LOC100014523</i>), mRNA	85%
BC159121.1	<i>Xenopus tropicalis</i> SRY (sex determining region Y)-box 1, mRNA (cDNA clone MGC:172383 IMAGE:7663004), complete cds	86%

Electronic Supplementary Material S2: Similarity of *Sox14* gene between *O. niloticus* and several vertebrates based on Blast analysis at NCBI.

Accession number	Nucleotide sequence information	Match score
EF431925.1	<i>Oreochromis aureus Sox14 (Sox14) gene, complete cds</i>	99%
EF431927.1	<i>Oreochromis mossambicus Sox14 (Sox14) gene, partial CDs</i>	99%
EF431926.1	<i>Oreochromis karongae Sox14 (Sox14) gene, partial cds</i>	99%
AY277955.1	<i>Takifugu rubripes</i> transcription factor <i>Sox14a (Sox14a) gene, complete cds</i>	90%
NM_001164872.1	<i>Oryzias latipes</i> SRY-box containing gene 14 (<i>Sox14</i>)	91%
NM_001194657.1	<i>Macaca mulatta</i> SRY (sex determining region Y)-box 14 (<i>SoxSox14</i>), mRNA	80%
AB063098.1	<i>Macaca fascicularis</i> brain cDNA clone:QtrA-14282, similar to human SRY-box 14 (<i>SoxSox14</i>), mRNA	80%
XM_526317.2	PREDICTED: Pan troglodytes SRY-box 14 (<i>SoxSox14</i>), mRNA	79%
NM_004189.2	<i>Homo sapiens</i> SRY (sex determining region Y)-box 14 (<i>SoxSox14</i>)	79%
NM_011440.1	<i>Mus musculus</i> SRY-box containing gene 14 (<i>Sox14</i>), mRNA	78%
NM_001163781.1	<i>Bos taurus</i> SRY (sex determining region Y)-box 14 (<i>SoxSox14</i>), mRNA	78%
XM_002742479.1	PREDICTED: <i>Callithrix jacchus</i> transcription factor <i>SoxSox-21</i> -like, transcript variant 1 (LOC100391037), mRNA	83%
XM_002814084.1	PREDICTED: <i>Pongo abelii</i> transcription factor <i>SoxSox-14</i> -like (LOC100460230), mRNA	79%
XM_001916428.1	PREDICTED: <i>Equus caballus</i> similar to rCG37201 (LOC100147540), mRNA	80%
AY612092.1	<i>Tetraodon nigroviridis</i> transcription factor <i>Sox 8 (SoxSox8) mRNA, complete cds</i>	100%

Electronic Supplementary Material S3: Abbreviations and names of genes located in the genomic blocks harbouring *Sox2* and *Sox14* among vertebrates.

ACTL6A - actin-like 6A
MRPL47 - mitochondrial ribosomal protein L47
NDUFB5 - NADH dehydrogenase (ubiquinone) 1 beta subcomplex
USP13 - ubiquitin specific peptidase 13 (isopeptidase T-3)
PEX5L - peroxisomal biogenesis factor 5-like
TTC14 - tetratricopeptide repeat domain 14
CCDC39 - coiled-coil domain containing 39
FXR1 - fragile X mental retardation, autosomal homolog 1
DNAJC19 - DnaJ (Hsp40) homolog, subfamily C, member 19
ATP11B - ATPase, class VI, type 11B
DCUN1 - DCN1, defective in cullin neddylation 1, domain containing 1
MCCC1 - methylcrotonoyl-CoA carboxylase 1 (alpha)
LAMP3 - lysosomal-associated membrane protein 3
B3GNT5 - UDP-GlcNAc:betaGal beta-1,3-N acetylglucosaminyltransferase 5
MCF2L2 - MCF.2 cell line derived transforming sequence-like 2
KLHL6 - kelch-like 6; KLHL24 - kelch-like 24
YEATS2 - YEATS domain containing 2
PPP2R3A - protein phosphatase 2, regulatory subunit B', alpha
M5L2 - male-specific lethal 2 homolog
PCCB - propionyl CoA carboxylase, beta polypeptide
STAG1 - stromal antigen 1
TMEM22 - transmembrane protein 22
NCK1 - NCK adaptor protein 1
IL20RB - interleukin 20 receptor beta
CLDN18 - claudin 18
DZIP1L - DAZ interacting protein 1-like
A4GNT - alpha-1,4-N-acetylglucosaminyltransferase
DBR1 - debranching enzyme homolog 1
ARMC8 - armadillo repeat containing 8
TXNDC6 - thioredoxin domain containing 6
MRAS - muscle RAS oncogene homolog.

Electronic Supplementary Material S4: Abbreviations and names of genes located in the genomic block harbouring *Sox2* among fish species.

ATP11B - AT11B - Probable phospholipid-transporting ATPase

CLD22 - Claudin-22

TIM14 - Mitochondrial import inner membrane translocase subunit TIM14

FXR1 - Fragile X mental retardation syndrome-related protein 1

TTC14 - Tetratricopeptide repeat domain 14

B3GA1 - Galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase 1

HPS3 - Hermansky-Pudlak syndrome 3 protein

BCAR3 - Breast cancer anti-estrogen resistance protein 3

FBP1L - Formin-binding protein 1-like

MCCB - Probable methylcrotonoyl-CoA carboxylase beta chain; mitochondrial

RPE65 - Retinoid isomerohydrolase

PAR1 - Proteinase-activated receptor - Abnormal spindle-like microcephaly-associated protein

FA55C - Protein FAM55C.