

A

1 GCA GAA ACG AGT AAC TGA TCC GAC AGT TTC CGC AAC TCG GAA TAG AAG CGA CTA
57 TTT GAC TGC GCG TGA GAA AAG AGC AGA ACA GTT TGT CTC GAT CGC TTC AGC GGA
111 ATA GTA ATG ATG GTG TTT CAC GGC AGC GAG CTA AAC TCC CCA AGT CCG CCA CCT

M M V F H G S E L N S P S P P P 16

165 CAA GTC CCA TCT GTG GAT GAG GCG GAA GAG GAA CCT AAG AAG TCA ATG CAA CAC

Q V P S V D E A E E E P K K S M Q H 34

219 GTT AAA CGC CCA ATG AAT GCC TTT ATG GTC TGG AGT CAG ATA GAG CGG AGA AAA

V K R P M N A F M V W S Q I E R R K 52

273 ATG GCC GAA GAG CAT CCA GAT ATG CAC AAC GCG GAG ATC AGC AAG CGA CTT GGA

M A E E H P D M H N A E I S K R L G 70

327 AAG CGC TGG AAG CTG TTA TCA GAG AGC GAG AAA AGA CCG TTC GTC GAG GAA TCC

K R W K L L S E S E K R P F V E E S 88

381 GAG CGT CTT CGT ATT AGA CAC ATG CAA GCG TAC CCT GAT TAC AAA TAC AGA CCT

E R L R I R H M Q A Y P D Y K Y R P 106

405 CGC AAA AAG AAA CAG CCT GCT AAG CAG AAG AAC GCG GGG GCG CAA GAT TCG AAA

R K K K Q P A K Q K N G G A Q D S K L 24

489 ACG TCT TCG TCA AAT TCT GAC GAG CAT TCG CAC TCA CGA AAA CAC ATT GCC GAG

T S S S N S D G H S H S R K H I A E 142

543 ACA GTA GGA ACG AAG CGA GAA GCT CTG CCC GGC GCG CAG GGA TAT CCT GGA TTG

T V G T K R E A L P G A H G Y P G L 160

597 GGC TCT TCC AAG AAG CAA GCT ACT TCG CAT GGT GAA GCG CCG TTC AAG AAA CAG

G S S K K Q A T S H G E A P F K K Q 178

651 CGC CGT GAC TTC GCC TCA CCG ATT ACA CCA CCG CCG AAC GTT CCA GAC GCG ATT

R R D F A S P I T P P P N V P D A I 196

705 GGC GTA ACA CCG GAT GAT ACG ATA GAT CAG TTA TCA CTG TAT GAG CAG TTT GAG

G V T P D D T I D Q L S L Y E D F E 214

759 CAC GCT TTT AAG CAA GAT CAA CAA GTC CCC TCC ATT AAC CAC ACT CAA CAG ATG

H A F K Q D Q Q V P S I N H T Q Q M 232

813 CCA AAC TCG TGG CCG AAT CTC GAA TTC GGC AAC ATG AAC TTC AAC GGA TTG CCC

P N S W P N L E F G N M N F N G L P 250

867 GTG TCC CCT TTA CTT GAT CTT CAC GGA ACA AGT GCT GGC CAA TTT GAT TTT CCA

V S P L L D L H G T S A G Q F D F P 268

921 GAT CTC TAC ACT CCG CCC GAG GTA TCC GAG TTG ATT CAA GGC CAA TGG CTG GAG

D L Y T P P E V S E L I Q G Q W L E 286

975 AAC TCT TTA GGT CAT TTG TAG CTA GCT GAT TTT CAA ATT TCG TTA CAA CCG ATG

N S L G H L * 292

1029 GAA ACA GAC TTG TTT AAC ATG TGT GGA ATT TTC GCT TCA AAT GAT TTT TAT AAA
1083 TTT TGA GAG ATT CGC GGA GTG AAC TTC TCG CAT TTT ACT AAC AAG CAG AGA TAC
1137 GTT TTT CAG TGT ATA TTT TTT TTT TGG CCG ATC GAG TTT ATT TTA TTT TGT ATC
1181 GAT TGA TAT GTA TTT CTT GTC TTA GAG AGC CTT TAC TAC GAC AAT TGT ACA AAA
1245 AGA AGA TTC TGA ATA ATT TAA TAG TGT TGT ATT TTT ATG AAT TTT TTT TCG TTG
1299 TGC CCT ACT TGG GAA AAG TGT TTT AAA GTG AAT CCG AAT TGA AGA GAA GAA AAT
1353 TAC ATA CCA TTA TGA AAA AAA AAA AAA 3'

B

AmSoxC 1 -----
NvSoxC 1 -----
DmSoxC 1 MIAKPNQATTEPPLSLRPGTVPTVATTARPATITIQRRHPAPKADSTPHLTPFFSPSPSPASSPSPAPAQTPGAQKTSQAAIHPAA
CiSoxC 1 -----MASTRLLTKISARLSLHAGQLNDGNPSPESGLGSDWDFDSDEIVQSS
DrSox4 1 -----
GgSox11 1 -----
FrSox12 1 -----
HsSox12 1 -----

AmSoxC 1 -----MMVFGHSELNSPSPPOVES-----VDEAEEEPK
NvSoxC 1 -----MMVLQNSEFRASPSHSLGGMDQVD-----ATQATHESKRKS
DmSoxC 91 VASFSAPVAAASTRDPDPGPKSTHSHSHSHSPFPFRESEMDGERSHSHGHMFLSMGDSSLVFSGARVPVNSSTYSDATRTLH
CiSoxC 50 SVATFRSRIIDAIHASSPFLICLDLDPVDSQMDVDPFKKEADYKPKRQSPKMKREMLNLKLVLPATIKLPPPSRSQKLLA
DrSox4 1 -----MDDLMDASPTGGSNSAG-----DKMDIAWCQT
GgSox11 1 -----MVQQAESAESENLPREAMDTREGEFMACSPV-----ALDESDPDWCQT
FrSox12 1 -----MVQKTSHTESTAEALSPFAVDSSSDSGTCMDLPAASLPGSGSTASSTAG-----DKLAEDPWCQT
HsSox12 1 -----MVQQRGARAKRDGGPPLPFGFAEAG-----ARFPWCQT

AmSoxC 31 SMQHVKRPMNAPMVWSQIERRKMAEHPDMHNAISKRLGKRWKLRSSEKRPFVSEERLRIHQAYPDYKYRPRKKKQPAKQKNGGA
NvSoxC 42 DMQHVKRPMNAPMVWSQIERRKMAEHPDMHNAISKRLGKRWKLRSSEKRPFVSEERLRIHQAYPDYKYRPRKKKQPAKQKNGGA
DmSoxC 181 SPGHVKRPMNAPMVWSQIERRKMAEHPDMHNAISKRLGKRWKLRSSEKRPFVSEERLRIHQAYPDYKYRPRKKKQPAKQKNGGA
CiSoxC 140 RPYGHVKRPMNAPMVWSQIERRKMAEHPDMHNAISKRLGKRWKLRSSEKRPFVSEERLRIHQAYPDYKYRPRKKKQPAKQKNGGA
DrSox4 29 PEGHVKRPMNAPMVWSQIERRKMAEHPDMHNAISKRLGKRWKLRSSEKRPFVSEERLRIHQAYPDYKYRPRKKKQPAKQKNGGA
GgSox11 45 ASGHVKRPMNAPMVWSQIERRKMAEHPDMHNAISKRLGKRWKLRSSEKRPFVSEERLRIHQAYPDYKYRPRKKKQPAKQKNGGA
FrSox12 63 PEGHVKRPMNAPMVWSQIERRKMAEHPDMHNAISKRLGKRWKLRSSEKRPFVSEERLRIHQAYPDYKYRPRKKKQPAKQKNGGA
HsSox12 36 PEGHVKRPMNAPMVWSQIERRKMAEHPDMHNAISKRLGKRWKLRSSEKRPFVSEERLRIHQAYPDYKYRPRKKKQPAKQKNGGA

HMG domain

AmSoxC 121 QGSKVSSNSDGD-----HSHSRKHIAETVGTREALPFGAHGYPLGSSKQKQATSHGEAIFKKQQRDFASP
NvSoxC 132 KFAASQSPF-----KMLTTLAATGSAALPFAQMGSYGSSAKRTFNSMSEPTTKKQRDLGP
DmSoxC 271 QADAGCEARNDDTTN-----NNSLTLIANGTTTAGKSSSTFCQSSGSAKRLRNDSDGTSSEKYEVRKMSAEQ
CiSoxC 230 STFSYLHMDPDQSMIEITN-----TEALKAERVVSSKRLPKRRLQKQKSPFEEKPVTVAVINFPQNNSGMSQAEKQVATTNN
DrSox4 119 AERVVSASFGAKS-----ASKSKTLRTRHRKSTTLBLTSHSFDHHLVKRSLVVS-AAQKIPK-KPAKGVHYGG
GgSox11 135 GQSPKKNAPGGGSKAKSSGKCKSLKAAAAPPKPKAKAAPHGDAQDGYVFGAKVSKYSAKCVFVDEEEDDEDEDELQLRKQAD
FrSox12 153 GEGKEXLHSSINTSTKTTSSSRKNGKPSSSKPKSLFGSSSTKASPFASBHQSEHNSLYSKSGVSSAAKQIPDGKPKRMYVYGS
HsSox12 126 PFGGSGGSGSS-----PGRSCLAAAGAEQREGIWAAGRPRFTTKMTTTRSCWKCWSRPRGSGGSGWRDGLPD

AmSoxC 186 ITPPPNVDA-----IGVPPDTHDOLSLYEDFEHAFKQOQVPSIN-----
NvSoxC 191 ITPPPNVDS-----IGVPEEPDOLSLYEDFPFTHQNAQGINS-----
DmSoxC 343 LNSADILPFS-----ADNLISYQSSEYLPPLSLNSADCDLHLSLSSGGLPE-----
CiSoxC 312 LLTAVKVEFR-----VPLRITPATIVTAVPSKGRFVHTPSPCTHQKIGLVHG-----
DrSox4 189 CSTSSSP-----SVAVPASPTLSSASDQDFKFEGLASGSGVAFAPAR-----FKYTRASPSTPSASH-----
GgSox11 225 DEEDDEEGQQLRRYVAVKVPASPTLSSASSTEGASLYEVRVAGGGRGLYSFKNIKTQGGPPPPQAGLSPASSRSISTSSAGSEE
FrSox12 243 SAALVQVFA-----SIVVVPASPTLSSASDQDFKFEGLASGSGVAFAPAR-----
HsSox12 199 KRSAQVSG-----EGAAAAAASPPSEDEPEEEBAAAAEREGREETVAS-----GGSSVRQGGHTYSRRASSTPSGSHSAS

AmSoxC 228 -----HTQQMNSWPNLEFGNINFGNVPVSLPLDLHLFQAGQDFPFLYFPVEVSELIQQWLENSLGH-----
NvSoxC 234 -----MHQQAANNHQIISGWNLELSLNLNSVPSPLFLGLDANGQDFPFLYFPVEVSELIQQWLENSLGH-----
DmSoxC 390 -----SRENLEVVNRFLFLFLGGNEDSGLGVSSFCQSHNQDPLAGLMDNLSLSPFNDRRSEITVEVMVVALPGGSEIERRAHP-----
CiSoxC 363 TVFKVTEGNRQFIIVSGANDVIVSDVTNSSLVSTORTSTPCQKQIINNFVPIITNDGQGRKYNVVRGCIQVVASQEPSTNSTQKRVPEP
DrSox4 252 -YSSQSSSDEFEDELADSPGDFSFSTFTGAPLDRDLRFNFSQSG-SHEPFPDYCFE-EVSEMTSODLLESTINVTFTY-----
GgSox11 315 ADDLLFDLSLNFSGHAAELGAGAAAGHSLGLYLDKDLDFSESLGSHPEPFPDYCFE-EVSEMTSODLLESTINVTFTY-----
FrSox12 327 SHSSSSSDEFEDELINSSSDMSLQSGSGLVDRDLDFNFSQSGSHPEPFPDYCFE-EVSEMTSODLLESTINVTFTY-----
HsSox12 248 -----GSESLFLSRLPPQAGIDCSALDRDPLQPPSGTSHPEPFPDYCFE-EVSEMTSODLLESTINVTFTY-----

AmSoxC -----
NvSoxC -----
DmSoxC -----
CiSoxC -----
DrSox4 -----
GgSox11 -----
FrSox12 -----
HsSox12 -----

453 VPTTFANMGDDFRNEYLK

