

Biglycan-Like Extracellular Matrix Genes of Agnathans and Teleosts

Seikou Shintani,* Akie Sato, Satoru Toyosawa,† Colm O'hUigin, Jan Klein

Max-Planck-Institut für Biologie, Abteilung Immungenetik, Corrensstrasse 42, D-72076, Tübingen, Germany

Received: 13 April 2000 / Accepted: 4 July 2000

Abstract. Biglycan and decorin are two members of a family of small extracellular matrix proteoglycans characterized by the presence of 10 leucine-rich repeats and one or two attachment sites for glucosaminoglycans. Both have thus far been described only from tetrapod species, mainly mammals. Because the extracellular matrix has played an important part in the evolution of Metazoa, the phylogeny of its components is of considerable interest. In this study, biglycan-like (BGL) cDNA sequences have been obtained from two teleost (Oreochromis cichlid and zebrafish) and two lamprey species. The analysis of the sequences suggests that, like tetrapods, the lampreys possess two types of proteoglycans, both of which are biglycan-like; decorin-like proteoglycans could not be identified in these species. The genes specifying these two types apparently arose by duplication in the lamprey lineage after its divergence from gnathostomes. The two teleost species possess a BGL proteoglycan and a bona fide decorin. The BGL proteoglycan is highly divergent from the tetrapod biglycan and related to the BGL proteoglycans of the lamprey. Hence, although the duplication generating the ancestors of biglycan and decorin genes occurred after the divergence of agnathans but before the emergence of teleosts, only decorin acquired its characteristic properties in the bony fishes. The BGL gene presumably turned into a typical biglycan only in the tetrapod lineages. The presumed

* Present address: Department of Pedodontics, Osaka University, Faculty of Dentistry, 1-8, Yamadaoka, Suita, Osaka 585-0871, Japan † Present address: Department of Oral Pathology, Osaka University, Faculty of Dentistry, 1-8, Yamadaoka, Suita, Osaka 585-0871, Japan Correspondence to: J. Klein acquisitions of new functions appear to have been accompanied by changes in the evolutionary rate.

Key words: *Petromyzon* — *Oreochromis* — Zebrafish — Biglycan — Decorin — Proteoglycans

Introduction

An essential innovation in the evolution of metazoan multicellularity was the appearance of the extracellular matrix, an intricate network of macromolecules deposited between cells, especially well developed in connective tissues (Gerhart and Kirschner 1997). The extracellular matrix isolates cells from one another, creates a medium through which cells can move and be guided to specific locations, participates in signal transduction, provides conditions for cell differentiation, and fulfills a host of other functions in the development and maintenance of the body (Alberts et al. 1989). Among the main macromolecular components of the noncollagenous extracellular matrix are proteoglycans composed of a core protein and a varying number of covalently linked glycosaminoglycan (GAG) side chains (Kresse et al. 1994; Gehron Robey 1996; Iozzo and Murdoch 1996). Proteoglycans are the main constituents of cartilage, in which they contain only a few such side chains. The two main species in this latter class of proteoglycans are biglycan (also called PG I; Fisher et al. 1983, 1989; Neame et al. 1989; Wegrowski et al. 1995) and decorin (PG II; Krusuis and Ruoslahti 1986; Day et al. 1986, 1987; Li et al. 1992; Scholzen et al. 1994). Biglycan has two and decorin a single potential GAG attachment site containing the Ser-Gly sequence preceded by one or more acidic amino acid residues (Bourdon et al. 1987). In some of the biglycan molecules, however, one of the two sites may be unoccupied.

Biglycan and decorin are members of the small leucine-rich proteoglycan (SLRP) family characterized by the leucine-rich repeat (LRR), usually present in a tandem array (Iozzo and Murdoch 1996). The LRR domain has also been found in other proteins that together with SLRPs form the LRR superfamily. The LRR domain of the SLRP family is, however, distinguished from that of the other members of the superfamily in that it is flanked by cysteine clusters and in that each LRR consists of a β -strand LxxLxLxxNyL running parallel to an α -helix xaxxyayyyyayyxyy (where x is any amino acid residue; L leucine or isoleucine; N asparagine, which can, however, be replaced by cysteine or threonine; a an aliphatic residue; and y indicates a loose amino acid requirement; Kobe and Deisenhofer 1993). The cysteine clusters at the N- and C-termini of the LRR domain are composed of four and two similarly spaced cysteine residues, respectively, the former in a stretch of 20 amino acid residues. The cysteines are believed to be involved in the formation of intrachain disulfide bonds. Based on the organization of the cysteine clusters, the SLRP family can be divided into three subfamilies, one of which is comprised of biglycan and decorin. The N-terminal cysteine cluster of these two proteoglycans has the structure Cys-X₃-Cys-X-Cys-X₆-Cys, where X is any amino acid residue (Hocking et al. 1998).

The human and mouse biglycan and decorin core proteins each consist of 10 LRRs, they share 55-57% amino acid identity (Fisher et al. 1989; Scholzen et al. 1994; Wegrowski et al. 1995), and their encoding genes have nearly identical exon-intron organizations (Fisher et al. 1991; Vetter et al. 1993; Scholzen et al. 1994). Despite these similarities, however, the two proteoglycans appear to have different functions, as suggested by their distinct patterns of temporal and spatial expression (Hocking et al. 1998). The human biglycan and decorin genes have been mapped to the Xq27-q28 (Traupe et al. 1992) and 12q21-q22 (Pulkkinen et al. 1992) chromosomal regions, respectively. The two genes (cDNAs) have been cloned from human (Fisher et al. 1989; Krusius and Ruoslahti 1986), mouse (Wegrowski et al. 1995; Scholzen et al. 1994), rat (Dreher et al. 1990; Kokenyesi and Woessner 1989), rabbit (Boykiw et al. 1998; accession code U03394), cattle (Marcum et al. 1993; Day et al. 1987), sheep (accession codes AF034842 and AF125041), and dog (accession codes U83140 and U83141); the biglycan gene (cDNA) has also been isolated from the African clawed toad (Xenopus laevis; accession code AB037269), and the decorin sequence is also available from the pig (accession code AF125537) and the domestic fowl (Li et al. 1992).

The emergence of extracellular matrix-rich cartilage, bone, dentin, and other forms of connective tissue represents a major evolutionary innovation associated with the appearance and rise of vertebrates. It is therefore of considerable interest to characterize the molecular basis of this innovation. With this goal in mind, we initiated a search for lower vertebrate homologs of genes that are known to play an important part in the manufacture of the extracellular matrix in tetrapods (Toyosawa et al. 1998, 1999a, 1999b, 2000). The aim of the present study was to trace the evolutionary origin of the biglycan and decorin genes in jawless vertebrates and bony fish.

Materials and Methods

Animals. Adults of the cichlid fish Oreochromis niloticus were obtained from a dealer (Limnotherm, Bergheim-Niederaussen, Germany). Stocks of O. niloticus were maintained at the Max-Planck-Institut für Biologie, Tübingen. Ammocoetes larvae and adults of the sea lamprey Petromyzon marinus were from Lakes Huron and Champlain, USA.

Isolation of DNA (RNA) and cDNA Library Construction. Total RNA was isolated from the gut of P. marinus larvae from Lake Huron, and a cDNA library was constructed as described in an earlier publication (Mayer and Tichy 1995). The library had a postamplification titer of 1.0×10^{10} plaque-forming units per ml (pfu/ml). Jaws from adult O. niloticus were frozen in liquid nitrogen immediately after excision and kept at -70°C until their use. Total RNA was extracted from the powder of the frozen tissue samples. Poly (A⁺) RNA was isolated, and cDNA was synthesized with the help of the mRNA purification kit (Amersham Pharmacia Biotech, Freiburg, Germany) and the TimeSaver cDNA synthesis kit (Amersham Pharmacia Biotech), respectively. The cDNA was inserted into the Eco RI-digested Agt10 vector (Stratagene, Heidelberg, Germany), in vitro-packaged with the help of the Gigapack® cloning kit (Stratagene), and used to transform competent E. coli MN514 bacteria. The initial titer of the library was 3.2×10^5 pfu/ml; the library was amplified once to a titer of 2.6×10^{11} pfu/ml.

Polymerase Chain Reaction (PCR) Amplification. The cDNA lysates were PCR-amplified in the PTC-100 Thermal Cycler (Biozym, Oldendorf, Germany). One microliter of a solution containing 1.0×10^7 or 2.6×10^8 pfu of the cDNA library was added to a reaction mixture consisting of the PCR buffer (50 mM KCl, 1.5 mM MgCl₂, 10 mM Tris buffer, pH 8.5, 0.001% gelatin), 0.2 mM of each of the four deoxynucleoside triphosphates (Amersham Pharmacia Biotech), 1 μ M of each of the sense and antisense primers, and 2.5 units of *Taq* polymerase (Amersham Pharmacia Biotech). The initial DNA denaturation for 3 min at 95°C was followed by 35 cycles, each consisting of 1 min denaturation at 95°C, 1 min annealing at the appropriate annealing temperature, and 2 min extension at 72°C; the final extension was for 10 min at 72°C.

Isolation of cDNA 5'- and 3'-Ends by Rapid Amplification of cDNA Ends (RACE). Reverse transcriptase (RT)-PCR was carried out using 5'- and 3'-RACE kits (Gibco BRL, Eggenstein, Germany). One microgram of total RNA prepared from *O. niloticus* jaws or *P. marinus* bodies was used for the first-strand cDNA synthesis in 5' RACE with *O. niloticus* decorin-specific and *P. marinus* biglycan- and decorinspecific primers; oligo (dT) primers with adapters were used in the 3' RACE. An intermediate terminal deoxynucleotidyl transferase tailing reaction was carried out prior to amplification in 5' RACE. The firststrand cDNA products were then amplified by PCR with other inner primers. For zebrafish, a cDNA preparation was made from the whole body of a single fish using the SMART cDNA kit (Clonetech, Palo Alto, CA). Short fragments of biglycan-like and decorin genes were obtained using the primers shown in Table 1.

Table 1. Oligonucleotide primers used in this study

Designation	Sequence	Orientation	Location
Lamprey BGL1			
G2	5'-GAT CTT GTT GTT RTG AAG RTA GAC-3'	А	E8
Dec-L7	5'-ATC GTT GCC ATT GAG GAT GAA GA3'	S	E6
Dec-L13	5'-AGG CTG CAC TGG CAA CCG AAG-3'	А	E2
Dec-L14	5'-AAG ATC GGA GCA CTG CAC CAC A-3'	А	E2
Dec-L16	5'-ATG GCA CAG ACT CGG GGT ATT T-3'	S	E2
Lamprey BGL2			
G2	5'-GAT CTT GTT GTT RTG AAG RTA GAC-3'	А	E8
Big-L1	5'-TCT CCG TGA TCT TGT TGT TCT GC-3'	А	E3
Big-L6	5'-CAG GTA CAG CGC ATA TAG TTT ATT-3'	А	E3←E4
Big-L7	5'-GC AAG AGG ATC TGA TCC GAT A-3'	S	E6
Big-L14	5'-TCG TGC AGT GCT CAG ACC TG-3'	S	E2
Oreochromis BGL3			
AM10	5'-CCT GGG TAT GTG AAC TTC AGT TAT GA-3'	А	E8
Big-F1	5'-TCT CAT ACG GTT GTT GTC CAG GT-3'	А	E7
Big-F2	5'-AGT GTG GGA ATC AAC GAC TTC TG-3'	S	E8
Zebrafish BGL3			
Big-F3	5'-AGT GCT CTG ACC AGG GTC TGA-3'	S	E2→E3
Big-F4	5'-GTC GAC AGG CCA TTA AAA GCT C-3'	А	E5
Zebrafish decorin			
BDG1	3'-CAG TGC TCT GAY CTV GGT CTG-3'	S	E2→E3
G2	5'-GAT CTT GTT GTT RTG AAG RTA GAC-3'	А	E8
Oreochromis decorin			
G2	5'-GAT CTT GTT GTT RTG AAG RTA GAC-3'	А	E8
Dec-F2	5'-GAT CTT GTT GTT RTG AAG RTA GAC-3'	А	E6
Dec-F4	5'-TTG ACA ACA ATG CTC TGA CCA G-3'	S	E8
Dec-F5	5'-GGC ATT TCC TTT AGC AAG TTC TT-3'	А	E4
Dec-F6	5'-CCT TTT TAA TCT TGG TGA TCT CAT T-3'	А	E4

A, antisense; S, sense; E, exon.

Cloning and Sequencing. Twenty microliters of the PCR amplification product were purified by 1.5% low-melting-point agarose (Gibco BRL) gel electrophoresis, and the bands were identified with ethidium bromide staining, excised, and isolated using the QIAEX II gel extraction kit (QIAGEN GmbH, Hilden, Germany). The isolated DNA was blunt-ended, phosphorylated, and ligated to SmaI-digested pUC18 plasmid vector with the SureClone ligation kit (Amersham Pharmacia Biotech). The reaction products were transformed into competent E. coli XL-1 blue bacteria by standard methods and plated on LB agar containing ampicillin (50 µl/ml) and X-GAL. Transformants were grown overnight in LB-ampicillin broth, and minipreps were prepared according to the standard protocol. Two to five micrograms of DNA were used in the dideoxy sequencing reactions with the AutoRead Sequencing kit (Amersham Pharmacia Biotech). The reactions were processed by the Automated Laser Fluorescent (ALF) sequencer (Amersham Pharmacia Biotech).

Southern Blot Hybridization. Seven micrograms of lamprey (*P. marinus*) genomic DNA were digested with 100 units of the restriction enzymes *Eco*RI, *Hind*III, *Bam*HI, and *Msp*I (Boehringer Mannheim) overnight. The recovered DNA was loaded onto 0.8% agarose gel and run overnight. Alkali blots were prepared using the HybondTM N⁺ nylon membrane (Amersham Pharmacia Biotech). Prehybridization, hybridization, and probe labeling were carried out using the AlkPhos DIRECT kit (Amersham Pharmacia Biotech). One hundred nanograms of DNA were used for the labeling of the probe. After the overnight hybridizations, the DNA was washed according to the AlkPhos DIRECT protocol. Following the application of the chemiluminescent detection reagent CDP-Star (from the kit), HyperfilmTM ECL (Amersham Pharmacia Biotech) was exposed to the blot for 2 h and developed.

Data Analysis. The nucleotide sequences were aligned with the aid of the SeqPup computer program (Gilbert 1995a). Sequence similarities were evaluated with the aid of the DottyPlot computer program (Gilbert 1995b). Substitution rates were estimated by the method of Li et al. (1985) from sequences aligned using the ClustalX program (Thompson et al. 1997). Phylogenetic reconstructions were made using the alignment of the protein segment downstream of the first conserved cysteine cluster. The neighbor joining algorithm of the MEGA program (Kumar et al. 1993) was used to make phylograms, which were bootstrapped to estimate reliability. Distances were estimated from the proportion of amino acid identity following the removal of gapped sites. Parsimony trees were made based on amino acid sequences using the PAUP program (Swofford 1998). A heuristic search method was used with a starting tree obtained by stepwise additions of sequences with one tree held per addition. Ten replications of the addition procedure were made. Optimization was by branch-swapping using tree bisection and reconnection.

Results and Discussion

Characteristics of Biglycan-Like cDNA Sequences

Two *Oreochromis*, two zebrafish, and four lamprey sequences homologous to tetrapod biglycan and decorin cDNA sequences were obtained by screening the corresponding cDNA libraries or cDNA preparations with degenerate primers based on the conserved regions of the 366

CONCENCIE	
Lamprey BGL1-C	
Lamprey BGL1-H	
Lamprey BGL2-C	
Lamprey BGL2-H	
Oreochromis BGL3	
Zebrafish BGL3	
Cattle BG	************WPP-AAIOAT-DBAFTT
Sheep BG	*********-CPAAL-O
Horse BG	*******MP
Dog BG	
Mouse BG	**********CP
Rat BG	*********RPTL-QT-DT-MSDT*************************
Human BG	**********WPR-VSL-QRT-DPF-MAD*L-PV**************************
Toad BG	*********V-LCSCIL-IHRSMDMA-MKV-IPTESIP-*VGLP***********************************
Cattle DC	**********TIIFVAQGQLF*******EEHF-EVPEIE-***********************************
Sheep DC	**********TIIFF-VAQGQLF*******EERFHEVPE-E-*********************************
Pig DC	**********TIVFAQGQLF*******EDRF-EVPE-E-********************************
Dog DC	**********TIIFAQG-Q-R-LF*******EDRAMPD-EL************************************
Human DC	**************************************
Mouse DC	**********TLIFFAQGR-LFI-YDPD*****N-******************************
Rat DC	*********TLV-FAQGR-LFI-YDPD*****N-******************************
Fowl DC	*********-RLVL-FVP-CL-TR-HLFI*******-G-ADMAPTDD***PVISG************************************
Oreochromis DC	**********-RSAC-SVTACR-SLM*******-GD*-VTES-LPPVIGG**********************************
Zebrafish DC	********SAC-SVS-CR-S-MVM******-PDGPEL-TTRKPHV************************************
Human keratocan	**************************************
Fowl keratan	**************************************
Human lumican	**************************************
Fowl lumican	**************************************
Rat fibromodulin	**************************************
Fowl fibromodulin	**************************************
1	11
CONSENSUS ==>	MCPFGCQCHLRFPTVVQCSDLGLKSVPKDIPPDTTLLDLQNNKITEIKEDDFKNLKHLYALILVNNKISNSKIHPKAFSPLVKLEKLYLSKNHIKEIPENLPKSLQELRI
Lamprey BGL1-C	GS***AMAGKNARMVSQG-AQ-HFL-A**AM-S-DI-H-R-T-V-TGPIV
Lamprey BGL1-H	GS***ASK-ARMVSQG-AQ-HFL-A**AM-S-DI-H-R-T-V-TGI-PIV
Lamprey BGL2-C	SSA-***VQAK-ARQQQQL-**-VASS-DI-H-Q-T-V-GSM-SV
Lamprey BGL2-H	SSA-***VQAK-ARQQQYYASS-DI-H-Q-T-V-GSM-SV
Oreochromis BGL3	DAH-H-SP-***Q-IDK-EVMIDCQQG-NKG-F-I**RRNMDN-RLY-M-NAPNVIF
Zebrafish BGL3	DER-SKK***-LQTKAN-LDG-DNF-LQ**RNMDKI-HY-L-TQMI-V-SL
Cattle BG	X***
Sheep BG	R-QIR-QIR-QIR-QIR-QIR-QIR-QIR-QIR-QIR-Q
Horse BG	**E
Dog BG	RQIRQIRQIRQIRQIRQIRQI
Mouse BG	R-Q1R-Q1R-Q1
RAL BG	RRQ1RQ1RQ1
Human BG	RRRQ1RQRQ1RQ1
Toad BG	L
Cattle DC	
nie po	
Pig DC	
Dog DC	
Mourae DC	
Pat DC	
Oreochromie DC	V
Zebrafish DC	
Human keratocan	
Fowl keratan	E-OF-F-PPS-NALV-DNKFI-A*-ARTWV-YI-FT-S-KD-V-ATPMIN-NKT-NG-FSGVI-K-KP-IV-F-FD-FE-V-ADVG-FSG-
Human lumican	N-APE-N-PESY-SAMY-DE-XM*VGIKY-Y-RO-DH-D-KA-K-VTD-OMDH-VI.EKGRVK-KO-KHINH-N-T-SVGPED-OI.
Fowl lumican	V-APE-N-P-SYAMY-DN-XTI-T*V-SGIKY-Y-RM-EA-E-NT-D-VTD-(WDH-HLEKGRVK-KN-KHINY-N-T-AVGPT-DD-OL
Rat fibromodulin	D-OE-D-PPNAMY-DNRNYL-F*V-SRMXYVYFO-AA-O-GV-D-ATG-LWIA-HG-O-TSDGR-VK-RHRDH-N-TRM-GPRRHL
Fowl fibromodulin	QOE-D-PPN-SSAMY-DTRN-RYL-F*V-TRMKYVYFO-A-O-GA-D-ATE-EW-A-HOSE-MGKRV-AK-KNRMNN-N-TKM-SPRHL
2:	21 231 241 251 261 271 281 291 301 311 32 <mark>1</mark>
CONSENSUS ==>	HENEITKVPKGVFSGLNNMIVIELGGNPIKNSGIENGAFQGLKKLSYLRISDTKLTGIPKGLPESLTELHLDCNKTAIEAEDLKGYPNLARLGLSHNQISMVENPGSLA
Lamprey BGL1-C	L-KRDT-INNGRLHKPSVND-T-IYSQLENLEE-VDFY-FY-K-TE-Q-*
Lamprey BGL1-H	L-KRDT-INNGQLHKPSVND T-I YSQLENLEE-VDFY-FY-K-TE-O-*
Lamprey BGL2-C	N-K-IDAMKRLHAL-MQSTVEER-V-V-VSARDN-IQEH-QL-QIRLIHYKVIQ-+E
Lamprey BGL2-H	N-K-IDAMKRLHAL-MQSTVEER-V-V-VSARDN-IQEH-QIC-QIRLIHYKVIQ-*E
Oreochromis BGL3	Q-DRIQ-DA-KRKLH-LATINS*TL-IG-AEASDF-S-IS-YSKV-IFIR-KQAFKY*
Zebrafish BGL3	-D-K-SRLA-K-MHDLN-LSAIADV
Cattle BG	-D-R-RR-NCMEF-PD*-NEADT-NHQLVR-SK-Y+GR-I*S
Sheep BG	-D-R-RR-NCMEF-PD*-NEADT-NHQLR-SK-YGR-I*S
Horse BG	-D-R-RR-NCMEFQPD*-NEADT-NHQLR-SK-YGR-I*\$
Dog BG	-D-R-RRNCMEF-PD*-NEADT-NHQLLR-SK-YGR-I*S
Mouse BG	-D-R-RR-NCMEF-PD*-NEADT-NHQLLR-SK-YGR-II*S
Rat BG	-D-R-RR-NCMEF-PD*-NEADT-NHQLLR-SK-YGR-I*S
Human BG	-D-R-R
Toad BG	KNCNCM
Cattle DC	R-SNQVTSMIANI-TQP

Fig. 1. Legend appears on p. 367.

	331	341	351	361	371	¥ 381	391	401	411	★421	431
CONSENSUS ==>	NVPHLI	RELHLDNNKL	TRVPPGLADLI	YLQVVYLHSN	NITKVGVNDI	FCPVGFGVKR	AYGLFGYGPYY	SGISLFSNP	VPYWEVQPAI	FRCVTDRSAI	QFGSNYKKK
Lamprey BGL1-C	VSGN-	L-	VSSK-H	S-N	K-KE-KPT	TV-SP	-*******Q-	AYD	-KP-SV	HNHN	HR-*
Lamprey BGL1-H	VSGN	L-	VSSK-H	S-N	K-KE-KPT	TV-SP	-********Q-	-AYD	-KP-SV	HNHN	HR-*
Lamprey BGL2-C	TC	s-v-	-QF	H	K-AA-KSD	SK-ASP	V*******L-	D	-ND-P-SA	NSV	'SQ-FR-*
Lamprey BGL2-H	TC	s-v-	-QF	H	K-AA-KSD	SK-ASP	V******L-	D	-ND-P-SA	A\$V	'SQ-FR-*
Oreochromis BGL3	-T-KI	IYRM	RKSS-1	RIFG-	K-SSI	IRADS-K	N*******P-	TA	-KAI	G-RGV	'-L-*-FR-*
Zebrafish BGL3											
Cattle BG	FL-T		SAP	LT-			-******	NN		L	**
Sheep BG	FL-T-		SAP	LT-			-******	NN		L	*
Horse BG	FL-T		SAP	LT-			-*******	NN		L	*
Dog BG	FL-T-		SSP	LT-			_*******	NN		L	*
Mouse BG	FL-T		SAP	L	I	M	_*******	NN		L	*
Rat BG	FL-T		SAP	.L	I	M	-******	NN		L	*
Human BG	FL-T-		ASP	L		M	-******	NN		L	*
Toad BG	FM-V		SKP-M-	L	Q	I	-*******	NN		L	*R-*
Cattle DC	-T	N	AKG-VH-		SAI-S	P-YNT-K	-*******S-	V	-QIS-	YV-A-V	'-L-***
Sheep DC	-T	N	VKGH-		SAI-S	P-YNT-K	-*******S-	V	-QIS-	YV-A-V	'-L-***
Pig DC	-T	N	NKGEH-	N-	SAS	P-YNT-K	-********S-	v	-QIS-	YV	-L-***
Dog DC	-T		IGEH-	- I N-	SAS	P - YNT - K	-*******S-	v	-QIS-	YV	-L-***
Human DC	-T		GEH-	N-	svss	P-HNT-K	-*******	v	-QI\$-	YV	-L-***
Mouse DC			LQH-	- I N-	SAQ	RA-HPSRK	-*******S-	-AVYG	-RÌF-N-	YV	-L-***
Rat DC			LAQH-	-vN-	SEQH	LPSYQTRK	T*******	TAVY	-RQIH-H-	FGT-	-L-***
Fowl DC		NE-	VSGEH-	-IN-	K-ASI-I	L-YN'T-K	(-********	v	-QISA	IHEV	'-I-***
Oreochromis DC	-A	A-	-SP-H-	IАН	K-AATE	PNT-K	(-*******M-		V-	F	-L-*R
Zebrafish DC		EA-	-AAH-	I							
Human keratocan	D-SSI	D-Q-SH-Q-	-KR***I\$2	HHLH-DH-	K-KS-N-SVI	ISPSMLPA	ERDS-SHI	RYLR-DG-E	IKP*PIPM-L	MT-FRLLQ-V	'II******
Fowl keratan	\$\$II	D-Q-SH-Q-	-KI***IN#	H-EHLH-DH-	R-KS-NGTQ1	ISIA-AE	DY-NI-RI	RYLR-DG-E	IQP*PIPLDI	MI-FQLLQ-V	VI******
Human lumican	SS-1	/D-SY	KNI-T***VNI	EN-ENYEV-	QLE-FDIKS-	KILGPLSY	S********	KHLR-DG-R	ISETSLP-DM	IYE-LRVANEV	TLN*****
Fowl lumican	-ITS-V	/D-SF-Q-	KSI-T***VSH	EN-ENFQV-	K-N-FPLSS-	K-VGPLTY	S*******KI	THLR-DG-N	LTRADLPQEM	IYN-LRVAAD-	SLE*****
Rat fibromodulin	-SSS-I	D-SY-Q-	QKI***VN7	N-ENLQG-	R-NEFSISS-	T-VDVMNF	'S*******KI	QVLR-DG-E	IKRSAMPVDA	PL-LRLA-L-	EI******
Fowl fibromodulin	n -SSSII	D-SY-R-	QKI-R***VS7	N-ENLQG-	Q-NEFSISS-	T-VDVMNY	S*******	QVLR-DG-E	IKRNAMP-DA	PL-LRRATV-	EI******

Fig. 1. Amino acid sequence alignment of SLRP family members. The fish and lamprey sequences are from this study; sources for the additional sequences are referenced in the text. The alignment of the initial 110 amino acids was adjusted to bring potential GAG attachment sites at position 52-53 into register. The conserved Ser-Gly residues are indicated with closed rectangles (\blacksquare). Closed arrows indicate the positions of conserved cysteine residues. The horizontal line with vertical dividers indicates the positions of 10 LRR repeats in the human biglycan (BG) and decorin (DC) sequences, as described by Hocking et al. (1998). A consensus sequence based on simple majority is shown. A dash indicates identity with the consensus, an asterisk (*) the position of an indel, and a dot (.) unavailability of sequence information.

Table 2. Characteristics of cichlid and lamprey proteoglycan cDNA sequences

					P	osition of		
			Length		Translation	Stop		GenBank accession
Species	Proteoglycan	Total	ORF†	3' UTR	start site	codon	PAS	number
P. marinus	BGL1 C	2240 (P)*	388	1073	?	1165	2207	AF247826
	BGL1 H	2535 (P)	347	1491	?	1042	2518	AF247825
	BGL2 C	3114 (F)	410	1791	91	1321	3068	AF247828
	BGL2 H	2558 (P)	309	1625	?	931	2502	AF247827
O. niloticus	BGL3	1634 (F)	370	136	385	1496	1592	AF247821
	Decorin	2072 (F)	359	865	128	1025	2048	AF247822

* F and P indicate full and partial length cDNA sequences, respectively; the length is given in base pairs (bp)

† Open reading frame given in the number of encoded amino acid residues

C and H indicate the source of P. marinus from Lakes Champlain and Huron, respectively; PAS, polyadenylation signals; UTR, untranslated region

tetrapod genes (Table 1). Amino acid sequences translated from nucleotide sequences, which can be accessed in the GenBank database under the numbers AF247821– AF247828, are shown in Fig. 1, together with the known tetrapod sequences. In each case, the initial clone isolated from anchored PCR amplification products provided a partial cDNA sequence. In the case of *Oreochromis* and lampreys, full-length or nearly full-length cDNA sequences were then obtained in three to four steps by anchored PCR using primers based on the partial sequences generated in the preceding step (Table 1). Blast search analyses, sequence comparisons (Fig. 1; Tables 2 and 3), and phylogenetic analyses (Fig. 2) indicate that the eight sequences are very closely related to the biglycan group, which also includes, in addition to biglycan itself, decorin. As discussed in the next section, however, only two of the eight sequences correspond unambiguously to tetrapod decorins and are therefore so designated. The remaining six sequences are designated as biglycan-like (BGL) and distinguished by numbers (see below). Like other members of the biglycan group, the eight sequences have 10 LRRs; potential GAGattachment sites in the N-terminal part of the protein; and two clusters of cysteine residues flanking the LRR region (Fig. 1). The N-terminal cluster consists of four closely spaced cysteine residues; by contrast, the two cysteine residues in the C-terminal region are separated by 33 amino acid residues. The cysteine residues are invariant in all the thus far identified members of the biglycan group.

	Lamprey	Oreochromis	Oreochromis	Tetrapod	Tetrapod	
	BGL2	BGL3	DC	BG	DC	Outgroup
Lamprey BGL1	66.7	55.4	55.4	59.6	54.0	30.3
Lamprey BGL2		55.7	57.3	62.1	54.1	28.3
Oreochromis BGL3			53.6	56.3	53.6	27.7
Oreochromis DC				60.4	70.9	28.3
Tetrapod BG					58.9	29.7
Tetrapod DC						29.2

The percentage identity was measured in the conserved portion of the protein, following the omission of the N-terminal sequence up to the conserved cysteine residues. Tetrapod biglycan (BG) includes human, cattle, sheep, horse, dog, mouse, rat, and toad sequences. Tetrapod decorin (DC) includes human, cattle, sheep, pig, dog, mouse, rat, and fowl sequences. The outgroup consists of human keratocan, human and fowl lumican, as well as rat and fowl fibromodulin sequences



368

Fig. 2. A neighbor-joining tree of SLRP family protein sequences. The protein alignment shown in Fig. 1 was used. The poorly conserved N-terminal part (initial 110 residues of the alignment) as well as sites containing indels were omitted. Distances were estimated from the proportion of amino acid identity between sequences and the dendrogram drawn using the MEGA program. Numbers on the nodes indicate the percentage recovery of that node in 500 bootstrap replications.

Identities of the Biglycan-Like cDNA Sequences

For convenience, we distinguish the biglycan group (which includes biglycans and decorins of tetrapods, and the BGL sequences of agnathans and teleosts described here) from what we refer to as the "outgroup." The latter consists of the closest known relatives of the biglycan group identified by phylogenetic analysis; it includes human keratocan (Tasheva et al. 1999) and lumican (Grover et al. 1995), fowl keratan (Dunlevy et al. 1998) and lumican (Blochberger et al. 1992), as well as rat (accession code X82152) and fowl (Nurminskaya and Birk 1996) fibromodulin. Furthermore, for the same reason we divide the biglycan coding sequence into an Nterminal part extending from the N-terminus to the first cysteine cluster, the LRR region, and a short C-terminal part. The latter two parts are not difficult to align; the alignment of the LRR regions does not require the introduction of indels at all and that of the C-terminal part requires only one to three indels, depending on the sequence (Fig. 1). There is also no difficulty in extending the alignment of these two parts to the outgroup sequences, because at the amino acid level many residues are invariant not only throughout the biglycan group but also in the outgroup.

In contrast to these two parts, the sequences of the N-terminal part are alignable among tetrapods, but virtually unalignable between agnathans/teleosts and tetrapods (Fig. 1). The part is of different lengths in the various sequences so that long indels must be introduced to obtain alignments based on similarity rather than identity of amino acid residues. The variation in length is concentrated in a segment immediately upstream from the cluster of the four conserved cysteine residues. This segment appears to be absent in the tetrapod sequences, but in the agnathan and teleost sequences it is dominated by a particular amino acid residue that varies from sequence to sequence. Thus, in the Oreochromis BGL3, the segment of ~20 residues contains 16 Asp residues, whereas in the lamprey BGL1 and BGL2 sequences, the segment is dominated by proline residues whose number varies among the sequences. Proline-rich segments are also present in some of the outgroup sequences. Outside of this segment, the N-terminal part sequences fall into four classes: the tetrapod biglycans, the tetrapod decorins, the teleost BGL3, and the lamprey BGL1 and BGL2. The outgroup sequences form classes of their own. Within each class, it is possible to align the sequences on the basis of amino acid identity or similarity; alignment between classes is usually equivocal.



Fig. 3. Southern blot hybridization with the probe of zebrafish *BGL1* and *BGL2*. Zebrafish genomic DNA was digested by *Eco*RI (E), *Hin*dIII (H), *Bam*HI (B), and *Msp*I (M) restriction enzymes.

The canonical GAG attachment site consists of a Ser-Gly pair preceded by acidic amino acid residues. The number of Ser-Gly pairs in the N-terminal part ranges from zero to three. Mammalian biglycans each contain two such pairs, but the toad biglycan contains only one pair, as does the lamprey BGL2. The mammalian decorins each contain one Ser-Gly pair, as does the fowl decorin, whereas the Oreochromis decorin contains two pairs. The fowl pair and one of the Oreochromis pairs are not preceded by acidic amino acids, however. The lamprey BGL1 contains three Ser-Gly pairs, of which two are preceded by acidic residues, whereas the Oreochromis BGL3 contains none, as do some of the outgroup sequences (human keratocan, fowl keratan, rat and fowl fibromodulins). Other outgroup sequences (human and fowl lumicans) contain one pair. The position of the potential GAG-attachment sites is conserved in tetrapod biglycans and decorins (with the exception of the fowl decorin), but not in the outgroup sequences. The number and positions of the GAG-attachment sites cannot, therefore, be taken for a diagnostic feature of the different proteoglycan classes; the position shifts among the sequences of both the biglycan group and the outgroup.

Pairwise comparisons with known sequences reveal that all the new sequences resemble the biglycan group more closely than the outgroup sequences (Table 3) in terms of amino acid similarity. This observation, in combination with the results of the phylogenetic analysis described in the next section (see Fig. 2), establish the membership of the agnathan and teleost sequences in the biglycan group. The comparison furthermore identifies one of the two *Oreochromis* sequences as decorin (~71% similarity with tetrapod decorin) rather than biglycan (~60% similarity with tetrapod biglycans). Here again, the assignment is supported by the phylogenetic analysis (Fig. 2). The second *Oreochromis* sequence (as well as

the short zebrafish sequence) has slightly higher sequence similarity with tetrapod biglycans (~56%) than with tetrapod decorins (~54%). On this basis, we have designated it as biglycan-like, BGL3 (but see below). The same is true for the two main types of lamprey sequences (Table 3) and so these, too, are denoted as biglycan-like, specifically BGL1 and BGL2.

Phylogenetic Relationships

In the neighbor-joining tree in Fig. 2, the tetrapod decorins form one clade, which is joined by the *Oreochromis* decorin sequence, and the tetrapod biglycans form a second major clade. The *Oreochromis* BGL3 and the lamprey BGL1 and BGL2 sequences lie outside of these two clades. The *Oreochromis* BGL3 sequence is inexplicably found in a sister-group position to this entire cluster when the tree is rooted by the outgroup sequences. A tree drawn by the maximum parsimony method has a similar topology to the neighbor-joining tree (data not shown). The two issues that arise from this phylogenetic reconstruction are, first, the relationship of the BGL3 sequence of the biglycan group of sequences, and second, the relationship of the lamprey sequences to the biglycan and decorin clades.

There are at least three explanations that may account for the outer position of the BGL3 sequence. The first explanation is that BGL3 is really a biglycan that, however, evolves faster than the tetrapod biglycans and that the accelerated evolutionary rate is responsible for its anomalous position on the tree. Affinity of BGL3 to tetrapod biglycan sequences is suggested not only by its overall closer sequence similarity with biglycans but also by shared conserved sites. In the entire BGL3 sequence there are 34 substitutions and 1 single-codon indel that BGL3 shares with the tetrapod biglycans, compared to 29 substitutions it shares with tetrapod decorins (Fig. 1). The second possibility is that BGL3 is a new member of the biglycan group that is neither biglycan itself nor decorin. If this were the case, the question would then have to be answered regarding the fate of the biglycan gene in teleosts. A search for a more orthodox version of cichilid fish biglycan failed to produce any candidates. When the search was extended to the zebrafish, only an apparent ortholog of the Oreochromis BGL3 sequence was found (Fig. 1). This result would then seem to suggest that either the "true" biglycan has been lost in at least some teleosts or that biglycans and decorins diverged after the divergence of teleosts and tetrapods. The latter possibility is, however, contradicted by the presence of a bona fide decorin in Oreochromis. The third possibility is that BGL3 is encoded in a pseudogene and so evolves anomalously from the presumably functional biglycans. The only observation that might possibly be interpreted as favoring this possibility is the apparent absence of a GAG-attachment site in BGL3, but we note 370

A

| Lamprey BGL2-C | 1
CAGTCI | 1
AGAGA | .1
CTCC

 | GGCG | 21
ЭСАЛ
 | AAGCG | GAGG | 31
9 CGT | гстсс
 | ACT | 11
CCAA | CTCC | 51
:CG 1 | CCAG
 | CACT | 61
C A0 | GCAC | AGAT | 71
г сто | ACGC | 8
ACA | 1
GCCC | CCCA
 | cc | | | | | | | | | | | | | | | | |
 | | | | | |
 | | |
 | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
|--|--|--
--
--
--
--
--
--
--
--
--
--
--|---|---|--
---|---|--|--|---|---
--|---|--|--|--|---|--
--	---	--
--	--	---
---	---	---
---	---	---
--	---	--
--	---	---
--	-------	--
---	--	--
---	--	---
--	---	--
---	--	---
Lamprey BGL2-C	1 ATG G	с тсе

 | TGC | GTG
 | TGC | GCT | TTG | CTC
 | 11
CTC | GTC | TGC | GCC | стс
 | тсс | TCC | ccc | тсс | ATC | 21
TCC | TCC | TCG | GTG
 | GCC | ACG · | 300
 | ACC | тсс | тсс | 31
AAA | CCG | TTC
 | GCC | CAG | CGC
 | CAG | ттс | ттс | ACC | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| PROTEIN | M J
41 | A S | L

 | c | v
 | с | A | L | L
 | L
51 | ۷ | с | A | L
 | s | s | P | s | I | s
61 | s | s | v
 | д | т | A
 | т | s | s | к
71 | P | F
 | A | õ | R
 | ð | F | F | т | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-C
PROTEIN | GAC T | FC ATG | ATG
M

 | GCC
A | GAC
D
 | GAC
D | GCC
A | GTG
V | GAG
E
 | TTC
F | TCG
S | GGA
G | GAC
D | GAC
D
 | TCC
S | GTG
V | GCC
A | CCC
P | AGC
S | GCC
A | ACG
T | CCC
P | AAG
K
 | AAG
K | GTC V | G
G
 | GGG
G | GAC
D | CGC
R | TCC | AAA
K | GCG
A
 | ACC
T | GCC
A | 666
6
 | AAG
K | CAG
Q | CCG
P | GGC
G | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-C
Lamprey BGL2-H | 81
CGC GC | GG GCC | . GCC

 | ACC | сс с
 | CCA | AAG | TCG | стс
 | 91
CCG | cc6 | ссө
 | сст | ccg
 | CCG | ccg | сст | ccc | GAC | 101
GCG | AGC | TGC | CCG
 | TTT
 | GGG | rgc
 | CAG | TGC | AGT | 111
GCG | cgg | GTC
 | GTG | CAG | TGC
 | TCA | GAC | CTG | GGC | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| PROTEIN | R (| 3 A | A

 | т | P
 | P | к | s | L
 | Р
131 | P | P | Ρ | P
 | P | P | P | P | D | р
141 | 5 | с | P
 | F | G | с
 | Q | с | 5 | A
151 | R | v
 | v | Q | с
 | s | D | L | G | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-C
Lamprey BGL2-H
PROTEIN | CTC G | rc TCC
 | GTG

 | CCC

P | CAG
 | GCC | ATC
 | CCC

P | ААА

К
 | GAC | GCC | CGC

R | CTC | СТС

 | GAT | CTG

L | CAG | AAC
N | AAC | AAG | ATC. | ACG
T | GAG

E
 | ATC | AAG
K | 2AG
 | GAC

D | GAC | TTC

F | AAG

K | GGC

G | CTC

 | AAT | AAA

K | CTA
L
 | TAT | GCG

A | CTG

L | TAC | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-C | 161
CTG G | rg aac | AAC

 | CTC | ATC
 | AGC | AAG | GTG | CAC
 | 171
CCC | AAG | GCG | ттс | GCA
 | CCG | стс | TCG | AGC | CTG | 181
GAC | AAG | CTG | TAC
 | ATC | TCG | AC
 | AAC | CAG | сто | 191
ACC | GAG | GTG
 | ccc | GGC | TCC
 | ATG | CCG | тсс | тсс | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-H
PROTEIN | L V | / N | N

 | L | r
 | s | к | v | н
 | P | к | A | F | A
 | P | Ŀ | 5 | s |
L | D | к | L | Ŷ
 | I | 5 | Н
 | N | Q | L | т | Е | v
 | P | G | s
 | M | P | s | s | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-C
Lamrey BGL2-H | 201
CTG GT | rg gag | стс

 | ccc | ATC
 | CAT | GAA | AAC | AAC
 | 211
ATC | AAG | AAG | атс
 | ccc .
 | AAG | GAC | GCG | TTC | TCT | 221
GGC | ATG . | AAG | AGA
 | CTC | CAC | GCG
 | стс | GAG | ATG | 231
GGA | GGG | AAC
 | ccc | CTG | CAG
 | TCG | ACA | GGG | ATC | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| PROTEIN | L 1
241 | ν E | L

 | R | I
 | н | Е | N | N
 | I
251 | ĸ | ĸ | I | P
 | к | D | А | F | s | G
261 | м | к | R
 | L | н | A
 | L | E | м | G
271 | G | N
 | P | L | Q
 | s | т | G | I | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-C
Lamprey BGL2-H
PROTEIN | GAG G | rg ggg

/ g | GCC

 | TTC
F | GAG
E
 | GGA

G | CTG

L | GAG

E | CGC
 | CTC | GTG

V | TAC

Y | GTG

V | CGC

R
 | GTG

V | TCC

S | GAC

D | TCG

S | AAG

K | CTC | GCG

A | CGC

R | ATC
 | CCC

P | ада

К | GAT
D
 | CTG

L | CCC .
P | AAC

N | TCG | ATC
I | CAG
 | GAG

E | CTG

L | CAC

H
 | TTG

L | GAG

E | CAC

H | AAC

N | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-C | 281
CAG A1 | IC ACA | GCT

 | сто | GAG
 | CAA | GAG | GAT | стб
 | 291
ATC | CGA | TAC | CCG | стс.
 | ATT | CAC | CGG | TTG | GGC | 301
TTG | AGC ' | TAT | AAC
 | CAG | ATC 1 | AG
 | GTG . | ATT | CAG | 311
AAC | GGA | -
AGC
 | CTG | GAG | ACC
 | TGC | -
CCA | CAC | стб | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-H
PROTEIN | Q : | ст. | A

 | L | E
 | Q | E | D | L
 | ï | R | Y | P | L
 | I | н | R | L | Ģ | L | s | Y | N
 | Q | I | к
 | v | I | Q | N | G | s
 | L | Е | т
 | с | P | н | L | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-C
Lamprey BGL2-H | 321
AGG G/ | AG CTG | CAC

 | стс | GAC
 | AGC | AAC | GTC | ста
 | 331
ACG | CAG | GTG | ccc | сст
 | GGC | CTG | GCC | ттс | стс | 341
AAG | CAC | CTG | CAG
 | GTG | GTC | TAC
 | CTG | CAC . | AGC | 351
AAC | AAG | ATC
 | GCG | GCC | GTT
 | AAA
 | TCG | GAC | GAC | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| PROTEIN | R 1
361 | C L | н

 | L | D
 | s | N | ۷ | L
 | т
371 | ç | v | 9 | P
 | Ģ | L | А | F | L | к
381 | н | L | Q
 | v | v | Y
 | L | н | s | N
391 | к | I
 | А | A | v
 | к | s | D | D | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-C
Lamprey BGL2-H
PROTEIN | TTC TC | BC AGC | AAG

K

 | GGC
G | GCG
 | AGT
S | CCC
 | ААG

К | CGA

R
 | GTC | CTT

L | TAC
Y | TCG

S | GGC
 | ATC
I | тсс

s | CTG

L | TTC

F | GAC | AAC | ссс
 | GTC .
V | AAC
 | ТАС

Ү | TGG (
 | D
 | GTG

V | CCG | ссс

Р | TCT

S | GCC

A | TTC

F
 | CGC

R | TGC

C | GTG

V
 | GCC

A | AGC

S | CGC

R | TCA

S | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Learning Dot 0 | 401
GCC 61 | |

 | 100 | 03.0
 | | | 000 | | | | | | | | | | | | | | |
 | 411 | | | |
 | | | | | | | | |
 | | |
 | | | | | |
 | | |
 | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-C
Lamprey BGL2-H
PROTEIN |
A \ |
/ Ω |
F

 | s | 2
2
 | N N |
F |
R | K K
 | TAG
 | | | | | | | | | | | | | | | | |
 | | | | | | | | |
 | | |
 | | | | | |
 | | |
 | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-C
PROTEIN
Lamprey BGL2-C
Lamprey BGL2-C | A V
AGGTGO | 7 Q | F
II
GAGG

 | s
GGCGC | 21
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
 | N | F |
R
31
C GGG | K
GGCA
 | TAG | 41
CGGG
A- | GGGG | 5
GG C | 1
AGGG
 | GCTC | 61
A CC | CTGG | GGGC | 71
CCA | TGGC | CCA | 81
GGAA | ACCC
 | 9
AT C | 1
CCAG/ | ACC.
 | 101
A CC | ACAG | GAAP | 111
\ GCA | .GGGT | rccc
 | 121
ACC | CCGA | I:
GCA
 | 31
FCACO | 30000 | A | | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-C
Lamprey BGL2-C
Lamprey BGL2-C
Lamprey BGL2-H
Lamprey BGL2-C
Lamprey BGL2-C | A V
AGGTGC | | F
F
GAGG
51
GAGT

 | GGCGC | 21
22
24 GP
161
 | N
N
N
N
GGGC | F
F | 31
GGG
171 | K
K
GGCP
 | TAG

AGGG

2CTT | 41
CGGG
A-
81
TTCC | GGGG

CCTG
 | 5
GG C
-* -
19
GC A | 1
AGGG

1
CAGC
 | GCTC
 | 61
A CC

201
C CA | CTGG | GGGC

GGGG | 71
CCA
211
CTC | TGGC

TCAC | CCA
-*
2:
GGT | 81
GGAA

21
ACGG | ACCC

AATT
 | 9
AT C

23
TT T
 | 1
CCAG/

1
TTAT: | ACC.
 | 101
A CC

241
T CC
 | ACAG | GAAP
GCGP | 111
GCA
251
4 TTT | GGGI | 7000

 | 121
ACC

261
CCAJ | CCGA(| 1:
GCA :

GAA (
 | 31
FCACO
71
CCAAF | 10000 | Ъ
G
- | | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-H
PROTEIN
Lamprey BGL2-H
Lamprey BGL2-C
Lamprey BGL2-C
Lamprey BGL2-C
Lamprey BGL2-C
Lamprey BGL2-C
Lamprey BGL2-C | A V
AGGTGC
141
CCGCCC
-T
281
CCACGC | 2
2
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3 | F
11
GAGG
51
GAGT
91
CTCG

 | GGCGC
CTCGI | 21
2A GP
161
TC CT
301
AC AT
 | N
AGGGC
GTGA | F
CGGC | 31
31
31
171
AAT
311
311 | K
GGCP
ACCC
 | TAG

AGGG

PCTT
-*
3TGC | 41
CGGG
A-
81
TTCC

21
CCCA | GGGG

CCTG

AGCA | GG C
-* -
19
GC A

33
GC A
 | 1
AGGG
1
CAGC
1
CCAG
 | GCTC

CGAG
 | 61
A CC

201
C CA

341
G GG
 | GCGT | GGGC

GGGG

ATCG | 71
211
CCA
211
CTC
351
TTA | TGGC

TCAC

TCGT
 | CCA (
2:
GGT)
 | 81
GGAA
21
ACGG

61
ACCC | ACCC

AATT

TGCG

 | 9

23
TT T

37
TC C
 | 1
CCAGJ
TTAT
TTAT | ACC.
 | 101
A CC

241
T CC

381
C CA | ACAG

ACCG

GCCG | GAAP
GCGP | 111
GCA
251
TTT
391
TGC | GGGT | 2227

2264¢

 | 121
ACC

261
CCAJ

401
GAT | CCGA | 1:
GCA 5
GAA 0
GAA 0
4:
CAT 0
 | 31
FCACO
71
5CAAF | 100000
10000
10000
10000 | 'G
 | | | | | | | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-H
PROTEIN
Lamprey BGL2-C
Lamprey BGL2-C
Lamprey BGL2-C
Lamprey BGL2-C
Lamprey BGL2-C
Lamprey BGL2-C
Lamprey BGL2-C
Lamprey BGL2-C | A V
1
AGGTGC
 | 10 CHO
5666GT
10 CCCCC
10 CCCCC
10 CCCCC
10 CCCCC
10 CCCCC
10 CCCCC
10 CCCCC
10 CCCCC
10 CCCCC
10 CCCCCC
10 CCCCCCC
10 CCCCCC
10 CCCCCCC
10 CCCCCCCC
10 CCCCCCCC
10 CCCCCCC
10 CCCCCCC
10 CCCCCCCC
10 CCCCCCCC
10 CCCCCCCC
10 CCCCCCCC
10 CCCCCCCCCC
10 CCCCCCCCC
10 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC | F
II
GAGGO
51
GAGTO
91
CTCGO
31
GTGGO

 | GGCGC
CTCGI | 21
2A GP
161
TC CT
301
AC AT
441
AC TC
 | AAT
N
GGGGC
GTGA
ATCA | ACCCA | 31
31
31
171
AAAT
311
AAAT
311
AAAT
451 | K
GGCA
CACCC
 | AGGG
AGGG
CTT
TGC
TGC
AGTG | 41
CGGG
A-
81
TTCC

21
CCCA

61
GAAA | GGGG

AGCA

AGAC | 5
GG C C
19
GC A

33
GC A

47
AA C
 | 1
AGGG
1
CAGC
1
CCAG
1
CTAG | GCTC

GCGAG

CACT | 61
A CC
201
C CP

341
G
GG

481
T CC
 | GCTGG | GGGG
GGGG
ATCG

CCGC | 71
: CCA
211
; CTC
351
; TTA
491
: ACC | TGGC

TCAC

GGGC
 | 2
2
2
3
3
3
2
0
7
3
5
7
5
4 | 81
GGAA
21
ACGG

61
ACCC

01
CTCG | ACCC
AATT
TGCG

 | 9
AT C
23
TT T

37
TC C

51
GG C
 | 1
TTAT
1
CATT
1
CATT | AACC. | 101
A CC

241
T CC

381
C CA

521
G GC

 | ACAG | GAAF
GCGF
AACC | 111
251
391
391
531
531 | | FCCC
3AGC
CGCT
4CGG
 | 121
ACCI
261
CCAJ

401
GATV

541
TGGG | CCGA+ | 1:
GCA 2:
GAA 0
 | 31
FCACC
71
5CAAF
11
5CTTF
51
50
51
 | 10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
1000000 | ХА
 | | | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | |
 | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-L
Lamprey BGL2-C
Lamprey BGL2-C | A V
1
AGGTGC
-T
281
CCACGC
-T
261
CCTGAC | 7 Q
GGGGT
CCCCCC
CTGAC
CTGAC
CTGAC
CCCGA
CCCGA | F TTC F
F
11
GAGG4
51
GAGT4
91
CTCG4

31
GTGG4
71
GAAC1

 | GGCGC
CTCGI
CCCGAGA
CCACGA | 21
22
23
24
24
161
700 CT
301
40 TC
581
44
44
581
44
581
 | AAT N
N
AGGGCC
AGGGCT
AATCA
CACAT | F
F
CCGGC
CCCCA | 31
31
31
31
171
311
311
311
311
451
591
3 GGA | K
SGGCP
RACCC
 | AGGG
AGGG
I
CCTT
STGC
AGTG
CTGT
CTGT | 41
CGGG
A-
81
TTCC

61
GAAA

601
GAAC | GGGG

AGCA

AGAC
 | 5
GG C
-* -
19
GC A
-
-
33
GC A
-
47
AA C
-
-
-
-
-
-
-
-
- | 1
AGGGG
1
CCAGC

1
CTAGG

1
GGGAG
 | GGTG
CGAG
GCGC
CACT
 | 61
A CC

201
C CP

341
G GG

481
T CC

621
T AC
 | GCGT
GCGT
TGAG
GTCC | GGGC
GGGGG
ATCG
CCGC
TACC | 71
211
5 CTC
351
7 TTA
491
8 ACC
631 | 7690

TCAC

GGGC
 | 2:
GGT :
-* | 81
GGAA
21
ACGQ
61
ACCC
01
CTCQ
41
TGGG | ACCC

TGCG
ACAA
* ACC
T
 | 9
AT C
23
TT T
 | 1
CCAG/
TTAT:

1
CCATT:

1
GAGGG
 | AACC.
TTTT
AGT
SAGT
SATC
 | 101
A CC

241
T CC

381
C CA

521
G GC

661
T GA
 | ACAG

GCCG

TATG
 | GAAF | 111
251
391
391
531
671
4 GTC | AGGC
AGGC
AGGC
AGGC
AGGC
AGGC
AGGC
AGGC | CCCC
SAGC
CCCT
ACCC
GCCC
 | 121
ACCI
261
CCAJ

401
GAT

541
TGG
GAC | CCGA | 11
GCA 5
GGA (
GGA (
4.
CAT (
55
SGA (

 | 31
TCACC
71
CCAAF
11
GGTTF
51
GGTAT
51
GGTAT
91
CCTGC | ACCGT | ж
19
19
19
19
19
19
19
19
19
19
19
19
19 | | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-H
PROTEIN
Lamprey BGL2-H
PROTEIN
Lamprey BGL2-H
Lamprey BGL2-H | A V
1
AGGTG
CCGCCC
-T
281
CCACGC
-T
561
CCTGAC

561
CCTGAC
 | 1 Question Carlos Carlo | FIC GAGGA
S1
GAGGA
91
CTCGG

31
GTGGG

11
GCTGG

 | CTCGT
CTCGT
CCGAGA
CGAGA
CGAGA
TTGGP | 21
22
23
24
24
161
75
20
20
20
20
20
20
20
20
20
20
20
20
20
 | AATAA | F
F
CCGGC
ACCCA
ACCCT
TTTG
TTTGA | 31
31
31
31
311
311
311
311
311
591
591
591
30
30
30
30
30
30
30
30
30
30
30
30
30 | AAA
K
SGGCP
CACCC
CCGCA | AGGG
AGGGG
1
CCTT
*
3
STGC
4
AGTG
CTGT
6
CTGT
7
505GG
 | 41
CGGGG
A-
81
TTCC
221
CCCA

61
GAAA

601
GAAC

41
* CAC.
G | GGGG
CCTG
AGCA
AGCA
TGCC | 5
GG C
-* -
19
GC A

33
3
GC A

61
GT G

5
CC A
 | 1
AGGG
1
CAGC

1
CTAG4

1
GGGA4

1
TGTGJ | CGAG | 61
A CC
201
C C
341
G GG

481
T CC

6211
T AC

761
T GG
 | GCGGT
GGCGT
TGAG
GGCCC
 | GGGC | 71
211
5 CTC
351
5 TTA
491
631
5 CCA
631
771
8 AGG | TGGC

TCQT

GGGC

56666
 | 2007 1
31
COT 1
51
COT 2
51
COT 2
COT 2 | 81
GGAA
21
ACGG
61
ACCC
01
CTCG
41
TGGG
81
* AGC
3 | ACCC
AATT
TGCG
ACAA
* ACC
T
AGGT | 9
AT C
23
TT T
 | 1
CCAG/
1
TTATY
1
CATTY
1
CCCCGO
CCCCGO
1
CCCCGO
 | AACC. | 101
A CC

241
T CC

381
C CA

661
T GA

801
G TG4
 | ACAG
ACCG
GCCG
TATG
TCCCG
GGTG | GAAF
GCGF
AACG
STTG
CGGP
AGTT | 111
251
391
531
671
4 GTC

 | .0307
 | CCCC
BAGC
CCCT
ACCG
GCCC | 121
ACCI
CCAJ
GAT-

691
GACI

821
TGAC | CCGA
 | 11
GCA 5
SAA (
CAT (
SAA (
CAT (
SAG (
CCA (
CCA (
SAG | 31
TCACC
71
CCAAF
11
CCAAF
31
31
CCTGG
31
CCTGG
31
CCTGG
31
CCTGG | ACCOT
ACCOT
ATCGC
TAAA
GACC | ж.
19
19
19
19
19
19
19
19
19
19
19
19
19 | | | | | | | | | | | |
 | | | | | | | | | | |
 | | | | | | | | | | | | | | | |
 | | | | | | | | | | | | |
 | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | |
| Lamprey BGL2-H
PROTEIN
Lamprey BGL2-H
PROTEIN
Lamprey BGL2-H
Lamprey BGL2 | A V
1
AGGTGG
AGGTGG
 | 2
2
2
2
2
2
2
2
2
2
2
2
2
2 | FIC
F
F
II
GAGG
91
CTCG
91
CTCG
31
GTGGC
71
II
GAACC
51
GTGAC

 | GGCGC
CTCGI
CTCGI
CGAGA
TTGGA
ATCAC | Q
21
CA GP
161
CC CT
CC CT
301
AC AT
581
CC TC
581
CC TA
581
CC TA
581
CCC TA
581
CC C | NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA
NGTGA |
F
F
CCGGCCAA
ACCCT
TTTTG
TTTTGAA | R
31
31
31
311
311
311
311
311 | K
GGGQP
RACCC
CCGCA
AACCC
CCGCA
GATG | AGGG
1
COTT
3
STGC
4
4
AGGG
5
5
5
5
5
5
5
5
5
5
5
5
5 | 41
CGGG
 | GGGGG

AGCA

TGCC

AGAT

TGGG
 | 5
GG C
-* -
19
GC A

33
GC A

47
AA C

61
GT
G

 | 1
AGGGG
CAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGCO
1
CCAGO | GGCTG
CGAG
GCGC
CACT

CCCCG

AGGG
 | 61
A CC

201
C CP

341
G GG

481
T CC

621
T AC

761
T AC

761
T GG

761
T GG
 | GCGT
GCGT
GGTTA
TGAG
GGGAG
GGGAG
 | GGGC
GGGGG
ATCG
CCGC
TACC
AGGA
GCCG | 71
211
5 CTC
211
5 TTA
491
2491
2 CA
631
2 CA
631
2 CA
911
3 TGT | TGGC
TCAC
TCGT
GGGC
GGGC
TGAG
TGAG | 2007 2
30
50 5
50 5
50 5
50 5
50 5
50 5
50 5
5
 | 81
GGAA

21
ACCGG

61
ACCCC
01
CTCG

81
* AGCG

81
* AGCG

21
81
* AGGG
 | ACCC

TGCG

ACCA
T
AGGT
 | 9
AT C

23
TT T

37
TC C

51
GG C

65
TA C

65
TA C

93
CT G

93
CT G

 | 1
CCAGJ
TTAT
1
CATTI
1
CATTI
1
CATTI
1
CATTI
1
CATTI
1
CATATI
1
CATATI | AACC. | 101
A CC
241
T CC
381
C CA
521
C CA
521
C CA
521
T GA
601
T GA
501
T GA
501
T GA
50
50
50
50
50
50
50
50
50
50
50
50
50 | ACAG

GCCG
TATG
TCCCC
GGTG | GAAF
GCGA
AACC
STTC
CGGA
AGTT | 111
A GCP
251
A TTT
391
531
531
531
671
C GTT
951
C GTA
 | GGGGT
GCCC
GCCC
GCCC
GCCC
GCCC
GCCC
GCC | ACGG | 121
ACCO
CCAA

401
GATU

541
TGGG
GACU

961
GCAU
 | CCGA
ACCG
STCT
AAAAA
GGTCC
CAAGT
 | 111
GCA :
2
GAA (
3
GAA (
4)
GAT (
5
5
SAG (
5
SAG (
5
S
SAG (
5
S
SAG (
5
S
SAG (
5
S
SAG (
S
SAG (
5
S
SAG (
5
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | 31
CCACC
71
71
71
71
71
71
71
71
71
71 | ACCGT
ACCGT
ATCGC
TTAAA
SGACC
JATGG | 2A
16
17
17
17
17
17
17
17
17
17
17
17
17
17 | | | | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | | |
 | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | |
| Lamprey BGL2-H
PROTEIN
Lamprey BGL2-T
Lamprey BGL2-T
Lamprey BGL2-T
Lamprey BGL2-T
Lamprey BGL2-C
Lamprey BGL2-C | A 1 A AGGTGG 141 CCGGCC cccaccc 281 CCCCCC 421 TCCCCAC 701 CCACGCC CCAGGTC 841 TOTOFGT 701 CCAGGTC 981 CTGTAN | A GGGGT | FIL F 11 GAGGG 51 GAGGT 91 CTCGGC 31 GTGGG 71 GAACT 51 GTGGG 51 GTGGA 991 GCTTT

 | GGCGC
CTCGI
CTCGI
CCAGP
TTGGP
ATCAC | Q
21
2A GP
161
CC CT
301
AC AT
4411
AC CC
581
AC CC
721
3G TA
581
BA CC
721
13G TA
440
100
CC TC

 | N
AGGGGC
VATCA
CATCA
CATCA
CATCA
CATCA
CATCA
CATCA
CATCA
CATCA
CATCA
CATCA
CATCA | F
CCGGCCA
ACCCCA
ACCCCA
TTTTG
TTTTG
TTTTGA
TTTGA | R
31
31
31
31
31
31
31
31
31
31 |
K
SGGCP
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGC | TAG
AGGG
CTTT
* | 41
CGGGG
A-
81
TTCC

61
GAAA

81
GTTA
81
GTTA

81
GTTA

921
CCAGT | GGGGG

AGCA
AGAC
TGCC
AGAT
TGCC

TGGG
 | 5
GG C
19
GC A
 | 1
AGGGG

1
CCAGGC

1
GGGGA

1
TGTGGGA

1
CTCCC
*
*
 | GCTC

GCGAG

CCCCG

AGGG

TGAG

TGAC | 61
A CC
201
C CP
341
G GG
481
T CC
621
T CC
761
T GG
901
T GT
761
T GG
901
T GT
104
C TG
 | GGGAG
GGAG
GGAG
GGAG
CONTA | GGGCC
GGGGG
ATCGG
CCGC
TACC
AGGA
GCCGG
GCCG | 71
211
CGPC
CTC
CTC
351
TTA
491
CCA
631
CCA
631
CCA
911
TGT
1051
AGC | TGGC
TCAC
TCGT
TCGT
TGAG
TGAG | CCA
2
GGT ;
 | 81
21
ACGG

61
ACCCC

01
CTCG

81
ACGC

81
ACCC

01
CTCG

01
CTCG

01
CTCG

01
CTCG

01
CTCG

01
CTCG

01
CTCG

01
CTCG

01
CTCG

01
CTCG

01
CTCG

01
CTCG

01
CTCG

01
CTCG

 | ACCC

TGCG

ACAA

ACGT

ACCT

ACCT | 9
AT C
TT T
37
TC C
51
C
GG C
51
C
GG C
51
C
C
GG C
51
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | 1
CCAG
1
TTAT
1
GATT
1
GATT
1
GAGGC

1
CCCCG

1
CACAT
A

7
A
A
7
A
A | AACC.
 | 101
A CC

241
T CC

381
C CA

661
T GA

801
G TG
G TG

941
T GC

108
A CC | ACAG
ACCG
TATG
TATG
GGTG
ATGCC
1
TGTA | GAAF
GCGF
AACC
STTC
CGGA
AGTT | 111
A GCP
251
391
5 TGC
531
5 CAG
671
4 GTC
951
7 GTA
951
7 GTA
10911
4 TTC | 100000

 | CGCT
ACGG
GGGC
ACCTT | 121
ACCI
CCAJ
GAT
GAT
541
TGG
GAC
541
TGG
GAC
961
GAC
101
TGA | CCGA
ACCGI
GTCTI
AAAAA
GGTCC
CAAGT
TGCAC | 11
GGA (
GGA |
31
FCACC
71
CCAAF
31
GGTAT
31
GGTAT
31
CCTGG
31
CCTGG
31
CCTGG
31
CCTGG
31
CCTGG
31
CCTGG
31
CCTGG
31
CCTGG
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAAF
31
CCCAF | ACCOT
ACCOT
ATCGC
TRAAA
SGACC
SGACC | 2A
16
17
16
17
17
17
17
17
17
17
17
17
17
17
17
17 | | | |
 | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | |
 | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-H
PROTEIN
Lamprey GGL2-C
Lamprey GGL2-C
Lampr | A 1
A 50
1
AGGTGK
 | A CARACT SCOTT | F 110
GAGGG
51
GAGTG
91
CTCGG
71
GAACT
71
GAACT
51
GCTGJ
991
GCTGT
991
GCTGT
991
S1
GTGAC
71
31
AACAA

 | GGCGC
GGCGC
CCCGAGP
CCACGP
TTGGP
ATCAC
TCCAG
GACTC | Q
21
2A GP
161
CC AT
301
AC AT
301
AC AT
581
AC CC
721
721
721
721
721
721
721
721
721
721
 | NATCA
NOTGA
NATCA
SGGCT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT
SACAT | F
CCGGCCA
JACCTA
TTTGGGC
TTTGGGG
TTTGGGG
ACAT | R
31
31
31
31
31
31
31
31
31
31 | |
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AGGG
AG
A | 41
CGGG
A-
81
TTCC
21
CCCA

61
GAAC

81
GAAC

81
CAGT.

81
CAGT.

161
AGTA-
 | GGGG
CCTG
AGCA
AGAC
TGCC
AGAT
CCTA
CCTA | 5
GG C A
19
GC A
33
GC A
47
6
GT -
10
GT G
-
10
CT G
-
11
10
-
-
-
-
-
-
-
-
-
-
-
-
- | 1
AGGG
AGGC
1
CTAGG
1
CTAGG
1
TGTGG
1
CTCC
3
1
TACG
7
7
1
AGGC
C | GGTC | 61
A CC
201
G GG
G GG
G GG
G GG
G GG
G GG
G GG
G |
GGTTA
GGTTA
TGAG
GGTCC
GGAG

I
ATGC
1
GGCAC | GGGC
GGGGG
ATCG
CCGC
TACC
AGGA
GCCG
AAAC
AAAC | 71
211
5 GTC
351
5 TTA
491
2 ACC
631
5 CCA
631
5 CCA
911
7 TGT
911
1051
1051
1191
1051 | TGGC

TCGT

GGGC

TGAC

TGAC

TGTA:
 | 2007 2
2007 2
3007 2
507 2 | 81
GGAA

61
ACCC

61
CTCG

81
AGCC

81
AGCC

81

81

81

81

21

01

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

21

201

201

201

201

201

201

201
 | ACCC
AATT
TGCG
ACAA
* ACCT
 | 9
AT C 23
TT T
 | 1
CCAGG
TTAT:
1
TTAT:
1
GATT:
1
GAGGC

1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
CGTC:
1
C
C
C
C
C
C
C
C
C
C
C
C
C | AACC. | 101
A CC

241
T GC

381
G GC

801
G GG

801
T GC

108
A TC

122
c GC | ACAG
ACCG
GCCG
TATG
TCCCC
GGTG
ATGCI
1
TGTA
1
CGTTT | GAAF
GCGF
AACC
CCGA
AGTI
AGTI
IGCA | 111
111
111
251
391
373
531
531
531
671
4 GTC
671
811
7 GTC
951
1091
4 TTC

1231
2 CCC
 | | CGCT
ACGG
GGCC
ACGG
AACGG
AACGG
CGCC
CATA | 121
ACC
CCAI
GCAI
GAT
541
GAC
541
GAC
541
GAC
541
GAC
541
GAC
100
100
100
100
100
100
100
100
100
10 | CCGA
ACCG
STCT
AAAA
GGTCC
CAAGT
CAAGT
TGCAG
 | 11
GCA 2
GCA 2
GAA 0
4
4
GAT 0
5
5
GAG 0
6
GCCA 0
6
GCCA 0
6
GCCA 0
7
GCCA 0
7
GCCA 1
7
CCCA 1
12
CCCA 1
CCCA 1
CCCC | 31
TCACC
71
11
53
53
53
53
53
53
53
53
53
53 | CCCCCC
ACCGT
TAAA
TCCCC
ATCGC
ATCGC
ATCGC
ATCGC
ATCGC
ACCCC
TCCCC | 2A
16
16
17
17
17
17
17
17
17
17
17
17
17
17
17 | | | | | | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | | |
 | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | |
| Lamprey BGL2-C
Lamprey BGL2-C | A \\ 1
AGGTGG
-T | 2
3
3
3
3
3
3
3
3
3
3
3
3
3 | F F F GAGG4 S1 GAGG4 S1 GAGG4 S1 GTCG4 S1 GTGG4 S1 GTGG4 S1 GTGG4 S1 GTGG4 S1 GTGG4 S1 GTG73 S1 GTG73 S1 GTG73 S1 GTG73 S1 GTG73 S1 GTG74 S1 GTG73 S1 GTG74 S1 GTG74 S1 GTG74 S1 GTG74 S1 GTG74 S1 GTG74 GTG74 GTG74 GTG74 GTG74 <tr tr=""> <tr< td=""><td>GGCGCGC
CCACGP
CCACGP
CCACGP
TTGGP
ATCAC
TCCAG
GACTG
GGCGA</td><td>Q
Q
Q
21
CA GF
161
CC CT
301
CC CT
301
CC CT
581
CC CT
581
CC CT
581
CC CT
721
S6 TA
CC
721
100
CC CT
100
CC CC CT
100
CC CT
100
CC CC CC CC
100
CC CC CC CC
CC CC CC
CC CC CC C</td><td>National Second
Second</td><td>F
CCGGCC
F
CCGGCC
ACCCT
TTTG
TTTGA
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGGG
CCGGCC
TTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TCGGGC
CCG
CCG
CCG
CCG
CCG
CC</td><td>R
311
311
311
311
311
311
311
31</td><td>K
SGGCZP
SGTCG
SGTCG
SGTCG
SGTCG
SGTTA
SGCCG
C</td><td>TAG
AGGG II
SCTT
*
JTGC
TTGT
55565

85675

1
S5675

1
S5675

1
S5675

1
S5675

1
S5675

1
S5775
</td><td>41
CGGG
A-
81
CCCA
61
GAAT

81
61
GAAT

81
91
41
* CACC
G
81
GTTA

91
161
AGTA

301
GCCA

</td><td>GGGGG
CCTG
AGCA
AGAC
TGCC
AGAT
TGCC
CCTA
GAAT</td><td>5
GG C</td><td>1
AGGG
GAGGC

1
CCTAGG

1
1
CTCCC

1
1
CTCCC

3
1
TACGC

1
1
AGGCC

1
1
CTACG

1
CTACG
</td><td>GGTC</td><td>61
A CC
201
C CP
341
T CC
481
T AC
621
T AC
621
T AC
761
T GT
901
T GT
104
C TG
112
C AC
C AC</td><td>GGTTGAG
GGTCC
GGTCC
GGAG
GGAG
J
GGAG
J
GGAC
J
J
GCAC
J
J
TGGAG</td><td>GGGGC
GGGGG
ATCG
CCGC
TACC
AGGA
GCCG
GCCG</td><td>71 211 211 771 351 777 491 7771 491 7771 493 7771 1051 7771 1051 1191 1331 776 776 777</td><td>TGGC</td><td>2007 2
2007 2
2007 2
30
2007 2
10
2007 2
10
10
10
10
10
10
10
10
10
10</td><td>81
GGAA
21
ACGG
61
ACCC
01
CTCC
81
ACCC
21
81
CTCC
81
ACCC
01
CTCC
81
ACCC
01
CTCC
81
CTCC
81
ACCC
01
CTCC
81
ACCC
91
CTCC
91
ACCC
91
CTCC
91
ACCC
91
ACCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCCC
91
ACCCCC
91
ACCCCC
91
ACCCCC
91
ACCCCCC
91
ACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC</td><td>ACCC
AATT
TGCG
ACAA
ACCA
ACCT
ACCT
ACCT</td><td>9
AT C 23
TT T</td><td>1
CCAGU
1
TTAT

1
GATT
1
GAGGC

1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACC
1
CACC
1
CACC
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT</td><td>AACC.</td><td>101
A CC
241
T C
381
C CA

521
G CC

521
G CC

801
G CG

801
G CG

108
A TC

122
C CC

136.5
S CT
</td><td>ACAG</td><td>GAAF
GCGA
STTC
CGGA
AGGTI
GCA
STCT
STCT</td><td>111
A GCA
251
A TTT
391
391
391
391
391
391
391
671
671
671
671
671
7
GTC
671
1091
1091
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095</td><td>43360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
443600
443600
443600
443600
443600
443600
4436000
4436000
443600</td><td>CCCC
CCCT
CCCCT
CCCCT
CCCCT
CCCTC
CCCCCC</td><td>121
ACCI
CCAJ
GCAT
GAT
GAT
GAT
GAC
GAT
GAC
GAT
GAT
GAT
GAT
GAT
GAT
TAC
GAT</td><td>CCGA
ACCGI
STCT-
CAAG
CAAG
CAAG
STTTJ</td><td>11
GCA (
GCA (
GCA (
GCA (
GCA (
GCCA
(
GCCA</td><td>31
FCACC
71
CCANF
11
36TTF
31
36TTF
31
36TTF
31
36TTCAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
3</td><td>ACCOT
ATCOC
TTAAA
SGACC
SGACC
COCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
SGACCA
CCCA
TTCCA
CCCA
SGACCA
CCCA
TTCCA
CCCA
TTCCA
SGACCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
SGACCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
CCCA
STCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
C</td><td></td><td></td></tr<></tr> <tr><td>Lamprey BGL2-H
LAMPrey BGL2-H
PROTEIN
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPR</td><td>A \\
1
A \\
1
A \\
1
A \\
1
A \\
1
A \\
1
A \\
1
A
\\
1
4
2
3
1
CCAGGTC

2
3
4
2
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
1
2
2
1
2
2
1
CCAGGTC

1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1

</td><td>SGGGT
SGGGT
SGGGT
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGA
SGGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA</td><td>F F F S1 GAGGU S1 GCTGGU S1 GCTGU S1 GAGUU S2 S1 GCTU S2 GAGUU GCTU GCTU GAUU GCTU GCTU GCTU GAUU</td><td>GGCGCC
CTCOT
CCACGZ
CCACGZ
TTTGGZ
ATCAC
TCCAG
GACTC
GACTC
GACTC</td><td>Q
Q
2
2
301
161
CC CT
301
161
CC CT
301
161
CC CT
441
40
CC AT
441
16C AT
581
100
581
100
56 AT
114
T AG
T AG
T AG
T AG
T AG
T AG
T AG
T
AG</td><td>N
N
GGGGC
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C</td><td>F
CCGGCC
F
CCGGCCA
ACCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA</td><td>R
31
31
7
311
7
300
311
7
300
311
7
300
301
7
311
7
300
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
30
7
30</td><td>K
SGGCP
SCCCCC
CCGCA
SCCCCCCCCCCCCCCCCCCCCCCCCC</td><td>TAG
AGGG
CTT
CTT
CTT
CTT
CTT
CTT
CTT
CTT
CTT</td><td>41
CGGG

21
CCCA
61
GAAC

81
GAAC

81
GAAC

81
GAAC

81
GAAC

81
GAAC

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

</td><td>GGGGG
CCTG
AGCA
AGCA
TGCC
TGCC
AGAT
CCTA
GAAT</td><td>5
GG C
- 19
GC A
- 33
GC A
- 47
AA C

GT G

- 61
GT G

</td><td>1
AGGG
1
CTAGG

1
CTAGG

1
TGTGJ
1
CTCCC

3
1
TGTGJ
3
1
CTCCC

1
AGGCC

1
AGGCC

1
CTAGG

1
CTAGG

1
CTAGG

1
CTAGG

1
CTAGG

1
CTAGG

1
CTAGG
</td><td>GGCTC

CGAG

CCCCG

AGGGG

TGAG

TGAG

CGAC

CGAC
</td><td>61
A CC
2011
G GP

491
T C

621
T AC

761
T GC

104
C T GC

132
C AC

146
C AC

132
C AC</td><td>GCTGG
GCGT
TGAG
GGAG
GGAG
GGAG
1
TGGAG
1
TGGGG
1
TGGGG
1
CAC*</td><td>GGGCC
GGGCG
ATCG
CCGC
TACC
AGGA
GCCG
GCCG</td><td>71 201 351 777 491 2491 2491 2491 2491 2491 2491 2491</td><td>TGGC
TCGT
TCGT
TGGG
TGGG
TGGG
TGGG
GGTG
CGCA</td><td>CCA (
</td><td>81
GGAA
GGAA
21
ACGG

61
ACGG

61
ACGG

81
AGGG

81

81

81

81

04
CTGG

81

04
CTGG

81

201
CTGG

81

201
CTGG

81

201
CTGG

81

201
CTGG

201
CTGG

</td><td>ACCC
</td><td>9
AT C
3
TT C
3
GG C
3
G
3
GG C
3
GG C
3
G
C
C
3
GG C
3
GG
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C</td><td>1
CCAG
1
TTAT:
1
CCCGG
CATT
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
C</td><td>AACC.</td><td>101
A CC
241
T CC
381
C CA
521
G CC
521
G CC
52
G CC
52
C
52
C
52
C
52
C
52
C
52
C
52
C
5</td><td>ACAG
GCCG
TATG
GCCG
TATG
GGTG
GGTG
ATGCI
1
TGTAT
1
CTTT
1
CACAG</td><td>GAAF
GCGA
AACC
STTC
CGGA
ATGT
IGCA
STCT
CCCG</td><td>111
A GCA
251
391
370C
531
531
531
531
531
531
531
531</td><td>19397</td><td>CCCC
CCCC
CCCCC
CCCCCCCCCCCCCCCCCCCCCC</td><td>121
ACC
CCA
401
GAT
541
541
GAT
681
GAC
681
GAC
681
GAC
100
1101
TAC
1241
GAT
1241
GAT
1241
GAT
1241
GAT
1241
GAT</td><td>CCGA
Accor
STCT
STCT
SGGTCC
SAAG
STCT
SAAG
STTT
ACCTC</td><td>11
GGA (
GGA (
GAT (</td><td>31
FCACC
71
CCAAF
31
30
31
30
31
30
31
30
31
30
31
30
31
30
31
30
31
30
31
30
30
30
30
30
30
30
30
30
30</td><td>ACCOT
ACCOT
ACCOT
ATCOC
TTAAA
SGACC
SGACC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC</td><td>A 19 19 19 19 19 19 19 19 19 19 19 19 19</td><td></td></tr> <tr><td>Lamprey BGL2-H
LAMPrey BGL2-H
PROTEIN
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPR</td><td>A N A N 1 AGGTGG 141 CCGCCG 1281 CCACGC 2421 TCCCACG 701 CAGGTG 701 CAGGTG 701 CAGGTG 701 CAGGTG 701 CAGGTG 701 CAGGTG 1121 GTATAC 1261 GGTACT 1261 CATACF 1401 CATACF 1541 CCTCCTAC</td><td>SGGGT
SGGGT
SGGGA
SGGA
SGGA
SGGA
SGGA
SG</td><td>11 F 11 GAGGA 91 CTCGG 31 GTCGGC 31 GTCGGC 93 11 GTCGCC 93 13 AACAC 271 131 AACAC 411 S51 CACAG</td><td>GGCCGC
CCCAGA
CCCCAGA
CCCCAGA
CCCCAGA
CCCCAGA
CCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCCAGA
CCCCCCAGA
CCCCCCAGA
CCCCCCAGA
CCCCCCAGA
CCCCCCAGA
CCCCCCCAGA
CCCCCCAGA
CCCCCCCAGA
CCCCCCCAGA
CCCCCCCAGA
CCCCCCCC</td><td>Q
Q
211
161
C C T
301
A GP
441
A C T
581
C
C</td><td>AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG</td><td>ITCGGC
CCGCCA
ACCCT
ACCCT
TTTGA
TTTGA
ACCT
TTTGA
ACCT
TTGGG
ACCT
ACCT
TGGGG
ACCT
TGGGA
ACCT
TGGGA
CCCCA</td><td>R
31
31
311
7 AGA
591
591
591
591
591
591
591
591</td><td>K
GGGCP
CACCCC
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA</td><td>TAG
AGGG
CTT
STGC
TTT
STGC
TTT
STGC
TTT
STGC
TTT
STGC
TTT
STGC
TTT
STGC
STGC</td><td>41
CGGGG
A-
81
TTCCC
CCCA

61
GATCC
41

81
GATCA

81
CAGTA

301
GCGGG

301
CCCA

301
CCCA

</td><td>GGGGG
GGGGG
AGCACA
AGACA
TGGC
TGGG
AGAT
CCTA
GAAT.
CACACA</td><td>5
GG C - 19
GG A - 33 3
GG A - 47
GG GT G - 61
GG GT G - 75
GG A - 61
GG C A - 75
GG A - 10
GT G - 10
GT G - 11
TG A A A - 14
CC - 15
TG T - 13
AA A - 14
TG C - 15
TG T -</td><td>1 AGGGG CTAGG CTAGG CTAGG TGTGGGGAG 1 GGGGAG 1 TGTGGG 1 TGTGG 1 TGTGG 1 TGTGG 1 TAGGC 1 1 AGGCC 1 AGGCC 1 AGGCC 1 AGGC 1 AGG</td><td>GGCTCC
GCGAG
GCGCC
CACT
CCCCG
CCCCG
CCCCG
CCCCG
CCGAC
CGAC</td><td>61
A CC
2011
G GC CP

341
G GC

621
T CC

621
T CC

761
T GC

104
C TG

104
C TG

104
C TG

106
C A

106
C A
</td><td>GGGGG
GGGGG
GGGGG
GGGGG
GGGGG
GGGGG
CGGGG
CGGGG
CGGC
CGGC
CAC*
</td><td>GGGC
GGGG
ATCG
CCGC
TACC
AGGA
ACCA
ACCA</td><td>71 2211 211 211 211 211 211 211 211 211</td><td>TGGC
TCQT
TCQT
TCQT
TGAG
TGAG
TGAG
TGAG
TGAG
TGAG
TGAG
TG</td><td>CCA
1
-*</td><td>81
GGAA
21
ACGG
61
ACCCC
01
CTCC
41
CTCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCCC
CTCCCC
CTCCCC
CTCCCC
CTCCCCCC
CTCCCC
CTCCCCCCC
CTCCCCCCCC</td><td>ACCC
AATT
TGCG
ACAA
ACCA
ACCT
ACCT
ACCT
ACCT
ACCA
ACCT
ACCA
ACCT
ACCA
ACCT
ACCA
ACCT
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA</td><td>9
AT C
23
TT T
37
TC -
37
TC -
51
GG C
-
51
GG C
-
51
CT G
-
51
CT G
-
51
CT G
-
51
CT G
-
-
51
CT G
-
-
-
-
-
-
-
-
-
-
-
-
-</td><td>1
CCCAGJ
1
TTATT
GATT:
GATT:
1
GAGGC
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCG
CCCCG
CCCCG
CCCG
CCCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCG
CCCG
CCCG
CCCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG</td><td></td><td>101
A CC
241
T CC

381
T GC
C CA

108
A TC

108
A TC

122C
C GC

136.
3 CT

136.
3 CT

136.
3 CT
</td><td>ACAG

GCCG

TATG
TCCCC

1
TGTA'
1
CTTTC
1
CCCCC
1
1
CCCCC
1
1
CCCCCC
1
1
CCCCCC</td><td>GAAF
GCGA
AACC
STTC
CGGA
AGTT
IGCA
STCT
CGCA
CCCG
CACC
CACC</td><td>111
A GCP
251
4 TTT

531
531
531
531
671
4 GTC
671
4 GTC
671
4 GTC

1091
4 TTC

1091
1
TTC

1511

1511

</td><td>IGGGT
GAGAGA
GAGAGA
CAGT
CAGT
ACACC
GCTT</td><td>ACGG
CGCT
CGCT
CGCT
CGCC
CCTG
CCTG</td><td>121
ACCI
CCA
CCA
GCA
GCA
GCA
GCA
GCA
GCA
GCA
GCA</td><td>CCGA
ACCGA
STCT4
SGTC4
CAAGT
CAAGT
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
S</td><td>11
GCA (
GAA (</td><td>31
PCACC
71
30
31
30
31
30
31
33
331
251
2551
260
391
332
331
331
331
331
331
331
33</td><td>ACCOT
ACCOT
TAAA
SGACC
SATGG
SGACC
SCACC
SCACC</td><td>A G A A A A A A A A A A A A A A A A A A</td><td></td></tr> | GGCGCGC
CCACGP
CCACGP
CCACGP
TTGGP
ATCAC
TCCAG
GACTG
GGCGA | Q
Q
Q
21
CA GF
161
CC CT
301
CC CT
301
CC CT
581
CC CT
581
CC CT
581
CC CT
721
S6 TA
CC
721
100
CC CT
100
CC CC CT
100
CC CT
100
CC CC CC CC
100
CC CC CC CC
CC CC CC
CC CC CC C | National Second | F
CCGGCC
F
CCGGCC
ACCCT
TTTG
TTTGA
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGGG
CCGGCC
TTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TCGGGC
CCG
CCG
CCG
CCG
CCG
CC | R
311
311
311
311
311
311
311
31 | K
SGGCZP
SGTCG
SGTCG
SGTCG
SGTCG
SGTTA
SGCCG
C
 | TAG
AGGG II
SCTT
*
JTGC
TTGT
55565

85675

1
S5675

1
S5675

1
S5675

1
S5675

1
S5675

1
S5775
 | 41
CGGG
A-
81
CCCA
61
GAAT

81
61
GAAT

81
91
41
* CACC
G
81
GTTA

91
161
AGTA

301
GCCA

 | GGGGG
CCTG
AGCA
AGAC
TGCC
AGAT
TGCC
CCTA
GAAT | 5
GG C | 1
AGGG
GAGGC

1
CCTAGG

1
1
CTCCC

1
1
CTCCC

3
1
TACGC

1
1
AGGCC

1
1
CTACG

1
CTACG

 | GGTC | 61
A CC
201
C CP
341
T CC
481
T AC
621
T AC
621
T AC
761
T GT
901
T GT
104
C TG
112
C AC
C AC | GGTTGAG
GGTCC
GGTCC
GGAG
GGAG
J
GGAG
J
GGAC
J
J
GCAC
J
J
TGGAG | GGGGC
GGGGG
ATCG
CCGC
TACC
AGGA
GCCG
GCCG | 71 211 211 771 351 777 491 7771 491 7771 493 7771 1051 7771 1051 1191 1331 776 776 777 | TGGC | 2007 2
2007 2
2007 2
30
2007 2
10
2007 2
10
10
10
10
10
10
10
10
10
10 | 81
GGAA
21
ACGG
61
ACCC
01
CTCC
81
ACCC
21
81
CTCC
81
ACCC
01
CTCC
81
ACCC
01
CTCC
81
CTCC
81
ACCC
01
CTCC
81
ACCC
91
CTCC
91
ACCC
91
CTCC
91
ACCC
91
ACCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCCC
91
ACCCCC
91
ACCCCC
91
ACCCCC
91
ACCCCCC
91
ACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC | ACCC
AATT
TGCG
ACAA
ACCA
ACCT
ACCT
ACCT
 | 9
AT C 23
TT T | 1
CCAGU
1
TTAT

1
GATT
1
GAGGC

1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACC
1
CACC
1
CACC
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT | AACC. | 101
A CC
241
T
C
381
C CA

521
G CC

521
G CC

801
G CG

801
G CG

108
A TC

122
C CC

136.5
S CT
 | ACAG | GAAF
GCGA
STTC
CGGA
AGGTI
GCA
STCT
STCT | 111
A GCA
251
A TTT
391
391
391
391
391
391
391
671
671
671
671
671
7 GTC
671
1091
1091
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095 | 43360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
443600
443600
443600
443600
443600
443600
4436000
4436000
443600 | CCCC
CCCT
CCCCT
CCCCT
CCCCT
CCCTC
CCCCCC
 | 121
ACCI
CCAJ
GCAT
GAT
GAT
GAT
GAC
GAT
GAC
GAT
GAT
GAT
GAT
GAT
GAT
TAC
GAT | CCGA
ACCGI
STCT-
CAAG
CAAG
CAAG
STTTJ | 11
GCA (
GCA (
GCA (
GCA (
GCA (
GCCA | 31
FCACC
71
CCANF
11
36TTF
31
36TTF
31
36TTF
31
36TTCAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
3 |
ACCOT
ATCOC
TTAAA
SGACC
SGACC
COCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
SGACCA
CCCA
TTCCA
CCCA
SGACCA
CCCA
TTCCA
CCCA
TTCCA
SGACCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
SGACCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
CCCA
STCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
C | | | Lamprey BGL2-H
LAMPrey BGL2-H
PROTEIN
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPR | A \\
1
A \\
1
A \\
1
A \\
1
A \\
1
A \\
1
A \\
1
A \\
1
4
2
3
1
CCAGGTC

2
3
4
2
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
1
2
2
1
2
2
1
CCAGGTC

1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1

 | SGGGT
SGGGT
SGGGT
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGA
SGGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA | F F F S1 GAGGU S1 GCTGGU S1 GCTGU S1 GAGUU S2 S1 GCTU S2 GAGUU GCTU GCTU GAUU GCTU GCTU GCTU GAUU | GGCGCC
CTCOT
CCACGZ
CCACGZ
TTTGGZ
ATCAC
TCCAG
GACTC
GACTC
GACTC | Q
Q
2
2
301
161
CC CT
301
161
CC CT
301
161
CC CT
441
40
CC AT
441
16C AT
581
100
581
100
56 AT
114
T AG
T AG
T AG
T AG
T AG
T AG
T AG
T AG | N
N
GGGGC
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C |
F
CCGGCC
F
CCGGCCA
ACCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA | R
31
31
7
311
7
300
311
7
300
311
7
300
301
7
311
7
300
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
30
7
30 | K
SGGCP
SCCCCC
CCGCA
SCCCCCCCCCCCCCCCCCCCCCCCCC | TAG
AGGG
CTT
CTT
CTT
CTT
CTT
CTT
CTT
CTT
CTT | 41
CGGG

21
CCCA
61
GAAC

81
GAAC

81
GAAC

81
GAAC

81
GAAC

81
GAAC

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

 | GGGGG
CCTG
AGCA
AGCA
TGCC
TGCC
AGAT
CCTA
GAAT | 5
GG C
- 19
GC A
- 33
GC A
- 47
AA C

GT G

- 61
GT G

 | 1
AGGG
1
CTAGG

1
CTAGG

1
TGTGJ
1
CTCCC

3
1
TGTGJ
3
1
CTCCC

1
AGGCC

1
AGGCC

1
CTAGG

1
CTAGG

1
CTAGG

1
CTAGG

1
CTAGG

1
CTAGG

1
CTAGG
 | GGCTC

CGAG

CCCCG

AGGGG

TGAG

TGAG

CGAC

CGAC
 | 61
A CC
2011
G GP

491
T C

621
T AC

761
T GC

104
C T GC

132
C AC

146
C AC

132
C AC | GCTGG
GCGT
TGAG
GGAG
GGAG
GGAG
1
TGGAG
1
TGGGG
1
TGGGG
1
CAC* | GGGCC
GGGCG
ATCG
CCGC
TACC
AGGA
GCCG
GCCG | 71 201 351 777 491 2491 2491 2491 2491 2491 2491 2491 | TGGC
TCGT
TCGT
TGGG
TGGG
TGGG
TGGG
GGTG
CGCA | CCA (
 | 81
GGAA
GGAA
21
ACGG

61
ACGG

61
ACGG

81
AGGG

81

81

81

81

04
CTGG

81

04
CTGG

81

201
CTGG

81

201
CTGG

81

201
CTGG

81

201
CTGG

201
CTGG

 | ACCC
 | 9
AT C
3
TT C
3
GG C
3
G
3
GG C
3
GG C
3
G
C
C
3
GG C
3
GG C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | 1
CCAG
1
TTAT:
1
CCCGG
CATT
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
C | AACC. | 101
A CC
241
T CC
381
C CA
521
G CC
521
G CC
52
G CC
52
C
52
C
52
C
52
C
52
C
52
C
52
C
5 | ACAG
GCCG
TATG
GCCG
TATG
GGTG
GGTG
ATGCI
1
TGTAT
1
CTTT
1
CACAG | GAAF
GCGA
AACC
STTC
CGGA
ATGT
IGCA
STCT
CCCG | 111
A GCA
251
391
370C
531
531
531
531
531
531
531
531 | 19397 | CCCC
CCCC
CCCCC
CCCCCCCCCCCCCCCCCCCCCC |
121
ACC
CCA
401
GAT
541
541
GAT
681
GAC
681
GAC
681
GAC
100
1101
TAC
1241
GAT
1241
GAT
1241
GAT
1241
GAT
1241
GAT | CCGA
Accor
STCT
STCT
SGGTCC
SAAG
STCT
SAAG
STTT
ACCTC | 11
GGA (
GGA (
GAT (| 31
FCACC
71
CCAAF
31
30
31
30
31
30
31
30
31
30
31
30
31
30
31
30
31
30
31
30
30
30
30
30
30
30
30
30
30 | ACCOT
ACCOT
ACCOT
ATCOC
TTAAA
SGACC
SGACC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC | A 19 19 19 19 19 19 19 19 19 19 19 19 19 | | Lamprey BGL2-H
LAMPrey BGL2-H
PROTEIN
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPR | A N A N 1 AGGTGG 141 CCGCCG 1281 CCACGC 2421 TCCCACG 701 CAGGTG 701 CAGGTG 701 CAGGTG 701 CAGGTG 701 CAGGTG 701 CAGGTG 1121 GTATAC 1261 GGTACT 1261 CATACF 1401 CATACF 1541 CCTCCTAC | SGGGT
SGGGT
SGGGA
SGGA
SGGA
SGGA
SGGA
SG | 11 F 11 GAGGA 91 CTCGG 31 GTCGGC 31 GTCGGC 93 11 GTCGCC 93 13 AACAC 271 131 AACAC 411 S51 CACAG | GGCCGC
CCCAGA
CCCCAGA
CCCCAGA
CCCCAGA
CCCCAGA
CCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCCAGA
CCCCCCAGA
CCCCCCAGA
CCCCCCAGA
CCCCCCAGA
CCCCCCAGA
CCCCCCCAGA
CCCCCCAGA
CCCCCCCAGA
CCCCCCCAGA
CCCCCCCAGA
CCCCCCCC | Q
Q
211
161
C C T
301
A GP
441
A C T
581
C C | AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG | ITCGGC
CCGCCA
ACCCT
ACCCT
TTTGA
TTTGA
ACCT
TTTGA
ACCT
TTGGG
ACCT
ACCT
TGGGG
ACCT
TGGGA
ACCT
TGGGA
CCCCA | R
31
31
311
7 AGA
591
591
591
591
591
591
591
591 |
K
GGGCP
CACCCC
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA | TAG
AGGG
CTT
STGC
TTT
STGC
TTT
STGC
TTT
STGC
TTT
STGC
TTT
STGC
TTT
STGC
STGC | 41
CGGGG
A-
81
TTCCC
CCCA

61
GATCC
41

81
GATCA

81
CAGTA

301
GCGGG

301
CCCA

301
CCCA

 | GGGGG
GGGGG
AGCACA
AGACA
TGGC
TGGG
AGAT
CCTA
GAAT.
CACACA | 5
GG C - 19
GG A - 33 3
GG A - 47
GG GT G - 61
GG GT G - 75
GG A - 61
GG C A - 75
GG A - 10
GT G - 10
GT G - 11
TG A A A - 14
CC - 15
TG T - 13
AA A - 14
TG C - 15
TG T - | 1 AGGGG CTAGG CTAGG CTAGG TGTGGGGAG 1 GGGGAG 1 TGTGGG 1 TGTGG 1 TGTGG 1 TGTGG 1 TAGGC 1 1 AGGCC 1 AGGCC 1 AGGCC 1 AGGC 1 AGG | GGCTCC
GCGAG
GCGCC
CACT
CCCCG
CCCCG
CCCCG
CCCCG
CCGAC
CGAC | 61
A CC
2011
G GC CP

341
G GC

621
T CC

621
T CC

761
T GC

104
C TG

104
C TG

104
C TG

106
C A

106
C A
 | GGGGG
GGGGG
GGGGG
GGGGG
GGGGG
GGGGG
CGGGG
CGGGG
CGGC
CGGC
CAC*
 | GGGC
GGGG
ATCG
CCGC
TACC
AGGA
ACCA
ACCA | 71 2211 211 211 211 211 211 211 211 211 | TGGC
TCQT
TCQT
TCQT
TGAG
TGAG
TGAG
TGAG
TGAG
TGAG
TGAG
TG | CCA 1
-* | 81
GGAA
21
ACGG
61
ACCCC
01
CTCC
41
CTCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCCC
CTCCCC
CTCCCC
CTCCCC
CTCCCCCC
CTCCCC
CTCCCCCCC
CTCCCCCCCC | ACCC
AATT
TGCG
ACAA
ACCA
ACCT
ACCT
ACCT
ACCT
ACCA
ACCT
ACCA
ACCT
ACCA
ACCT
ACCA
ACCT
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA | 9
AT C
23
TT T
37
TC -
37
TC -
51
GG C
-
51
GG C
-
51
CT G
-
51
CT G
-
51
CT G
-
51
CT G
-
-
51
CT G
-
-
-
-
-
-
-
-
-
-
-
-
- |
1
CCCAGJ
1
TTATT
GATT:
GATT:
1
GAGGC
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCG
CCCCG
CCCCG
CCCG
CCCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCG
CCCG
CCCG
CCCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG | | 101
A CC
241
T CC

381
T GC
C CA

108
A TC

108
A TC

122C
C GC

136.
3 CT

136.
3 CT

136.
3 CT
 | ACAG

GCCG

TATG
TCCCC

1
TGTA'
1
CTTTC
1
CCCCC
1
1
CCCCC
1
1
CCCCCC
1
1
CCCCCC | GAAF
GCGA
AACC
STTC
CGGA
AGTT
IGCA
STCT
CGCA
CCCG
CACC
CACC | 111
A GCP
251
4 TTT

531
531
531
531
671
4 GTC
671
4 GTC
671
4 GTC

1091
4 TTC

1091
1 TTC

1511

1511

 | IGGGT
GAGAGA
GAGAGA
CAGT
CAGT
ACACC
GCTT | ACGG
CGCT
CGCT
CGCT
CGCC
CCTG
CCTG | 121
ACCI
CCA
CCA
GCA
GCA
GCA
GCA
GCA
GCA
GCA
GCA | CCGA
ACCGA
STCT4
SGTC4
CAAGT
CAAGT
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
S | 11
GCA (
GAA (| 31
PCACC
71
30
31
30
31
30
31
33
331
251
2551
260
391
332
331
331
331
331
331
331
33 | ACCOT
ACCOT
TAAA
SGACC
SATGG
SGACC
SCACC
SCACC | A G A A A A A A A A A A A A A A A A A A | |
| GGCGCGC
CCACGP
CCACGP
CCACGP
TTGGP
ATCAC
TCCAG
GACTG
GGCGA | Q
Q
Q
21
CA GF
161
CC CT
301
CC CT
301
CC CT
581
CC CT
581
CC CT
581
CC CT
721
S6 TA
CC
721
100
CC CT
100
CC CC CT
100
CC CT
100
CC CC CC CC
100
CC CC CC CC
CC CC CC
CC CC CC C | National Second | F
CCGGCC
F
CCGGCC
ACCCT
TTTG
TTTGA
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGGG
CCGGCC
TTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGGG
CCGGC
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TTGA
ACCT
TCGGGC
CCG
CCG
CCG
CCG
CCG
CC

 | R
311
311
311
311
311
311
311
31 | K
SGGCZP
SGTCG
SGTCG
SGTCG
SGTCG
SGTTA
SGCCG
C | TAG
AGGG II
SCTT
*
JTGC
TTGT
55565

85675

1
S5675

1
S5675

1
S5675

1
S5675

1
S5675

1
S5775

 | 41
CGGG
A-
81
CCCA
61
GAAT

81
61
GAAT

81
91
41
* CACC
G
81
GTTA

91
161
AGTA

301
GCCA

 | GGGGG
CCTG
AGCA
AGAC
TGCC
AGAT
TGCC
CCTA
GAAT | 5
GG C | 1
AGGG
GAGGC

1
CCTAGG

1
1
CTCCC

1
1
CTCCC

3
1
TACGC

1
1
AGGCC

1
1
CTACG

1
CTACG
 | GGTC | 61
A CC
201
C CP
341
T CC
481
T AC
621
T AC
621
T AC
761
T GT
901
T GT
104
C TG
112
C AC
C AC |
GGTTGAG
GGTCC
GGTCC
GGAG
GGAG
J
GGAG
J
GGAC
J
J
GCAC
J
J
TGGAG | GGGGC
GGGGG
ATCG
CCGC
TACC
AGGA
GCCG
GCCG | 71 211 211 771 351 777 491 7771 491 7771 493 7771 1051 7771 1051 1191 1331 776 776 777 | TGGC | 2007 2
2007 2
2007 2
30
2007 2
10
2007 2
10
10
10
10
10
10
10
10
10
10 | 81
GGAA
21
ACGG
61
ACCC
01
CTCC
81
ACCC
21
81
CTCC
81
ACCC
01
CTCC
81
ACCC
01
CTCC
81
CTCC
81
ACCC
01
CTCC
81
ACCC
91
CTCC
91
ACCC
91
CTCC
91
ACCC
91
ACCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCC
91
ACCCCC
91
ACCCCC
91
ACCCCC
91
ACCCCC
91
ACCCCCC
91
ACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC | ACCC
AATT
TGCG
ACAA
ACCA
ACCT
ACCT
ACCT | 9
AT C 23
TT T |
1
CCAGU
1
TTAT

1
GATT
1
GAGGC

1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACT
1
CACC
1
CACC
1
CACC
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT
1
CACAT | AACC. | 101
A CC
241
T C
381
C CA

521
G CC

521
G CC

801
G CG

801
G CG

108
A TC

122
C CC

136.5
S CT
 | ACAG
 | GAAF
GCGA
STTC
CGGA
AGGTI
GCA
STCT
STCT | 111
A GCA
251
A TTT
391
391
391
391
391
391
391
671
671
671
671
671
7 GTC
671
1091
1091
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095
1095 | 43360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
44360
443600
443600
443600
443600
443600
443600
4436000
4436000
443600 | CCCC
CCCT
CCCCT
CCCCT
CCCCT
CCCTC
CCCCCC | 121
ACCI
CCAJ
GCAT
GAT
GAT
GAT
GAC
GAT
GAC
GAT
GAT
GAT
GAT
GAT
GAT
TAC
GAT | CCGA
ACCGI
STCT-
CAAG
CAAG
CAAG
STTTJ
 | 11
GCA (
GCA (
GCA (
GCA (
GCA (
GCCA | 31
FCACC
71
CCANF
11
36TTF
31
36TTF
31
36TTF
31
36TTCAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
31
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
37CAF
3 | ACCOT
ATCOC
TTAAA
SGACC
SGACC
COCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
CCCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTAAA
SGACCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
SGACCA
CCCA
TTCCA
CCCA
SGACCA
CCCA
TTCCA
CCCA
TTCCA
SGACCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
SGACCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
CCCA
STCCA
CCCA
TTCCA
CCCA
TTCCA
CCCA
C | | | | | | | | | | | | | | | |
 | | | | |
 | | | | | | | |
 | | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | | |
 | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | |
| Lamprey BGL2-H
LAMPrey BGL2-H
PROTEIN
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPR | A \\
1
A \\
1
A \\
1
A \\
1
A \\
1
A \\
1
A \\
1
A \\
1
4
2
3
1
CCAGGTC

2
3
4
2
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
6
1
CCAGGTC

5
1
2
2
1
2
2
1
CCAGGTC

1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1
2
1

 | SGGGT
SGGGT
SGGGT
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGA
SGGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA
SGGGA | F F F S1 GAGGU S1 GCTGGU S1 GCTGU S1 GAGUU S2 S1 GCTU S2 GAGUU GCTU GCTU GAUU GCTU GCTU GCTU GAUU

 | GGCGCC
CTCOT
CCACGZ
CCACGZ
TTTGGZ
ATCAC
TCCAG
GACTC
GACTC
GACTC | Q
Q
2
2
301
161
CC CT
301
161
CC CT
301
161
CC CT
441
40
CC AT
441
16C AT
581
100
581
100
56 AT
114
T AG
T AG
T AG
T AG
T AG
T AG
T AG
T AG |
N
N
GGGGC
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | F
CCGGCC
F
CCGGCCA
ACCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGGCCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA | R
31
31
7
311
7
300
311
7
300
311
7
300
301
7
311
7
300
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
301
7
30
7
30 | K
SGGCP
SCCCCC
CCGCA
SCCCCCCCCCCCCCCCCCCCCCCCCC | TAG
AGGG
CTT
CTT
CTT
CTT
CTT
CTT
CTT
CTT
CTT
 | 41
CGGG

21
CCCA
61
GAAC

81
GAAC

81
GAAC

81
GAAC

81
GAAC

81
GAAC

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

81

 | GGGGG
CCTG
AGCA
AGCA
TGCC
TGCC
AGAT
CCTA
GAAT | 5
GG C
- 19
GC A
- 33
GC A
- 47
AA C

GT G

- 61
GT G

 | 1
AGGG
1
CTAGG

1
CTAGG

1
TGTGJ
1
CTCCC

3
1
TGTGJ
3
1
CTCCC

1
AGGCC

1
AGGCC

1
CTAGG

1
CTAGG

1
CTAGG

1
CTAGG

1
CTAGG

1
CTAGG

1
CTAGG
 | GGCTC

CGAG

CCCCG

AGGGG

TGAG

TGAG

CGAC

CGAC
 | 61
A CC
2011
G GP

491
T C

621
T AC

761
T GC

104
C T GC

132
C AC

146
C AC

132
C AC | GCTGG
GCGT
TGAG
GGAG
GGAG
GGAG
1
TGGAG
1
TGGGG
1
TGGGG
1
CAC*
 | GGGCC
GGGCG
ATCG
CCGC
TACC
AGGA
GCCG
GCCG | 71 201 351 777 491 2491 2491 2491 2491 2491 2491 2491 | TGGC
TCGT
TCGT
TGGG
TGGG
TGGG
TGGG
GGTG
CGCA | CCA (
 | 81
GGAA
GGAA
21
ACGG

61
ACGG

61
ACGG

81
AGGG

81

81

81

81

04
CTGG

81

04
CTGG

81

201
CTGG

81

201
CTGG

81

201
CTGG

81

201
CTGG

201
CTGG

 | ACCC

 | 9
AT C
3
TT C
3
GG C
3
G
3
GG C
3
GG C
3
G
C
C
3
GG C
3
GG C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | 1
CCAG
1
TTAT:
1
CCCGG
CATT
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
1
CCCGG
C | AACC. | 101
A CC
241
T CC
381
C CA
521
G CC
521
G CC
52
G CC
52
C
52
C
52
C
52
C
52
C
52
C
52
C
5 | ACAG
GCCG
TATG
GCCG
TATG
GGTG
GGTG
ATGCI
1
TGTAT
1
CTTT
1
CACAG | GAAF
GCGA
AACC
STTC
CGGA
ATGT
IGCA
STCT
CCCG | 111
A GCA
251
391
370C
531
531
531
531
531
531
531
531 | 19397
 | CCCC
CCCC
CCCCC
CCCCCCCCCCCCCCCCCCCCCC | 121
ACC
CCA
401
GAT
541
541
GAT
681
GAC
681
GAC
681
GAC
100
1101
TAC
1241
GAT
1241
GAT
1241
GAT
1241
GAT
1241
GAT | CCGA
Accor
STCT
STCT
SGGTCC
SAAG
STCT
SAAG
STTT
ACCTC | 11
GGA (
GGA (
GAT (
 | 31
FCACC
71
CCAAF
31
30
31
30
31
30
31
30
31
30
31
30
31
30
31
30
31
30
31
30
30
30
30
30
30
30
30
30
30 | ACCOT
ACCOT
ACCOT
ATCOC
TTAAA
SGACC
SGACC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC
ACCOC | A 19 19 19 19 19 19 19 19 19 19 19 19 19 | | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | |
 | | | | | | | | | |
 | | | | | | | |
 | | | | | | | | | | | |
 | | | | | | | | | | | | | | |
| Lamprey BGL2-H
LAMPrey BGL2-H
PROTEIN
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-H
LAMPrey BGL2-C
LAMPrey BGL2-C
LAMPR | A N A N 1 AGGTGG 141 CCGCCG 1281 CCACGC 2421 TCCCACG 701 CAGGTG 701 CAGGTG 701 CAGGTG 701 CAGGTG 701 CAGGTG 701 CAGGTG 1121 GTATAC 1261 GGTACT 1261 CATACF 1401 CATACF 1541 CCTCCTAC | SGGGT
SGGGT
SGGGA
SGGA
SGGA
SGGA
SGGA
SG | 11 F 11 GAGGA 91 CTCGG 31 GTCGGC 31 GTCGGC 93 11 GTCGCC 93 13 AACAC 271 131 AACAC 411 S51 CACAG

 | GGCCGC
CCCAGA
CCCCAGA
CCCCAGA
CCCCAGA
CCCCAGA
CCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCAGA
CCCCCCAGA
CCCCCCAGA
CCCCCCAGA
CCCCCCAGA
CCCCCCAGA
CCCCCCAGA
CCCCCCCAGA
CCCCCCAGA
CCCCCCCAGA
CCCCCCCAGA
CCCCCCCAGA
CCCCCCCC | Q
Q
211
161
C C T
301
A GP
441
A C T
581
C C |
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG
AATAG | ITCGGC
CCGCCA
ACCCT
ACCCT
TTTGA
TTTGA
ACCT
TTTGA
ACCT
TTGGG
ACCT
ACCT
TGGGG
ACCT
TGGGA
ACCT
TGGGA
CCCCA | R
31
31
311
7 AGA
591
591
591
591
591
591
591
591 | K
GGGCP
CACCCC
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA
CCGCA | TAG
AGGG
CTT
STGC
TTT
STGC
TTT
STGC
TTT
STGC
TTT
STGC
TTT
STGC
TTT
STGC
STGC
 | 41
CGGGG
A-
81
TTCCC
CCCA

61
GATCC
41

81
GATCA

81
CAGTA

301
GCGGG

301
CCCA

301
CCCA

 | GGGGG
GGGGG
AGCACA
AGACA
TGGC
TGGG
AGAT
CCTA
GAAT.
CACACA | 5
GG C - 19
GG A - 33 3
GG A - 47
GG GT G - 61
GG GT G - 75
GG A - 61
GG C A - 75
GG A - 10
GT G - 10
GT G - 11
TG A A A - 14
CC - 15
TG T - 13
AA A - 14
TG C - 15
TG T - | 1 AGGGG CTAGG CTAGG CTAGG TGTGGGGAG 1 GGGGAG 1 TGTGGG 1 TGTGG 1 TGTGG 1 TGTGG 1 TAGGC 1 1 AGGCC 1 AGGCC 1 AGGCC 1 AGGC 1 AGG | GGCTCC
GCGAG
GCGCC
CACT
CCCCG
CCCCG
CCCCG
CCCCG
CCGAC
CGAC | 61
A CC
2011
G GC CP

341
G GC

621
T CC

621
T CC

761
T GC

104
C TG

104
C TG

104
C TG

106
C A

106
C A
 | GGGGG
GGGGG
GGGGG
GGGGG
GGGGG
GGGGG
CGGGG
CGGGG
CGGC
CGGC
CAC*
 | GGGC
GGGG
ATCG
CCGC
TACC
AGGA
ACCA
ACCA | 71 2211 211 211 211 211 211 211 211 211 | TGGC
TCQT
TCQT
TCQT
TGAG
TGAG
TGAG
TGAG
TGAG
TGAG
TGAG
TG | CCA 1
-*
 | 81
GGAA
21
ACGG
61
ACCCC
01
CTCC
41
CTCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
41
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCC
CTCCCC
CTCCCC
CTCCCC
CTCCCC
CTCCCCCC
CTCCCC
CTCCCCCCC
CTCCCCCCCC | ACCC
AATT
TGCG
ACAA
ACCA
ACCT
ACCT
ACCT
ACCT
ACCA
ACCT
ACCA
ACCT
ACCA
ACCT
ACCA
ACCT
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA
ACCA | 9
AT C
23
TT T
37
TC -
37
TC -
51
GG C
-
51
GG C
-
51
CT G
-
51
CT G
-
51
CT G
-
51
CT G
-
-
51
CT G
-
-
-
-
-
-
-
-
-
-
-
-
- |
1
CCCAGJ
1
TTATT
GATT:
GATT:
1
GAGGC
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCCG
CCCG
CCCCG
CCCCG
CCCG
CCCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCG
CCCG
CCCG
CCCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG
CCG | | 101
A CC
241
T CC

381
T GC
C CA

108
A TC

108
A TC

122C
C GC

136.
3 CT

136.
3 CT

136.
3 CT
 | ACAG

GCCG

TATG
TCCCC

1
TGTA'
1
CTTTC
1
CCCCC
1
1
CCCCC
1
1
CCCCCC
1
1
CCCCCC | GAAF
GCGA
AACC
STTC
CGGA
AGTT
IGCA
STCT
CGCA
CCCG
CACC
CACC | 111
A GCP
251
4 TTT

531
531
531
531
671
4 GTC
671
4 GTC
671
4 GTC

1091
4 TTC

1091
1 TTC

1511

1511

 | IGGGT
GAGAGA
GAGAGA
CAGT
CAGT
ACACC
GCTT | ACGG
CGCT
CGCT
CGCT
CGCC
CCTG
CCTG | 121
ACCI
CCA
CCA
GCA
GCA
GCA
GCA
GCA
GCA
GCA
GCA |
CCGA
ACCGA
STCT4
SGTC4
CAAGT
CAAGT
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
SGTC4
S | 11
GCA (
GAA (| 31
PCACC
71
30
31
30
31
30
31
33
331
251
2551
260
391
332
331
331
331
331
331
331
33 | ACCOT
ACCOT
TAAA
SGACC
SATGG
SGACC
SCACC
SCACC | A G A A A A A A A A A A A A A A A A A A | |
 | | | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | | | | | | | | | |
 | | | | | | | | | | | | |
 | | | | | | | | |
 | | | | | | | | | | | | | |
 | | | | | | | | | | | |
 | | | | |

Fig. 4. Alignment of **(A)** *BGL1* and **(B)** *BGL2* nucleotide sequences obtained from *P. marinus* species from Lakes Huron and Champlain. A dash (–) indicates identity with the upper sequence, an asterisk (*) an alignment gap, and a dot (.) unavailability of sequence information.

Coding regions are presented and numbered by codon beginning at the first nucleotide. 5' and 3' noncoding regions are numbered separately in blocks of 10 nucleotides. A translation in the single letter amino acid code is given below the sequences.

that there are other, apparently functional proteoglycan genes that similarly lack this site. The contraindications to the pseudogene status of *BGL3* are, first, the fact that the gene is expressed in both *Oreochromis* and zebrafish and has no obvious inactivating defects, and second, that

the ratio of synonymous to nonsynonymous substitutions in comparisons of zebrafish and *Oreochromis BGL3* is at least 4. (Synonymous substitutions are saturated, giving Ks values of >2, while nonsynonymous substitutions give a Ka value of 0.24.) The fact that synonymous sub-

TTC GCG GCG GCG F A A A CTC TCT 999 GAT 990 Lamprey BGL1-C THE ASC CCG GCG GAT GAG GCC CTG GAA 61 CCC AAA CCA CCG GCG GCG CCC AAA CCA CCG Lamprey BGL1-C Lamprey BGL1-H PROTEIN TCG GCG GCG CCC AAA CCA CCG ACC TCC CCC GTG GCG CCC CCT A A ĸ P P P P P к P ---I R v A p т s G Ð D P к P А р v 0 D А P Lamprey BGL1-C Lamprey BGL1-H PROTEIN с F v D s o А 121 A4 151 151 ACC ARG CAA GAC CAC CAC ATG CGC CAG CTG CAC CGC CTG TC CTG GTG AAC AAC CTC ATG CGG AAG ATA CAC CCC AAG GGC TTT GCG CCC ATG GTG AGC Lamprey BGL1-C Lamprey BGL1-H PROTEIN E т D D L LH ALF LVNN T. τ. Δ. ĸ p к A F к 0 ۴ к G А ٥ т н Lamprey BGL1-C Lamprey BGL1-H ---C G I S н т v т P LIELRV HEN v Y N R т. Е P P s L I KR P к ____ m F T N PROTEIN . Lamprey BGL1-C Lamprey BGL1-H PROTEIN т. н v Е L G к ₽ L ---s G I E v A FN G ____ к T I N s G L T. ¥ I R т. 241 251 261 271 261 271 CAG CTG CCC AAG GAG CTG CCC CTG GAG GGA CTG GAC CTG GAG GGA AAT GAG ATC GTT GCC ATT GAG GAT GAA GAC CTC TTC GGG TAC CCC TTC CGA CTG GGT TTG AGC Lamprey BGL1-C Lamprey BGL1-H PROTEIN 0 τ. L G E L L E 1 I D т. ₽ N s E L н Е G N v A Е Е п L F Y 281 TAC AAG AAG ATC ACA GAG GTG CAG AAT GGC AGC CTG GCC GTG AGC GGC AAC CTG CGC GAG CTT CAC CTG GAT AAC AAC CTC CTG GTC AGC GTG CGG CCG GGC CTG TCC AAG X N K I T E V O N G S T D Y G G Y T Lamprey BGL1-C Lamprey BGL1-H PROTEIN 331 341 TTC TGC CCA 321 CTC AAC Lamprey BGL1-C Lamprey BGL1-H PROTEIN GAG GTG AAG CCC ACC ACG GTG TTC AGC CCC AAG CGC GCC CAG TAT AAG AAG GAC GCC GGC ATC т. к ----Ε ---ĸ P ----D F с P v --Q 371 381 AAG TAC TGG GAG GTC CCC AGC GTC TTC CGC TGC GTC CAC AAC CAC AAC GCA ATC CAC TAT GGC AGC AAC TAC CGC AAA TAG Lamprey BGL1-C Lamprey BGL1-H PROTEIN ----v Y w v ----D ----D s v F R с н N н M А r - ----u F G 21 31 41 51 61 CGCATTTACA CTCGCCGGGT AATCCGGCAA CCCAAGCGGA CACTTACACG ccoccococ croccococ басааста Lamprey BGL1-C Lamprey BGL1-H 211 171 181 191 201 221 231 151 161 241 GCCAGCCAGC Lamprey BGL1-C Lamprey BGL1-H ACCAAACAGA CGCTTACAAA GCCAGCTCGC CTGCTAGCTT GCTAACCAGC CAAAGAGACA тарасатата 321 CA GCC 331 Lamprey BGL1-C Lamprey BGL1-H 451 461 471 481 491 501 511 521 9 GTCCGTTTTG AAAGAAACAA AACAAAATCA TTCCAATAAT CGCTAAAAGT CTGAGAAAGG CGGTAGTCAC CGCATCGCGC Lamprey BGL1-C Lamprey BGL1-H Lamprey BGL1-C Lamprey BGL1-H AATACTCTTC CTTCCGATAT TAGGAAACCA CTGGAGCCAC GCACTTTTAA AATCTACATT GAAAACTAAT TTCTTTAGAA CTGCTTTGTG TGTGTTAGTT 741 751 ACACGTACAA AGTGATGTAA AGGGGTAGCG TAGCAGACAC AGCTGTTAAC CAGTGCTCAT GGCACCAACT Lamprey BGL1-C Lamprey BGL1-H TACGTA CCCGCGTCAA GTTTCACACC GCACCGGTAA AGTCAACCAT GTTTGGTTCG ACCGTGT 1031 1041 GAGAATAACA ATAAATCATA 1051 ATACCATG** 1061 1071 1081 1091 annaa Lamprey BGL1-C Lamprey BGL1-H AG TAAATTTTCC TAAATGCTGT ATGGTATTTA ACAGCTAACC AATCGGGCAA GCATT 1141 1151 1161 1171 1181 1191 1201 1211 1221 1221 1241 1251 Lamprey BGL1-C Lamprey BGL1-H CACCCTTACA TITTAACGCC CECETTACTT TEAACTCCCC TECERACACT CEETCCAACA AAAAACCER GEREGEGEGE CASTCTTECA AACCCETEGA ACCTCTCCCC CEACAATCCT TECERACAC CCEACCATT

 1261
 1271
 1281
 1291
 1301
 1311
 1321
 1331
 1341
 1351
 1361
 1371
 1391

 Lamprey BGLI-C
 TTCAAAGTCA AGAACAACT THTCCAAATC GAATGCAAAC TAACAAATA GCGGCCCCCC CCCTCGTGGG GTCTCCCCCG AAGGTCACAT GACCGGCTCC ATCCTCCTCC ACGCGTGTGT CAAATGTTTT CCGCCTCGCGC

Fig. 4 Continued.

stitutions predominate indicates that the gene has been functional for much if not all of the time since the divergence of zebrafish and *Oreochromis*. Taken together, the most likely of the three possibilities is, at present, the first one.

As for the relationship of the lamprey *BGL1* and *BGL2* sequences to biglycans and decorins, we note that all efforts to find a decorin-like gene in the lamprey have failed. To determine whether any other biglycan- or decorin-related genes might be present in the lamprey genome, a Southern blot was made under conditions of reduced stringency. The probes used were >300-nucleotide-long fragments from conserved parts of either

the *BGL1* or *BGL2* genes (Fig. 1; positions 125–242 for *BGL1* and 125–237 for *BGL2*). The Southern blot (Fig. 3) shows between one (*Bam*HI digestion) and five (*MspI* digestion) hybridizing bands of varying intensity to be present in the lamprey genome, regardless of which probe was used. An internal *MspI* site is found in both *BGL1* and *BGL2* probes, producing the weaker additional bands seen in the *MspI* digestion. For both probes with any of the enzymes, the hybridization pattern of bands and their positions are strikingly similar—the major difference appears to be the strength of the hybridization signal rather than the band positions. The relative strength of the bands is also antithetical—bands that hy-

bridize strongly to the *BGL1* probe appear to do so weakly to the *BGL2* probe and vice versa. We interpret this to indicate that each probe hybridizes to both *BGL1* and *BGL2* under the conditions used and further that any similarly related genes should give a hybridization signal. The absence of any bands additional to *BGL1* and *BGL2* indicates that additional loci do not exist in the lamprey or, alternatively, that any such loci are very dissimilar to *BGL1* and *BGL2*. Using the enzyme *Bam*HI (Fig. 3), only one strongly hybridizing band of large size appears to be present in the digest. One possible explanation of this is that the *BGL1* and *BGL2* loci are linked on a single digestion fragment.

At present, therefore, it seems reasonable to conclude that the gene duplication leading to the emergence of biglycans and decorins occurred after the divergence of Agnatha and Gnathostomata and that *BGL1* and *BGL2* are more closely related to the ancestor of biglycan and decorin than to either of the two. The slightly higher similarity of lamprey *BGL* sequences to biglycan than to decorin might be the result of *BGL1* and *BGL2* having biglycan-like (rather than decorin-like) functions. As for the origin of *BGL1* and *BGL2*, it seems that these two genes are the result of an independent duplication that occurred in the agnathan (lamprey) lineage after its divergence from gnathostomes.

Sequence Variation Within Lampreys

Sequences of both BGL1 and BGL2 determined from P. marinus species collected in Lakes Champlain and Huron were not identical. Figure 4 shows nucleotide alignments of sequences obtained from both sources. In the case of BGL1, two synonymous and four nonsynonymous differences are found in the 348 codons of shared coding sequence. A further six differences occur in more than 1 kb of 3' untranslated (UT) sequence. This corresponds to an overall nucleotide distance of 0.0057 substitutions per site. Furthermore, the polyadenylation site appears to differ between the sequences. In the case of BGL2 no differences are found in the 310 shared codons. The 3' UT sequence of BGL2 appears to contain an extended segment of CA repeats, as well as runs of single-nucleotide repeats. Polyadenylation also appears to occur in different positions in sequences obtained from Lakes Champlain and Huron, and this makes alignment of the 3' end difficult. In the approximately 1.5 kb of clearly alignable 3' UT sequence, 12 substitutions as well as many indels are found. This corresponds to an overall distance of 0.0049 substitutions per site, similar to that found for BGL1. Although little is known about rates of molecular evolution in lampreys, the degree of difference in sequence between lamprey from Lakes Champlain and Huron is surprisingly high and presumably has taken many years to accumulate. The population from Lake Huron is thought to have been introduced in

the 1920s from an unknown source, whereas the Lake Champlain population has probably existed since the last glaciation. The specimens may represent sibling species of *P. marinus*. However, more genes will need to be characterized to determine the true degree of differentiation between the populations.

Interpretation of Biglycan-Decorin Evolution

On the strength of the available evidence it seems reasonable to suggest that agnathans, as represented by the two Petromyzon species used in the present study, possess small proteoglycans that may not have as yet differentiated functionally into biglycan and decorin. Instead, the gene encoding the ancestral proteoglycan independently duplicated into two copies, BGL1 and BGL2, which appear to have remained functionally undifferentiated. The duplication of the ancestral gene that ultimately led to biglycan and decorin appears to have taken place after the divergence of agnathans from gnathostomes but before the divergence of teleostei (or at the latest before the divergence of Perciformes here represented by the Oreochromis cichid and Cypriniformes represented by the zebrafish). Teleosts have two presumably functionally differentiated proteoglycans, one closely related to the tetrapod decorin and the other (BGL3) allied with the BGLs of the lamprey. The acquisition of the decorin function by one of the duplicated genes was accompanied by a slowdown of its evolutionary rate with the result that this gene can now be identified unambiguously as decorin. The second of the duplicated teleost genes may have kept its original function and a relatively rapid evolutionary rate until the emergence of tetrapods. This supposition may explain why the teleost BGL3 appears to have diverged earlier than it in fact apparently did. In tetrapods, the BGL3 acquired a biglycan function and slowed down its evolutionary rate. This slowdown was more dramatic than that of the decorin gene, with the result that the branches in the biglycan clade are shorter than in the decorin clade, especially in mammals.

Supporting this view, the evidence from mammalian sequences suggests that although the overall evolutionary rates are now slow, there is some rate variation. The rat and mouse biglycan have only two residue differences over their entire length, and the sheep-cattle biglycan only three differences. Mammalian decorins appear to evolve more rapidly than mammalian biglycans. The sheep and cattle decorin sequences have five differences, whereas the rat and mouse sequences differ in 44 residues. The mouse and rat decorin sequences appear quite distinctive and are almost as distant as fowl from other mammals even in their conserved regions. The implication is that the evolutionary rate for the sequences can change, perhaps as a consequence of altered functions. In this view, the changing evolutionary rates can explain the complex relationships found in the gene family in gnathostomes.

References

- Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD (1989) Molecular biology of the cell. New York: Garland Publishing
- Blochberger TC, Vergen JP, Hempel J, Hassel JR (1992) cDNA to chick lumican (corneal keratan sulfate proteoglycan) reveals homology to the small interstitial proteoglycan gene family and expression in muscle and intestine. J Biol Chem 267:347–352
- Bourdon MA, Krusius T, Campbell S, Schwartz NB, Ruoslahti E (1987) Identification and synthesis of a recognition signal for the attachment of glycosaminoglycans to proteins. Proc Natl Acad Sci USA 84:3194–3198
- Boykiw RH, Sciore P, Reno C, Marchuk L, Frank C, Hart DA (1998) Altered levels of extracellular matrix molecules mRNA in healing rabbit ligaments. Matrix Biol 171:31–378
- Day AA, Ramis CI, Fisher LW, Gehron-Robey P, Termine JD, Young MF (1986) Characterization of bone PG II cDNA and its relationship to PG II mRNA from other connective tissues. Nucleic Acids Res 14:9861–9876
- Day AA, McQuillan CI, Termine JD, Young MR (1987) Molecular cloning and sequence analysis of the cDNA for small proteoglycan II of bovine bone. Biochem J 248:801–805
- Dreher KL, Asundi V, Matzura D, Cowan K (1990) Vascular smooth muscle biglycan represents a highly conserved proteoglycan within the arterial wall. Eur J Cell Biol 53:296–304
- Dunlevy JR, Neame PJ, Vergnes JP, Hassell JR (1998) Identification of the N-linked oligosaccharide sites in chick corneal lumican and keratocan that receive keratan sulfate. J Biol Chem 273:9615–9621
- Fisher LW, Termine JD, Dejter Jr SW, Whitson SW, Yanagishita M, Kimura JH, Hascall VC, Kleinman HK, Hassell JR, Nilsson B (1983) Proteoglycans of developing bone. J Biol Chem 258:6588– 6594
- Fisher LW, Termine JD, Young MF (1989) Deduced protein sequence of bone small proteoglycan I (biglycan) shows homology with proteoglycan II (decorin) and several nonconnective tissue proteins in a variety of species. J Biol Chem 264:4571–4576
- Fisher LW, Heegaard A-M, Vetter U, Vogel W, Just W, Termine JD, Young MF (1991) Human biglycan gene. Putative promoter, intron-exon junctions, and chromosomal localization. J Biol Chem 266:14371–14377
- Gehron Robey P (1996) Vertebrate mineralized matrix proteins: Structure and function. Connect Tissue Res 35:131–136
- Gerhart J, Kirschner M (1997) Cells, embryos, and evolution. Toward a cellular and developmental understanding of phenotypic variation and evolutionary adaptability. Malden, MA: Blackwell Science Publications
- Gilbert DG (1995a) SeqPup version 0.4j: a biosequence editor and analysis application. http://iubio.bio.indiana.edu/soft/molbio
- Gilbert DG (1995b) DottyPlot, version 1.0c. http://iubio.bio.indiana. edu/soft/molbio
- Grover J, Chen XN, Korenberg JR, Roughly PJ (1995) The human lumican gene. Organization, chromosomal location, and expression in articular cartilage. J Biol Chem 270:21942–21949
- Hocking AM, Shinomura T, McQuillan DJ (1998) Leucine-rich repeat glycoproteins of the extracellular matrix. Matrix Biol 17:1–19
- Iozzo RV, Murdoch AD (1996) Proteoglycans of the extracellular environment: clues from the gene and protein side offer novel perspectives in molecular diversity and function. FASEB J 10:598– 614
- Kobe B, Deisenhofer J (1994) The leucine-rich repeat: a versatile binding motif. Trends Biochem Sci 19:415–421

Kokenyesi R, Woessner JFJ (1989) Purification and characterization of

a small dermatan sulphate proteoglycan implicated in the dilatation of the rat uterine. Biochem J 260:413–419

- Kresse H, Hausser H, Schönherr E (1994) Small proteoglycans. In: Jollès P (ed) Proteoglycans. Basel, Switzerland: Birkäuser Verlag, pp 73–100
- Krusius T, Ruoslahti E (1986) Primary structure of an extracellular matrix proteoglycan core protein deduced from cloned cDNA. Proc Natl Acad Sci USA 83:7683–7687
- Kumar S, Tamura K, Nei M (1993) MEGA: Molecular evolutionary genetic analysis, version 1.01. University Park, PA: The Pennsylvania State University
- Li W-H, Wu C-I, Luo C-C (1985) A new method for estimating synonymous and nonsynonymous rates of nucleotide and codon changes. Mol Biol Evol 2:150–174
- Li W, Vergnes J-P, Cornuet PK, Hassell JR (1992) cDNA clone to chick corneal chondroitin/dermatan sulfate proteoglycan reveals identity to decorin. Arch Biochem Biophys 296:190–197
- Marcum JA, Torok M, Evans S (1993) cDNA sequence for bovine biglycan (PGI) protein core. Biochim Biophys Acta 1173:81–84
- Mayer WE, Tichy H (1995) A cDNA clone from the sea lamprey *Petromyzon marinus* coding for a scavenger receptor Cys-rich (SRCR) domain protein. Gene 164:267–271
- Neame PJ, Choi HU, Rosenberg LC (1989) The primary structure of the core protein of the small, leucine-rich proteoglycan (PGI) from bovine articular cartilage. J Biol Chem 264:8653–8661
- Nurminskaya MV, Birk DE (1996) Differential expression of fibromodulin mRNA associated with tendon fibril growth: isolation and characterization of a chicken fibromodulin cDNA. Biochem J 317: 785–789
- Pulkkinen L, Alitalo T, Krusius T, Peltonen L (1992) Expression of decorin in human tissues and cell lines and defined chromosomal assignment of the gene locus (DCN). Cytogenet Cell Genet 60: 107–111
- Scholzen T, Solursh M, Suzuki S, Reiter R, Morgan JL, Buchberg AM, Siracusa LD, Iozzo RV (1994) The murine decorin. Complete cDNA cloning, genomic organization, chromosomal assignment, and expression during organogenesis and tissue differentiation. J Biol Chem 269:28270–28281
- Swofford DL (1998) PAUP: Phylogenetic analysis using parsimony (and other methods), Version 4. Sunderland, MA: Sinauer Associates
- Tasheva ES, Funderburgh JL, Funderburgh ML, Corpuz LM, Conrad GW (1999) Structure and sequence of the gene encoding human keratocan. DNA Seq 10:67–74
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) the ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882
- Toyosawa S, O'hUigin C, Figueroa F, Tichy H, Klein J (1998) Identification and characterization of amelogenin genes in monotremes, reptiles, and amphibians. Proc Natl Acad Sci USA 95:13056–13061
- Toyosawa S, O'hUigin C, Klein J (1999a) Dentin matrix protein 1 gene of prototherian and metatherian mammals. J Mol Evol 48:160–167
- Toyosawa S, O'hUigin C, Tichy H, Klein J (1999b) Characterization of dentin matrix protein 1 gene in crocodilia: implications for the evolution of mineralized tissues. Gene 234:301–314
- Toyosawa S, Sato A, O'hUigin C, Tichy H, Klein J (2000) Expression of the dentin matrix protein 1 gene in birds. J Mol Evol 50:31–38
- Traupe H, van den Ouweland AMW, van Oost BA, Vogel W, Vetter U, Warren ST, Rocchi M, Darlison MG, Ropers H-H (1992) Fine mapping of the human biglycan (BGN) gene within the Xq28 region employing a hybrid cell panel. Genomics 13:481–483
- Vetter U, Vogel W, Just W, Young MF, Fisher LW (1993) Human decorin gene: intron-exon junctions and chromosomal localization. Genomics 15:161–168
- Wegrowski Y, Pillarisetti J, Danielson KG, Suzuki S, Iozzo RV (1995) The murine biglycan: complete cDNA cloning, genomic organization, promoter function, and expression. Genomics 30:8–17