

Matrix Biology 19 (2000) 693-698



www.elsevier.com/locate/matbio

Short note

Molecular cloning and relative tissue expression of keratocan and mimecan in embryonic quail cornea

Lolita M. Corpuz^a, Jane R. Dunlevy^b, John R. Hassell^c, Abigail H. Conrad^a, Gary W. Conrad^{a,*}

^aDivision of Biology, Kansas State University, Manhattan, KS 66506-4901, USA ^bDepartment of Anatomy and Cell Biology, University of North Dakota, Grand Forks, ND 58203, USA ^cShriners Hospital for Children-Tampa and University of South Florida, College of Medicine, 12502 North Pine Drive, Tampa, FL33612-9499, USA

Received 4 August 2000; received in revised form 28 August 2000; accepted 29 August 2000

Abstract

We have cloned and sequenced the cDNAs for quail cornea keratan sulfate proteoglycan core proteins, keratocan and mimecan. The deduced quail keratocan protein contains a single conservative amino acid difference from the chick sequence, whereas quail mimecan protein contains a 58 amino acid-long avian-unique sequence that shares no homology with mammalian mimecan. Ribonuclease protection assay of Day 16 embryonic quail tissues reveals that keratocan and lumican are expressed at highest levels in cornea, whereas mimecan mRNA is expressed at a much lower level. Keratocan is expressed only in quail cornea, whereas mimecan is expressed in many different tissues as four transcripts of different sizes. Both lumican and mimecan are expressed at lowest levels in brain, liver and sternum. © 2000 Elsevier Science B.V./International Society of Matrix Biology. All rights reserved.

Keywords: Cornea; Extracellular matrix; Keratocan; Mimecan; Proteoglycans; Quail

1. Introduction

The remarkable transparency of the cornea is a function of the very small, uniform diameter and regular spacing of stromal collagen fibrils and minimal hydration of the extracellular matrix. These properties are regulated in part by the keratan sulfate proteoglycans (KSPGs) of the corneal stroma. Three major corneal KSPGs have been identified in vertebrates: lumican (Blochberger et al., 1992), keratocan (Corpuz et al., 1996; Dunlevy et al., 1998) and mimecan (Funderburgh et al., 1997; Dunlevy et al., 2000). The gene that encodes mimecan (Funderburgh et al., 1997) was originally identified as that for osteoglycin (Madisen et al., 1990). Core protein analyses have shown that KSPGs belong to the small leucine-rich proteoglycan (SLRP) family (Iozzo, 1997). In adult mammals, keratocan mRNA is expressed almost exclusively in cornea and sclera (Corpuz et al., 1996), whereas lumican and mimecan mRNAs are expressed in many different tissues (Corpuz et al., 1996; Funder-

^{*}Corresponding author. Tel.: +1-785-532-6662; fax: +1-785-532-6653.

E-mail address: gwconrad@ksu.edu (G.W. Conrad).

	-91
C B GGAAGCAAACACAGTGGTTAGCAGCCAAAGGCTGTAACAGTGATTTCTAACTGGCTGCAGACAGCTCCAGGAACTTGACTAGCAAGGCAA	-1
M M T L K V C P S L L L F L V H S V W T R T V R Q V Y N E	90
L D P E H W S H Y T F E C P Q E C F C P P S F P N A L Y C D	180
9	
AACAAAGGACITAAAGAAATACCIGCAATCCCAGCAAGAATTIGGTACCICTATCITCAAAACAACCTAATIGAAACAATTICAGAGAAG N K G L K E I P A I P A R I W Y L Y L Q N N L I E T I S E K	270
g C	
CCTTTTGTTAATGCCACTCATTTAAGATGATAAATTTGAACAAGAATAAAATCACCAACAATGGAATTGAGAGTGGTGTGCCTCAGCAAG	360
t g a	
CIGAAAAGGCIGITGIACTIATICCITIGAAGATAATGAAGAGGIGCCAGCACCACTACCAGTGGCCITIGGAACAGCIGAGGCIG	450
LKRLLYLFLEDNELEEVPAPLPVGLEQLRL	
C t	
GCTAGAAACAAAATCTCCAGAATCCCAGAAGGAGTCTTCAGCAATTTGGAAAAATCTCACTATGTTAGATTTGCACCAGAATAACTTGTTG	540
A R N K I S R I P E G V F S N L E N L T M L D L H Q N N L L	
α	
GACAGCGCTCTTCAAAGTGACACCTTCCAAGGACTCAACAGTCTTATGCAACTCAACATAGCAAAAAATTCCCTCAAAAAGATGCCTTTA	630
D S A L Q S D T F Q G L N S L M Q L N I A K N S L K K M P L	
C	
AGCATTCCAGCTAACACACTGCAGCTGTTTTTAGACAACAACAACACTTCAAGTTATACCAGAAAACTACTTCAGTGCAATACCCCAAAGTG	720
S I P A N T L Q L F L D N N S I E V I P E N Y F S A I P K V	
그는 것 같은 것 같은 것이 같은 것 같은 것 같은 것 같은 것 같은 것 같	
ACTTTCCTTAGGCTGAACTACAATTAACTTATCTGATGAGGCATTCCCCAAATGGGTTTAATGTTTCAACTACTACTACTACTACAACTACAACTTA	
	810
TFLRLNYNKLSDEGIPPNGFNVSSILDLQL D	810
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D	810
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D TCTCATAACCAGCTCACTAAAATTCCACCGATCAATGCTCATCTCGAGCACCTTCACCTTGATCACAACAGAATCAAAAGTGTCAATGGC S H N Q L T K I P P I N A H L E H L H L D H N R I K S V N G	810 900
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D TCTCATAACCAGCTCACTAAAATTCCACCGATCAATGCTCATCTCGAGCACCTTCACCTTGATCACAACAGAATCAAAAGTGTCAATGGC S H N Q L T K I P P I N A H L E H L H L D H N R I K S V N G	810 900
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D TCTCATAACCAGCTCACTAAAATTCCACCGATCAATGCTCATCTCGAGCACCTTCACCTTGATCACAACAGAATCAAAAGTGTCAATGGC S H N Q L T K I P P I N A H L E H L H L D H N R I K S V N G t ACTCAGATATGCCCAGTCTCAATTGCAGTAGCAGAAGACTATGGTCTTTATGGTAACATCCCTCGCCTTCGTTTATCTTCGCTTGGATGGA	810 900 990
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D TCTCATAACCAGCTCACTAAAATTCCACCGATCAATGCTCATCTCGAGCACCTTCACCTTGATCACAACAGAATCAAAAGTGTCAATGGC S H N Q L T K I P P I N A H L E H L H L D H N R I K S V N G t ACTCAGATATGCCCAGTCTCAATTGCAGTAGCAGAAGACTATGGTCTTTATGGTAACATCCCTCGCCTTCGTTATCTTCGCTTGGATGGA	810 900 990
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D TCTCATAACCAGCTCACTAAAATTCCACCGATCAATGCTCATCTCGAGCACCTTCACCTTGATCACAACAGAATCAAAAGTGTCAATGGC S H N Q L T K I P P I N A H L E H L H L D H N R I K S V N G t ACTCAGATATGCCCAGTCTCCAATGCCAGTAGCAGAAGACTATGGTCTTTATGGTAACATCCCTCGCCTTCGTTATCTTCGCTTGGATGGA	810 900 990 1080
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	810 900 990 1080
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	810 900 990 1080
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	810 900 990 1080 1170
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D TCTCATAACCAGCTCACTAAAATTCCACCGATCAATGCTCATCTCGAGCACCTTCACCTTGATCACAACAGAATCAAAAGTGTCAATGGC S H N Q L T K I P P I N A H L E H L H L D H N R I K S V N G t ACTCAGATATGCCCAGTCTCCAATGCCAGTAGCAGAAGACTATGGTCTTTATGGTAACATCCCTCGCCTTCGTTATCTTCGCTTGGATGGA	810 900 990 1080 1170
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D TCTCATAACCAGCTCACTAAAATTCCACCGATCAATGCTCATCTCGAGCACCTTCACCTTGATCACAACAGAATCAAAAGTGTCAATGGC S H N Q L T K I P P I N A H L E H L H L D H N R I K S V N G t ACTCAGATATGCCCAGTCTCAATTGCCAGTAGCAGAGACTATGGTCTTTATGGTAACATCCCCCGCCTTCGCTTATCTTCGCTTGGATGGA	 810 900 990 1080 1170 1260
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D TCTCATAACCAGCTCACTAAAATTCCACCGATCAATGCTCATCTCGAGCACCTTCACCTTGATCACAACAGAATCAAAAGTGTCAATGGC S H N Q L T K I P P I N A H L E H L H L D H N R I K S V N G t ACTCAGATATGCCCAGTCCTCAATGCCAGTAGCAGAAGACTATGGTCTTTATGGTAACATCCCTCGCCTTCGTTATCTTCGCTTGGATGGA	 810 900 990 1080 1170 1260
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D TCTCATAAACCAGCTCACTAAAATTCCACCGATCAATGCTCATCTCGAGCACCTTCACCTTGATCACAACAGAATCAAAAGTGTCAATGGC S H N Q L T K I P P I N A H L E H L H L D H N R I K S V N G t ACTCAGATATGCCCAGTCTCAATTGCAGTAGCAGAAGACTATGGTCTTTATGGTAACATCCCTCGCCTTCGTTATCTTCGCTTGGATGGA	 810 900 990 1080 1170 1260 1350
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D TCTCATAACCAGGTCACTAAAATTCCACCGATCAATGCTCATCTGGAGGACCTTCACCTTGATCACAAAAGAGATCAAAAGTGTCAATGGC S H N Q L T K I P P I N A H L E H L H L D H N R I K S V N G ACTCAGATATGCCAGTCCAATTGCAGTAGCAGAAGACTATGGTCTTTATGGTAACATCCCTCGCCTTCGTTATCTTCGCTTGGATGGA	 810 900 990 1080 1170 1260 1350 1440
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D TCTCATAACCAGGTCACTAAAATTCCACCGATCAATGCTCATCTCGAGCACCTTCACCTTGATCACAAACGAAATGAAAAGTGTCAATGGC S H N Q L T K I P P I N A H L E H L H L D H N R I K S V N G ACTCAGATATGCCCAGTCTCAATTGCAGTAGCAGAAGACTATGGTCTTTATGGTAACACCCTCGCCTTCGTTATCTTCGCTTGGATGGA	 810 900 990 1080 1170 1260 1350 1440
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D TCTCATAACCAGCTCACTAAAAATTCCACCGATCAATGCTCATCTCGAGGCACCTTCACCTGATCACAACAGAAATCAAAAGTGTCAATGGC S H N Q L T K I P P I N A H L E H L H L D H N R I K S V N G ACTCAGATATGCCCAGTCTCAATTGCAGTAGCAGAAGACTATGGTCTTTATGGTAACATCCCTCGCCTTCGTTATCTTCGCTTGGATGGA	 810 900 990 1080 1170 1260 1350 1440 1520
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D TCTCATAACCAGCTCACTAAAATTCCACCGATCAATGCTCATCTCGAGCACCTTCACCTTGATCACAAAAGAGAGATCAAAAGTGTCAATGGC S H N Q L T K I P P I N A H L E H L H L D H N R I K S V N G ACTCAGATATGCCCAGTCCAATGCGCGAGAAGACTATGGTCTTATGGTAACATCCCTCGCCTTCGCTTATCTTCGCTTGGATGGA	 810 900 990 1080 1170 1260 1350 1440 1520 1610
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D TCTCATAACCAGCTCACTAAAATTCCACCGATCAATGCTCATCTCGAGCACCTTCACCTTGATCACAAACAGAATCAAAAGTGTCAATGGC S H N Q L T K I P P I N A H L E H L H L D H N R I K S V N G ACTCAGATATGCCCAGTCCAATTGCGATAGCAGAAGACTATGGTCTTTATGGTAACATCCCTCGCCTTCGTTATCTTCGCTTGGATGGA	 810 900 990 1080 1170 1260 1350 1440 1520 1610 1691
T F L R L N Y N K L S D E G I P P N G F N V S S I L D L Q L D TCTCATAACCAGCTCACTAAAATTCCACCGATCATGCTCATCTCGAGCACCTTCACCTTGATCACAAACGAATCAAAAGTGTCAATGGC S H N Q L T K I P P I N A H L E H L H L D H N R I K S V N G t ACTCAGATATGCCCAGTCTCAATGCCAGTAGCAGAAGACTATGGTCTTTATGGTAACATCCCCCGCCTTCGTTATCTTCGCTTGGATGGA	 810 900 990 1080 1170 1260 1350 1440 1520 1610 1691

Fig. 1. Nucleotide and deduced amino acid sequences of quail keratocan (a) and mimecan (b), and comparison with chick cDNA. GenBank accession numbers are: keratocan, AF128223 (quail), AF022890 (chick); and mimecan, AF128224 (quail), AF126963 (chick). The nucleotides are numbered beginning at the translation start site. Bold, lowercase letters above the quail nucleotide sequence and bold, uppercase letters below the amino acid sequence provide the corresponding data from the chick sequence and mark sites of variations in nucleotides and amino acids, respectively, from the quail sequence. A hyphen denotes absence of a nucleotide or an amino acid. The 5'-end of the chick sequence is indicated by an asterisk (*). The 3'-ends of both chick keratocan and mimecan cDNAs, which include poly(A) tails, are at quail positions 2306 and 1745, respectively.

			a	τ	С			e c	j a	C	;	t			c c	;	t							
TTT	AA	TTT	TGAC	ACCA	TTTT	TCCI	FTCAG	CTTT	GGT	TGGT	CATI	ICAGI	ATTIC	GTTC	TATA	TTIC	GAT	ICTA	CAGA	GAC	AAAA	CCAA	AGA	1871
AAAC	TA	g ATC	a a	ac PTTA	TTTC2	CAGZ	AGTAG	GIGTI	TTA	CAGA	AATI	g Facaa	g c Agatz	AGTT	a AATO	a (State	t g	t IGAA	g AAAC	a ATTZ	ATT	TAAG	a GAG	1961
t ATTZ	AA -	gaa	ttta A	g - Aga	g 	ACCC	ICCCI	C CCTAC	a CCAC	TAAC	CA CGCT2	AACTI	a t	CTGA	ACTZ	g AAAC	AAA	TATT	TICA	TAG	GAAG	TGTZ	ATC	2043
AGT	8 rgg	I AGC	t AAAT	саа А	IC -'AAA(TCAC	GGGCA	a GIGT	FTAG	CTAI	CTG	tg CCAAC	TCATZ	C G TAAA	TAC	C ATATZ	ac ACT	ICIG	a GACI	CTA	gaga	TGTC	C CTA	2129
GTT	a TT	TTA	GATT	a ICAG	TAGC2	AACZ	t AAACA	ATAT	t ATAA	GCAG	GCA	ICITI	t tg 'ACACZ	C AAAT	CIGO	t FTGG2	ATT	TIGA	TTTI	TGT	IGTI	GTIC	- Gat	2219
			t	a	с	t				t			t	888	aaa	aaaa	aaa	aaaa	888	aaaa	1888	aaa	888	
GTT	этт	TGG	ATGT	rr-1	TTAT	rigit	GTTT	GTTT	AGTA	AAAC	AATZ	ACTGC	GAGT	CAAT	ACA	AAT	ICA	AATA	TTTC	AGC	CAAC	AAT	ICA	2308
CIC	FAT	TCA	CIGT	ГААТ	TTTI	JTTA'	ITICI	TAATA	CCCA	ACAA	AGC	AGACI	ATCA	TCAG	CAAT	ATT?	ATT	TTAC	TTTI	CAT	JTAT	TIGI	rag	2398
ATA	SAA	GTA	ATTA	GTTG	AATCO	CATTO	CCTA	CATT	FTAG	AAGA	GAAJ	TTAG	AATT	'AGAA	CATI	CTA	AAT.	TTTA	AAAC	TGG	CCTC	AGCA	AC	2488
AAT	'AA	TTT	AAAC	IGTA	TATG	IGGC/	ATAGI	CTTC	FICT	TIGG	AGG	TAAA	CATT	TCTT		AGCA	GAT	CTTT	AGAA	GCA	FIGA	GAAT	TT	2578
CAT	TT	CTA	ACCT	ICAG	TCIC	ATAC.	IGTTI	GAAA/	ATTG	GGAC	TTAT	IGCIC	TCIGO	CACT	TACA	\GA'I'	ATA/	ACCC	AGTA	TAG	AAGC	CCAG	FIG	2668
TAG	\AT	TAA	CTTA'	FTCA	GAAA'	ICCI	CTTT	GAAA	ACCT	TTAT	TAT		TICAT	TGAA	TATO	ATTO	GCT	GGAA	AATI	TTA	JCAA	AATO	JTT	2758
TCC.	IGI	'ACC	TTTA	TATI	CATI		A'IGA/		ATAC	ACA1	CTA/			AGTA	ALGIX	AIGC/	ATG:	TICA	CGCA	TGI.		TCA'		2848
TAA	AI	AGT	AA1-1-	TATI	-1-1-1-10	JAAT	AAGAA	AIGA	1-1-1-1	6666	ACAC	STACT	ACTA	(I-IAA	A1-17	ATAA'	IGA	TCAT	CAAP	AAA	AAAA	AAAA	AAA	2938
AAAA	MA	AAA																						2941
(b)																								
പ്പും സമനന	ייידי	TTTV-V	~ አጥጥ	יוויויוי	ልጣጥጋር	יאיינגי	പപപ	ദവംദവ	יאיזיעי	יייע	ል አጥጥ	γ m γ m γ m γ m γ m γ m γ m γ m γ m γ m γ γ m γ γ m γ γ γ γ γ γ γ γ γ γ γ γ γ	സ്ഥാന	C2C2	Con Con	സമഹ	γm·	יייאריב	አስጥ	GT. גידיג ג	CAA	JAGA	AT A	-523
GAGA	AA	AAG	CCATZ	GCA	ATTAG	AAAG	GGTA	AATCA	CCA	CACA	GCAG	AAAA	CIGCC	ACTT	TCAC	CAAG	CAG	GAT	GAAC	CAGA	CTG	GATA	AC	-455
ATTI	GT	AAA	GGTT7	AAG	CTAAG	TACA	GCCC	CCTGC	TCC	AGGA	TATI	ACCT	CTTAC	CCTA	GGGA	CAGA	ATG	CTG	GCAA	AGCA	TTT	ITTA	CC	-253
											~	•					++~	otta	ttaa	atta				
CCAT	GG	AAA	GCAC'	TTT	AAGAC	TGIC	CAATC	TTGGZ	AAC	CTTA	TTCA	TIGC	AGCAA	CCTG	GATC	TAA				any 	TIC	ATAA	CT	-178
t TCAC	TC	TTT	TATC	AGTC	g i TACAC	i g TGGI	TTAA	a .cgtai	IGCA	TAAG	t AACA	t ACTG	C GTTTI	CCA AG	a - CAC	AGCO	g CACZ	AGAA	a GAGC	t CTGC	at BACC	CIGC	TA	-91
AAAC	GT	CAC	AAAG	AAAG	TATIC	TCCF	ACAGI	CAACI	CAC	GAGC	AGAC	CTCC	ATCAG	CTCA	GTGC	AAGC	CAZ	AGGA	CATT	GĊĪ	C CTGT	gaga	AA	-1
AAAC	GT	CAC	AAAG g	AAAG a	TATIC g	TCCZ	ACAGI	CAACI	CAC	GAGC	AGAC	CTCC	ATCAG	CTCA	GTGC	CAAGO	CCAZ	AGGA	CATT	GĊŦ	C TGT	gaga	AA	-1
AAAC ATGJ M H	GT AG	CAC ACT T	AAAG g TTAC L Q	AAAG a AGGC A	G TACTI T F	TCC/	ACAGI ICCTG	CAACZ GTTGC V A	ACAC CATT F	GAGC TGTA V	AGAC CCTI P I	CTCC TGGI V	ATCAG GAAGO K F	CTCA CAGC A	GTGC 8 ACCI P	CAAGO CCTZ P 1	CCAP ATAC I (AGGA CAGC 2 Q	CATT AAGA D	GCTC TTCI S	C TGT ACCC P	gaga Aagi K F	AAA PTT	-1 90
AAAC ATG/ M P	GT	CAC ACT T	g TTACZ L Q att	AAAG AGGC A	G G TACTI T I A		ACAGI ICCIG L	CAAC# GTTG0 V A	ACAO CATT F	GAGC TGTA V	AGAC	CTCC TIGGI V	ATCAG GAAGO K I	CTCA CAGC A	GTGC 8 ACCI P	CAAGO I ICCTZ P]	CCAZ ATAC I (AGGA CAGC 2 Q a	CATT AAGA D	GCTX TTCI S	C CTGT ACCC P	gaga 'Aagi K F	AAA PTT ?	-1 90
AAAC ATGZ M P TATC	GT AAG C	CAC ACT T t	AAAG g TTAC L Q gtt G	AAAG a AGGC A g ATAC	TATTO g TACTI T I A 1 AGATI	TTCP TTTT F F (ATTG)	ACAGI ICCTG L CACG	CAACA GTTGC V A GGAAC	ACAO CATT F	GAGC TGTA V GATC	AGAC CCTI P I a CCAGC	CCTCC PTGGI V LAGGA	ATCAG GAAGC K I	CTCA CAGC A	GTGC 8 ACCI P GCTC	CAAGO P 1 P 1 GCCC2	CCAF ATAC I (AAGC	AGGA CAGC 2 Q a GATG	CATT AAGA D CAAT	GCTX TTC/ S	C TGT ACCC P BGAT	gaga Aagt K F Gga <i>p</i>	AAA TTT ?	-1 90 177
AAAC ATG/ M H TATC Y H	GT AAG C BAG	CAC ACT T t	AAAG TTACJ L Q gtt G V	AAAG AGGC A g ATAC T A	TATIC TACTI T I A I AGATZ D I	TTCC#	ACAGI ICCTG L S ICACG T	CAACA GTIGO V A GGAAC G S	CATT F SCCT L	GAGC TGTA V GATC I	AGAC CCTI P I a CAGC Q (CCTCC TTGGI U V CAGGA Q D	ATCAG GAAGG K H TTATC Y H	CTCA CAGC A AAAT M	GTGC ACCT P GCTC L	CAAGO P I GCCCZ P I	ATAC I (AAGC K I	AGGA CAGC 2 Q B GATG D A T	CATT AAGA D CAAT I	GCTX TTCI S AAAA K	C TGT ACCC P SGAT D	GAGA AAGI K F GGAA G I	AAA PTT S ACA	-1 90 177
AAAC ATG/ M P TATC Y P	GT AAG C BAG	CAC ACT T t TAC	AAAG TTAC LQ gtt G V	AAAAG AGGCC A g ATAC T A	TATIC TACTI T I A 1 A A I A A A A C	TTCC# TTTTT F F (ATTG: I V F A	ACAGI ICCTG L C ICACG T	CAACZ GTTGC V A GGAAC G S	ACAO F SCCT L	GAGC TGTA V GATC I	AGAC CCTI P I CAGC Q (CCTCC TIGGI V CAGGA 2 D a	ATCAG GAAGC K I TTATC Y I	CTCA CAGC A AAAT	GTGC ACCT P GCTC L	CAAGO P I SCCCZZ P I	ATAC I C AAGC X I	AGGA CAGC 2 Q BATG D A T	CATT AAGA D CAAT I	GCTX TTCI S AAAA K	C TGT ACCCC P SGAT D	GAGA AAGI K F GGA# G 1	AAA TTT ? ACA T	-1 90 177
AAAC ATGA M P TATC Y P AATC N V	CGT VAG C EAG STT	CAC ACT T TAC Y TAC Y	AAAG TTACJ L Q gtt G - D V CTIGJ L D	AAAAG AGGCC A g ATAC T A ACAC T	TATIC G TACTI T H A 1 A 1 C TGGCC G I A	TTCCF TTTTT F MTTGI V TGAC L R	ACAGI ICCTG L ICACG T GACTG L	CAACZ GTIGO V A GGAAC G S KCAAGO Q A	ACAO F SCCT L CAGA D	GAGC TGTA V GATC I TGAC D	AGAC PI B CAGC Q S I	CCTCC TIGGI J V CAGGA Q D BAGCI E L	ATCAG GAAGC K F TTATC Y F GAGTC S J	CTCA CAGC A AAAAT M CCCAG	GTGC ACCT P GCTC L ACCC P	CAAGO P J SCCC2 P J SCCC2 P J CACAJ T J	ATAC I C AAGC K I AAGC K I	AGGA CAGC 2 Q BATG D A T GATA D T	CATT AAGA D CAAT I CAAA N	GCTX S PAAAC K TTTX L	C TGT ACCC P GGAT D SCCT P	GAGA AAGI K F GGA G I ACTI T (AAA FTT F ACA F	-1 90 177 267
AAAC ATGZ M H TATC Y H AATC N V	CGT AAG C SAG STT 7	CAC ACT T TAC Y TCT S	AAAG TTACJ L Q gtt GJ - D V CTTGJ L D	AAAG AGGC A g ATAC T A ACAC T	TATTO TACTI T F A 1 PAGATZ D 1 C TGGCC G 1 A	TTCC FTTTT F TTC TTC ATTC A TTC A CTCAAC	ACAGI L CACG T GACTO L	CAACZ GTIGO V A GGAAC G S CAAGO Q A	ACAO CATT F SCCT L CAGA D	GAGC TGTA V GATC I TGAC D	AGAC PI a CAGC Q (AGCC S I	CCTCC TIGGI L V CAGGA 2 D BAGCI 5 L	ATCAG GAAGC K H TTATC Y H GAGTC S J	CTCA CAGC A A AAAAT C M CCAG	GTGC ACCI P GCTC L ACCC	CAAGO P 1 CCTX P 1 CCCTX P 1 CCCTX P 1 CCCTX	ATAC ATAC I (AAGC K I AAGC K I	AGGA CAGC. 2 Q BATG D A T GATA D T	CATT AAGA D CAAT I CAAA N	GCTX S AAAA K TTTX L	C TGT ACCC P SGAT D SCCT P	GAGA AAGI K F GGAA G I ACTI T C	AAA TTT ACA T IGT	-1 90 177 267
AAAC ATGZ M P TATC Y P AATC N V CTGC L I	CGT AAG C SAG STT V TC'	CAC ACT T TAC Y TCT S TCC C	AAAG TTACJ L Q gtt Gj Gj V V CTIGJ L D GTGTC V C	AAAG AGGC A AGGC T ATAC T A ACAC T L	TATIC TACTI T I A 1 C C TGGCC G I A AAGCC S C	TTCC TTTTT F T T T T T T T T T T T T T T	ACAGI L CACG T SACTO L CAGIG V	CAACA CAACA V A CGGAAC G S CAAGC Q A TACTC Y C	ACAO F SCCT L CAGA D STGA0 E	GAGC TGTA V GATC I TGAC D GGAA E	AGAC PI 8 CAGC Q (AGC SI ATAC I I	CCTCC TTGGT V CAGGA 2 D BAGCT 5 L SATAT) I	ATCAG GAAGG K I TTATG Y I GAGTC S J TGAAG E A	CTCA CAGC A AAAAT M CCAG A R CTGT	GTGC ACCI P GCTC L ACCC P ACCC	CAAGO P I CCT# P I CCCAC P I	ATAC ATAC I C AAGC K I AAGC K I CTGC	AGGA CAGC. 2 Q BATG D A GATA D T GATA D T	CATT AAGA D CAAT I CAAA N AGGA E	GCTX S AAAA K TTTX L TTTX L AAAC2 T	C TGT P GGAT D GCCT P AGCA A	GAGA AAGI K F GGAA G J ACTJ T C TATC Y L	AAA PTT ACA F NGT C TTT	-1 90 177 267 357
AAAC ATGA M H TATC Y H AATC N V CTGC L I	CGT AAG C SAG STT J	CAC ACT T TAC Y TCT S TCC	AAAG TTAC2 L Q gtt G V CTIG L D STGTC V C	AAAAG AGGCC A ATAC T A ACACC T T L	TATIC TACTI TACTI A 1 A A AGATZ D I C C C G I A A A G C C C C C C C C C C C C C C C C	TTCC TTTTT F C TTCC C TTCC C TTCC C TTCC C TTCC C TTCC C TTTTTT	ACAGI L CACG T GACTO L CAGIG V	CAACZ GTTGC V A CGGAAC G S CAAGC Q A TACTC Y C	ACAC F SCCT L CAGA D STGA(E	GAGC TGTA V GATC I TGAC D GGAA E	AGAC CCTTI P I CAGC Q (CAGCC S I ATAG I I	CCTCC TIGGI V CAGGA 2 D 8 3 AGCI 5 L 3 ATAT) I	ATCAG GAAGG K I TTATC Y I GAGTC S J TGAAG E A	CTCA CAGC A A A A A A A A CTGT V	GTGC 8 ACCT P GCTC L ACCC P	CAAGO P I GCCCIP P I CACAM T I CACAM T I	ATAC ATAC I (AAGC X I AAGC X I AAGC X I E TGC	AGGA CAGCA 2 Q a SATG SATG D A T GATA D T CCAAA	CATT AAGA D CAAT I CAAA N CAAAA N AGGA E	GCTX TTC2 S AAAA K TTTX L AAC2 T	C TGT ACCCC P GGGAT D GGCCT P AGCA A	GAGA AAGT K F GGAA G 1 ACT T T C C T ACT Y L	AAA PTT P ACA F IGT C TTT	-1 90 177 267 357
AAAC ATGZ M H TATC Y H AATC N V CTGC L I	CGT AAG C SAG STT Z TC C CA	CAC ACT T TAC Y TCT S TCC C C S	AAAG g TTACJ L Q gtt G V CTIG L D GTGTC V C TTTA	AAAAG AAGGC A T AATAC T A ACACC T L	TATIC TACTI T I A 1 CAGATZ D 1 C TGGCC G 1 A AAGCC S C AAACCA	TTCC# TTTTT F C VTTG: C V C TGA C C TGA C C TGA C S S AGAC	ACAGI L C C C C C C C C C C C C C C C C C C	CAACZ GTIGO V A GGAAO G S CAAGO Q A TACTO Y C	ACAC F GCCT L CAGA D STGA(E	GAGC TGTA V GATC I TGAC D GGAA E	AGAC CCTTI P I CAGC Q (S I ATAG I I TTTC	CCTCC TIGGI V CAGGA 2 D BAGCI 5 L SATAT) I	ATCAG K I TTATC Y I GAGTC S J TGAAG E A C	CTCA CAGC A A AAAAT M CCCAG A R CTGT V CTGT	GTGC P GCTC L ACCC P ACCC P	CAAGO CCCTX P 1 CCCTX P 1 CCCCA T 1 CCCCAC P 1 CCCCAC	ATAC ATAC I C AAGC K I AAGC K I CTGC C F	AGGA CAGCA 2 Q BATG D A T GATA D T CCAAA 2 K t t	CATT AAGA D CAAT I CAAA N CAAAA N AGGA E	GCTX S AAAA K TTTT L AACZ T	C TGT P GGAT D GGCT P AGCA A	GAGA AAGI K F GGAA G I T T C C T T T C C T A T T A A T A	AAA TTT 7 ACA F NGT C TTT 4	-1 90 177 267 357
AAAA ATGY M H TATC Y H AATC N V CTGC L I L TATC Y A	CGT AAG C SAG STT V C CA	CAC ACT T TAC Y TCT S TCCC C C C C C C C C C C C R	g TTACI L Q gtt Gi V CTTGI L D STGTC V C TTTAA	AAAAG AGGGC A G AATAC T A ACAC T T L ATAA K	TATIC TACTI TACTI TACATI D TAGATIZ D TGGCC G I A A AAGCC S C AAACCZ I F	TTCCA TTTTTT F F C TTCAA TTCAA C TTCAA C TTCAAA TTCAAA TTCAAA TTCAAAA TTCAAAAA TTCAAAAAAAAAA	ACAGI L CCTG L CACG T CACGG L CAGTG V CAGTG V CAGTG V	CAACE V A GGAAC G S CAAGC Q A TACTO Y C GCTO71 A V	ACAO CATT F GCCT L CAGA D STGAA E STGAA S CTCC S	GAGC TGTA V GATC I TGAC D GGAA E AGAT D	AGAC CCTT P I CCAGC Q (CAGCC S I ATAG I I TTTC F A	CCTCC FIGGI J V CAGGA Q D BAGCI E L BAGCI E L SATAT J I SCTGA Q D	ATCAC GAAGC K I TTTATC Y I GAGTC S J TGAAG E A C CATR I I	CTCA CAGC A A AAAAT CTGT V CTGT T	GTGC & ACCT P GCTC L ACCCC P ACCCC P TTTC L	CAAGO P I CCCTM P I CACAM T I CCCAC P I CCCAC P I R F	ATAC ATAC I C AAGC K I AAGC K I CTGC C F AGAZ R I	AGGA CAGCA 2 Q BATG D A T GATA CCAAA P K CCAAA P K t ATCCG I D	CATT AAGA D CAAAT I CAAAA N ACGA E ATTTT F	GCTX S AAAAX K TTTTX L AAACZ T TAG7 S	C CTGT ACCCC P SGAT D SGCCI P AGCA A CGGG G	GAGA AAGI K F GGAA G I ACTI T C C TATC Y L AATA N M	AAA PTT 7 ACA F TGT C TTT	-1 90 177 267 357 447
AAAA ATGA M H TATC Y H AATC N V CTGCC L L L TATC Y A	CGT AAG C SAG STT T C C C C C C A C	CAC ACT T TAC Y TCT S TCC C C C C C C C C C C C C C C C	AAAG TTAC: L Q gtt G V CTTG: L D CTTG: V C TTTA: V C TTTA: V C	AAAAG AAGGC A MATAC T A ACAC T L STTTT. L ATAA K	g TATTO TACTN T A 1 AAGATO C G TGGCCC S C S C AAAGCCC A AAAGCC S C I F	TTCC/ FTTTT F C MTTG2 C V C TGAC C R C C GATC S S AGATC C R	ACAGI L CCTC L CACC T CACC L CACTC L CACTC L CACTC L CACTC L CACTC L L CACTC T L CACTC T T CACC T T CACTC L C C C C C C C C C C C C C C C C C	CAACE V A GGAAC G S CAAGC Q A TACTO Y C GCTG7 A V	ACAC F GCCT L CAGA D STGAA E CTCA S	GAGC TGTA V GATC I TGAC D GGAA E AGAT D	AGAC CCTTI P I CCAGC Q (CAGCC Q (CAGCC S I ATAG I I TTTCC F A	CCTCC FIGGI V CAGGA 2 D 8 8 AGCI 5 L 1 SATAT 0 I SCTGA 2 D	ATCAC GAAGC K I TTIATC Y I GAGTC S J TGAAG E A C CATTA I I	CTCA CAGC A A AAAAT ; M CCCAG A R CTGT V CTAC	GTGC a ACCI P GCTC L ACCCC P TTTC L	CAAGO P I SCCCI P I CACAA T I CCCAC P I CCCAC P I SAGAA R F	ATAC ATAC I C AAGC K I AAGC K I TTCC I E ACGAZ R I I	AGGA CAGCA 2 Q BATG D A T GATA D T CCAAA C CCAAA C CCAAA C CCAAA C CCAAA C CCAAA C CCAAA C CCAAA C CCAAA C	CATT AAGA D CAAT I CAAA N AGGA E ATTT F	GCTX S AAAAX K TTTTX L AAACZ T S AAACZ	C CTGT ACCCC P GGAT D GGCCT P AGCA A CGGG G C t	GAGA AAGTI K F GGAA G I ACTI T C C TATC Y L AATA N M	AAA TTT 7 ACA T TGT 2 TTT 4	-1 90 177 267 357 447
AAAA ATGJ M P TATC Y P AAATC Y P TATC Y P ATAC	CGT AAG SAG STT ZTC' CTC' SCA4	CAC ACT T T TAC Y TCT S TCC C C C R I C C C C R I C C C C C C C C	AAAG TTAC: L Q gtt G V V CTTG: L D STGTC V C TTTA: F N ATAGZ I E	AAAAG AGGCC A ATAC T A ACAC T L STTT: L K AAGAA K AAGAA	TATTC G TACTJ T A A A A A A A A A A A A C C C A A A A A A C C A A A A A C C A A A A A A A A A A A A A	TTCCF	ACAGI L CACCIC T GACTC L CACGIC V SGATA I S	CAACE GTICK V A GGAAC G S CAAGC Q A TACTC Y C GCTG7 A V AAGC7 K L	ACAC CATT F GCCT L CAGA D STGA E STGA S TCTC S	GAGC TGTA V GATC I TGAC D GGAA E AGAT D GTTA L	AGAC CCTTI P I CCAGC Q (CAGCO S I ATAG I I TTTCC F A TTTCC L E	CCTCC FTGGI V CAGGA 2 D 8 6 6 6 6 6 7 1 1 1 1 1 1 1 1 1 1 1 1 1	ATCAC GAAGC K F TTATC Y F GAGTC S J TGAAG E A CATTA I I I I S	CTCA CAGC A A A A A A A A CCCAG A R CTGT V CTAC C T C CTCT L	GTGC ACCI P GCTC L ACCCC P ACCCC P TTTC L	CARACA P I GCCCI P I CACAM T I CCCAC P I CACAM R F CCAAP R F	ATAC ATAC AAAGC X I AAAGC X I AAAGC X I AGAA X I AGAA X I X I X I X I X I X I X I X I X I X I	AGGA CAGCA 2 Q BATG D A T GATA D T CCAAM 2 K t ATCGA I D g Q Q Q Q Q Z L	CATT AAGA D CAAT I CAAA N AGGA E ATTT F F	GCTX S AAAA K TTTX L AAACZ T TAG7 S GAAZ K	C TGT ACCCC P GGAT D GGCT P AGCA A TGGGG G C t t	GAGA AAGTI K F GGAA G I T C TATCY Y L AATA N M CCAG P U	AAA FTT ACA F GT C STT KTG STT 7	-1 90 177 267 357 447 537
AAAA ATG3 M P TATC Y P AATC Y P TATC Y P ATAC I F	CGT AAG SAG STT ZTC ZTC X X X X X X X X X X X X X X X X X X X	CAC ACT T TAC Y TCT S TCC C S CCG R I CCG R I CCG C R I CCG C R I CCG C C C	AAAG TTAC: L Q gtt G V V CTTG: L D STGTC V C TTTA: F N ATAG: I E	AAAAG AGGCC A ATAC T A ACACAC T L ATAA K AAAGA D	g TATTO TACTI T A 1 A C TTGGCCC G A A A C T C A B A B A A B A B <td>TTCCF TTTTT F TTTTT TTTTT TTTTTT TTTTTTTT</td> <td>ACAGI L CACGO T GACTO L GACTO V SGATA I S</td> <td>CAACI GTIGO VA GGAAC G S CAAGO V C GCAGO Y C GCTG7 A V AAGC7 K L</td> <td>ACAC F F CAGA D STGAG E CAGA S S TCTCI L</td> <td>GAGC TGTA V GATC I TGAC D GGAA E AGAT D GTTA L</td> <td>AGAC CCTTI P I CCAGC Q (CAGCC S I ATAG I I TTTGG F A TTTGG L E</td> <td>CCTCC PTGGI V CAGGA 2 D 8 GAGCA 5 L 9 I 9 I 9 I 9 CCTGA 4 D 1 9 CCTGA 6 CTGA 6 C 1 9 CCTCC 1 9 CAGGA 1 9 CAGGA 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1</td> <td>ATCAC GAAGCC K F TTTATC Y F GAGTC S J TGAAG E A CATTA I I I S</td> <td>CTCA CAGC A A A A A A A A CTGTA C T CTCT L</td> <td>GTGC ACCT P GCTC L ACCC P TTTC L TTTC L TTTC L</td> <td>CACAA P I GCCCIA P I CACAA T I CACAA R F CACAAA R F CGAAAA E N</td> <td>ATAC ATAC I (AAGC K I AAGC K I CTGC CTGC CTGC Z AGAA R I AGAA R I AGAA R I AGAA R I I I I I I I I I I I I I I I I I I I</td> <td>AGGA CAGC. 2 Q BATG D A T GATA D T CCAAM ? K t ATCG I D g CGACT R L</td> <td>CATT AAGA D CAAT I CAAA N AGGA E ATTT F F TIGT V</td> <td>GCTX S AAAAX K TTTX L AAACZ T TAACZ S GAAZ K</td> <td>C TGT ACCCC P GGAT D GGCT P AGCA A CTG L</td> <td>GAGA AAGTI K F GGAA G 1 TACTI T (C C TATC Y L AATA N M CCCAG</td> <td>AAA FTT F ACA F GT C STT ATG 1 STT 7</td> <td>-1 90 177 267 357 447 537</td>	TTCCF TTTTT F TTTTT TTTTT TTTTTT TTTTTTTT	ACAGI L CACGO T GACTO L GACTO V SGATA I S	CAACI GTIGO VA GGAAC G S CAAGO V C GCAGO Y C GCTG7 A V AAGC7 K L	ACAC F F CAGA D STGAG E CAGA S S TCTCI L	GAGC TGTA V GATC I TGAC D GGAA E AGAT D GTTA L	AGAC CCTTI P I CCAGC Q (CAGCC S I ATAG I I TTTGG F A TTTGG L E	CCTCC PTGGI V CAGGA 2 D 8 GAGCA 5 L 9 I 9 I 9 I 9 CCTGA 4 D 1 9 CCTGA 6 CTGA 6 C 1 9 CCTCC 1 9 CAGGA 1 9 CAGGA 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	ATCAC GAAGCC K F TTTATC Y F GAGTC S J TGAAG E A CATTA I I I S	CTCA CAGC A A A A A A A A CTGTA C T CTCT L	GTGC ACCT P GCTC L ACCC P TTTC L TTTC L TTTC L	CACAA P I GCCCIA P I CACAA T I CACAA R F CACAAA R F CGAAAA E N	ATAC ATAC I (AAGC K I AAGC K I CTGC CTGC CTGC Z AGAA R I AGAA R I AGAA R I AGAA R I I I I I I I I I I I I I I I I I I I	AGGA CAGC. 2 Q BATG D A T GATA D T CCAAM ? K t ATCG I D g CGACT R L	CATT AAGA D CAAT I CAAA N AGGA E ATTT F F TIGT V	GCTX S AAAAX K TTTX L AAACZ T TAACZ S GAAZ K	C TGT ACCCC P GGAT D GGCT P AGCA A CTG L	GAGA AAGTI K F GGAA G 1 TACTI T (C C TATC Y L AATA N M CCCAG	AAA FTT F ACA F GT C STT ATG 1 STT 7	-1 90 177 267 357 447 537
AAAA ATGJ M H TATC Y H AATC L I TATC Y A ATAC I H	CGT AAG C SAG STT C C C C C C C C C C C C C C C C C C	CAC ACT T TAC Y TCT S TCC C S C C C C C C C C C C C C C	AAAG TTTAC: L Q gtt - G V V CTTG: L D STGTC V C TTTAM F N ATAGMIE I E	AAAAG AGGCCA A G G ATACC T A A ACACC T L STTTTL L ATAAA K D	TATTC G TACTJ T A A A A A A A A A A A A C C C A A A A A A A C C A A A A A A A A A A A A A	TTCC2 TTTTT F F (V TTGAC CTGAC	ACAGI L CACGO T GACTO L GACTO V SGATA I L C C C	CAACE GTTGO V A GGAAG G S CAAGO Q A TACTO Y C GCTG7 A V AAGCTI K L	ACAC CATT F GCCT L CAGA D CAGA CAGA CAGA CAGA CATT L G G CATT F CATT F CAGA CATT F CATT F CATT F CATT F CATT F CATT F CATT F CATT F CATT F CAGA CATT CAGA CAG	GAGC TGTA V GATC I TGAC D GGAA E AGAT D GTTA L	AGAC CCTTI P I CCAGC Q (AGGCC S I ATAG I I TTTGG F A TTTGG L F	CCTCC PTGGI L V CAGGA Q D BAGCI S L BAGCI S L SATAT D I SATAT D I SATAAT D I SAGAA S D S CTCA	ATCAC K I TTATC Y I GAGTO S J TGAAG E A C CATTA I I L S a	CTCA CAGC A A A A A A A A CTGTA C CTCT T L	GTGC ACCTI P GCTCC L ACCCC P TTTCC L TTTCC L A COLORING	CACAA P I CCCTX P I CCCTX P I CCCAC P I CCCAC P I CCCAC P I CCCAC P I CCCAC P I CCCAC P I CCCAC P I CCCTX	ATAC ATAC I (AAAGC K I AAAGC K I AGAA R I AGAA R I AGAA R I AGAA C I C	AGGA CAGC 2 Q BATG D A T GATA D T CCAAA C CCAAA 2 K 1 D Q CGAC 2 L	CATT AAGA D CAAT I CAAA N AGGA E ATTT F TTGT V	GCTX TTC: S AAAAX K TTTT TAG7 S 8 GAAA K	C TGT ACCCC P GGAT D GCCT P AGCAA CTG G L	GAGA AAGI K F GGAA G I TACTI T C C TATC Y L AATA N M CCCAG P V	AAA PTT 7 ACA F CACA F CTT C TTT ATG 1 STT 7	-1 90 177 267 357 447 537
AAAAA ATGZ M H TATC Y I AAAA CTGC L I TATC Y A ATAC I E CTAC L T	CGT AAG GAG STT STT CA A SCA CA CA CA CA CA CA CA CA CA CA CA CA C	CAC ACT T TAC Y TCT S TCC C S TCC C C S C C C S C C C T C C C C	AAAG TTAC: L Q gtt G V V CTTG: L D STGTC V C TTTA: F N ATAGZ I E AAAAC	AAAAG AGGCCA A G G ATACC T A ACACC T C A ACACC T C A ACACC T C A C C C T T C C C T T C T T T C C C C	TATTC G TACTI T I A A A A A A A A A A A A A	TTCCA TTTTT F F C CTGAC C C C C C C C C C C C C C C C C C C C	ACAGI L CACGO T GACTO L GACTO V GGATA S C C ATTCCG A	CAACI GTICK V A GGAAC G S CAACA Q A TACTO Y C GCTG7 A V AACCI K L AACCI	ACACO CATT F GCCT L CAGA D STGAA E S S CTCTCL S S S S S S S S S S S S S S S S S S S	GAGC TGTA V GATC I TGAC D GGAA E AGAT D GTTA L CAGA R	AGAC CCTTI P I CCAGC Q (ACAGCO S I ATAGCI I I TTTGG F A TTTGG L F ATCAA	CCTCC PTGGI J V CAGGA 2 D 8 6 6 6 6 1 1 1 1 1 1 1 1 1 1 1 1 1	ATCAC K I TTATC Y I GAGTO S J TGAAG E A CATTA L S a CAGGG CAGGG CAGGG	CTCA CAGC A A A A A A A A A A A CTGT T CTCT T L CTCT T L CTCT T	GTGC ACCTO P GCTCO L ACCCO P TTTCO L TGCTI A CAAAA K	CAAGA P I GCCCI P I CACAM T I CACAM R F I GAGAM R F I GAGAM R F I GAGAM	ATAC AAAGC K I AAAGC K I TGC F AGAA R I AAACC V F C C	AGGA CAGC. 2 Q BATG D A T GATA D T CCAAA C CCAAA C Q Q Q Q CGAC C Q CGAC C Q CGAC C C	CATT AAGA D CAAT I CAAA N AGGA E ATTT F F ITGT V V ICAA	GCIX S AAAAX K TTTTV L AACZ T S GAAZ K GAAZ K GAAZ	C TGT ACCCC P GGAT D GCCT P AGCA A CTG G L	GAGA AAGTI K F GGAZ G T ACTI T C C C C C C C C C C C C C C C C C C C	AAA FTT 7 ACA F GT C CTT C CTT C MTG 1 STT 7	-1 90 177 267 357 447 537 627
AAAAA ATGM M H TATX Y H AATX CTGC L I TATC Y M ATAG I F CTAC L F	CGT AAG C SAG STT Z CA A CA A CA CA CA CA CA CA CA CA CA CA	CAC ACT T TAC Y TCT S TCC C C S C C C C C C C C C C C C	AAAG TTAC: L Q gtt G y V CTTG: L D STGTC V C TTTA: F N ATAGZ I E AAAAC	AAAAG AAGGCA A AAGGCA T AATAC T AACAC T STTT: L AAAGA K AAAGA D TAACC T	TATTC G TACTJ T A A A A A A A A A A A A C C C A A A A A A A C C A A A A A A A A A A A A A	TTCCA TTTTT F F C C C TTGAC C C C C C C C C C C C C C C C C C C	ACAGI L CACGO T CACGO T CACGO C CACGO C C C C C C C C C C C C C C C C C C	CAACE GTICK V A GGAAC G S CAACC Q A TACTO Y C GCTG7 A V AACCI K L AACCI N Q	ACACO CATT F GCCT L CAGA D STGAC E S TCTC: S S TCTC: S S S TTTTC L Q AAAM	GAGC TGTA V GATC I TGAC D GGAA E AGAT D GTTA L CAGA R	AGAC CCTTI P I CCAGC Q (AGGCC S I ATAG I I I TTTGG F A TTTGG L F ATCA I F	CCTCC PTGGI J V CAGGA 2 D 8 BAGCI 5 L 9 I 8 SATAT D I 8 CTGA 2 E 8 AGAGA 3 S	ATCAC K I TTATC Y I GAGTO S J TGAAG E A CATTA L S a CAGGG R G	CTCA CAGC A A A A A A A A A A A CTGTA C CTCT T CTCT L GAATI A I	GTGC ACCTO P GCTIC L ACCCO P TTTIC L TTGCTI A CAAAA K	CAAGO P I COCTX P I COCTX	ATAC AAAGC K I AAAGC K I AAAGC J F AGAAA AACC J F C C AATC J A	AGGA CAGCA Q Q BATG D A T GATA D T CCAAA CCAAA CCAAA CGACT Q Q CGACT Q CGACT Q F	CATT D CAAT I CAAA N AGGA E ATTT F F ITGT V V CAAA K	GCTK TTCI S AAAAK K TTTTK T AAACI T S AAAA K GAAA K GAAA K	C TGT ACCCC P GGAT D GCCT P AGCAA CTG L	GAGA AAGTI K F GGAZ G T ACTI T C C C C C C C C C C C C C C C C C C C	AAA FTT C C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C	-1 90 177 267 357 447 537 627
AAAA ATGJ M H TATX Y H AATX CTGC L I TATC Y P ATAC I F CTAC L F	CGT	CAC ACT T TAC Y TCT S TCC C C C C C C C C C C C C C C C	AAAG TTAC: L Q gtt G y V CTTG: L D STGTC V C TTTA: F N ATAGZ I E AAAAC	AAAAG AAGGCA A G AATACC T A AACAC T L AATAA K AAAGA D TAACC T	TATTC G TACTJ T A A A A A A A A A A A A C C C A A A A A A A C C A A A A A A A A A A A A A	TTCCA TTTTT F F C C TTGAC C C C C C C C C C C C C C C C C C C	ACAGI ICCICG L C CAGIG C C A C C A C C A C C A C C A C C A C C C C C C C C C C C C C	CAACE CAACE V A CGAAC G S CAACE Q A TACTO Y C GCTG77 A V AAGC71 K L AAGC71 K L AAGC72 N Q	ACACO CATT F GCCT L CAGA D STGAA E CTCTC: S S TTTTI L g AAAN	GAGC TGTA V GATC I TGAC D GGAA E AGAT D GTTA L CAGA R	AGAC CCTTI P I CCAGC Q (ACAGCO S I ATAG I I I TTTGG F A TTTGG L F ATCA I F I	CCTCC PTGGI L V CAGGA Q D BAGCI S L BAGCI S L SATAT D I SATAT D I SATAT D I SAGAGA C S	ATCAC K I TTATC Y I GAGTO S J TGAAG E A CATTA I I ACTTI L S a CAGGG R G	CTCA CAGC A A A AAAAT CTGTA CTGTA CTCT L GAAT I	GTGC ACCTO P GCTIC L ACCCO P TTTIC L TGCTI A CAAAA K	CAAGO P I COCTX P I COCTX	ATAC ATAC I (AAGC K I AAAGC K I AAAGC K I AAAGC I F C C AAATC I Z	AGGA CAGCA Q Q BATG D A T GATA D T CCAAA CCAAA CCAAA CGACCA R CGACT R CGACT R CGACT R CGACTA R CGACTA R CGACTA R CGACTA R CGACCA R C CAGCA C Q Q C C C Q C Q C C C C Q C C C C C	CATT D CAAT I CAAA N CAAA N CAAA E ATTT F F V TCAA K	GCTK TTCI S AAAAK K TTTTK T AAACI T S AAAAI K GAAAI K	C TGT ACCCC P GGAT D GCCT P AGCAA CTG L	GAGA AAGT K F GGAZ G T ACTT T C C C C C C C C C C C C C C C C C	AAA PTT 7 ACA F C TTT C TTT 7 ATG 1 STT 7 AAT	-1 90 177 267 357 447 537 627
AAAAA ATGZ M H TATX Y H AATX CTGC L I TATC Y A ATAC I F CTAC L F		CAC T T TAC TAC Y TCT S TCC C S S C S C S C S C S C S C S	AAAG TTAC: TTAC: L Q gtt G V V CTTG: L D STGTC V C TTTA: F N ATAG? I E AAAACI K L CTCT?	AAAAG AAGGCA A AAGGCA T AATAC T AACAC T AACAC T AAAAGA K AAAGA D TAACCTT	g TATTY g TACTY T A 1 AAGATY D TGGCCC G TGGCCC G AAAGCC G AAACCA T F AAACAT T F AAACAA T F AAGCGAC	TTCCA TTTTT F F C C TTGAC TTGAC TTGAC S S C CTGAC S S C C TTGAC S S C C TTTAA F N S C C TTTAA F N	ACAGI ICCICG L C C C C C C C C C C C C C	CAACE CAACE V A CGGAAC G S CCAAGC Q A CTCGA CTCGA	ACACO CATT F GCCT L CAGA D STGAA E STGAA S	GAGC TGTA V GATC I TGAC D GGAA E AGAT D GTTA R CAGA R	AGAC CCTT P I CCAGC Q (CAGCC S I CAGCC S I I I I I I I I I I I I I I	CCTCC PTGGI J V CAGGA Q D BAGCI S L BAGACI S CTGA C S CTTAA	ATCAC K I TTATC Y I GAGTO S J TGAAG E A CATTA I I ACTTI L S a CAGGG R G CCCTAC	CTCA CAGC A A A AAAAT CAAAAT CTGT CTCT CTC	GTGC ACCTO P GCTIO L ACCCO P TTTIO L TGCTI A CAAAA K AAGTI	CAAGO P I COCTX P I COCTX P I CACAA T I CACAA P I CACAA R F N CACAA N N	ATAC ATAC I (AAGC K I AAAC K I AAAC C I I AAAC I I C C C C AAAC I I I I I I I	AGGA CAGCA Q Q BATG D A T GATA D T CCAAA CCAAA CCAAA CCAAA CGACT Q GGACT Q GGACT CGACT CGACT	CATT AAGA D CAAT I CAAA N AGGA E ATTT F F TTGT V ICAA K IGCA	GCTK TTCI S AAAAK K TTTT TAGI S GAAA K GAAA K CCTT	C TGT ACCCC P GGAT D GGCT G G C C ACTG L	GAGA AAGTI K F GGAZ G T T C C C C C C C C C C C C C C C C C C	AAA FTT ACA T IGT C TTT ATG I STT V AAT	-1 90 177 267 357 447 537 627 717
AAAAA ATGM M H TATX Y H AATX CTGC L I TATC Y A ATAC I F CTAC L F TTGCC L A		CAC T T TAC Y TCT S TCC C C S S CCC R TCC C C S S CCC T C C C T C C T C C T C T	AAAG TTAC: L Q gtt G V V CTTG: L D GTGTC V C TTTA: F N ATAGZ I E AAAAC I E AAAAC L Y	AAAAG AAGGCA A G ATACCT T A AACAC T L ATTAA K AAAGA D T AACTT. L	TATTC G TACTI T A A A A A A A A A A A A A	TTCCA TTTTT F F C C TTGAC C C C C C C C C C C C C C C C C C C	ACAGI ICCICG L C C C C C C C C C C C C C	CAACE V A GGAAC G S CAAGC Q A TACTO Y C GCTGT A V AAGCTI K L AAGCTI K L AAGCTI K L AAGCTI K L AAGCTI K L AAGCTI K L AAGCTI K L CTGGA	ACACO EATT F GCCTT L CAGA D STGAA E S TTTA L S S S AAAA N AATC S	GAGC TGTA V GATC I TGAC D GGAA E AGAT D GTTA L CAGA R R TGTT V	AGAC CCTTI P I CCAGC Q (CAGCC S I CAGCC S I I I I I I I I I I I I I I I I I I I	CCTCC PTGGI J V CAGGA GAGGA GAGGA GAGAGA C E CTTAA	ATCAC K I TTATC Y I GAGTO S J TGAAG E A CATTA L S a CAGGG R G CCTAC L F	CTCA CAGC A A AAAAT CAGAAT CTGT CTCT L CTCT L CAGAAT CAGAA	GTGC ACCTO P GCTIO L ACCCO P TTTIO L TTTIO L TGCII A K AAGGI S	CAAGO P I GCCCZ P I CACAM T I CACAM T I CACAM R F I GAGAM R F N N N N N N N N N	ATAC ATAC I (AAGC K I AAGC K I AAGC K I AGAA R I AAC C I C C C I C C C I I AGC I I I I I I I I I I I I I I I I I I I	AGGA CAGCA Q Q BATG D A T GATA D T CCAAA P K CGACA C Q CGACA Q CGACT Q CGACT C CGACT C CGACT C CGACT C CGACA C L L	CATT D CAAT I CAAA N CAAA N CAAA N CAAA E ATTT F F V ITCAA K IGCA H	GCTK TTCI S AAAAK K TTTTL TAGI S GAAA K GAAA K CCTTI L	C TGT ACCCC P GGAT D GCCT P AGCAA G C ACTG L CAGG Q L	GAGA AAGTI K F GGAZ G T ACTI T C C C C C C AATA N M CCAGG P V ACAAA T N CACAA	AAA FTT C TGT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C T T T C T T T C T T T C T T T C T T C T T C T C T T C T T C T T C T C T T C T T C T T C T T C T T C T T C T T T C T T C T T T C T T T C T T T C T T T C T T T T T C T	-1 90 177 267 357 447 537 627 717
AAAAC ATGZ M H TATC Y H AATC L I TATC Y A ATAC I H CTAC L A CTAC L A C		CAC T T TAC Y TCT S TCC C S S TCC C S S C S S C S S C S S T C C T S T C C T S T C C T T T T	AAAG TTAC: L Q gtt G V V CTTG: L D GTGTC V C TTTA: F N ATAG I E AAAAC I E AAAAC I Y	AAAAG AAGGCA A AAGGCA T AATAC T AACAC T AACAC T AAAGA D TAAGA T AACTT. L AACTT. L AACTT. L	TATTC G TACTI T A A A A A A A A A A A A A	TTCCA TTTTT F F C C TTGAC C F C C C C C C C C C C C C C C C C C	ACAGI ICCICG L C C C C C C C C C C C C C	CAACE V A GGAAC G S CAAGC Q A TACTO Y C GCTGT K L AAGCTI K L AAGCTI K L AAGCTI K L AAGCTI K L AAGCTI K L AAGCTI K L CTGGZ	ACACO CATT F GCCT L CAGA D STGA S TTGA L S S TTTT L g AAAN N AATC S	GAGC TGTA V GATC I TGAC D GGAA E AGAT D GTTA R CAGA R TGTT V t	AGAC CCTTI P I CCAGC Q (CAGCC S I I I I I I I I I I I I I I I I I I I	CCTCC TIGGI V CAGGA BAGCI SACCI SATAT D I SATAT D I SATAT D I SATAT D I SACGA C S C C C C C C C C C C C C C	ATCAC GAAGC K I TTATC Y I GAGTC S J TGAAG E A C CATTA I I L S a CAGGG R G CCTAC L F	CTCA CAGC A A AAAAT CAGAAAT CTGTAC CTGT CTCT L CAGAAT CAGAAT CAGAAAT	GTGC a ACCTO P GCTIO L ACCCO P TTTIO L TTTIO L TGCTI A CAAAA K AAGTI S a a AAGTI A AA A AAGTI A A AAGTI A AAGTI A AAGTI A AA A AAGTI A A AAGTI A AAGTI A AAGTI A AAGTI A AAGTI A AAGTI A AAGTI A AAGTI A AAGTI A AAGTI A AAGTI A AAGTI A AAGTI A A A A A A A A A A A A A	CAAGO P I CCCTX P I CACAM T I CACAM T I CACAM R F N CCCAC P I CACAM R F N N N N N N N N N	ATAC ATAC I (AAGC K I AAGC K I AAGC K I AGAA R I AACC V F C C C C C C C I I I C C C C I I I I I	AGGA CAGC. 2 Q BATG D T T GATA D T CCAAN C	CATT D CAAT I CAAA N CAAA N CAAA N CAAA E ATTT F F V V CAAA K IGCA H	GCTK TTCI S AAAAK K TTTTK TAGI S AAAZ K GAAZ K GAAZ K CCTTI L	C TGT ACCCC P GGAT D GGCT P AGCA G C C ACTG L C ACTG Q D	GAGA AAGT K F GGAZ G T ACTT T C C C C C C AATA N M CCAG P V ACAAA T N CACAA H N	AAA FTT 7 ACA T IGT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C T T T C T T T C T T T C T T T C T T T T C T	-1 90 177 267 357 447 537 627 717
AAAAC ATGZ M H TATC Y H AATC N Y CTGC L I ATAC Y A ATAC L F TTGC L A C AATA	CGT	CAC T T T TAC Y TCT S TCC C S S TCC C S S C S S C S S C S S T C C T S T C C T S T C C T T T T	AAAG TTAC: L Q gtt G V V CTTG: L D GTGTC V C TTTA: F N ATAG2 I E AAAAC I E AAAAC	AAAAG a AGGGC A AGGGC T A ACAC T A AACAC T A AACAC T A AACAC T A AACAC T A AACAC T A AACAC T A ACACC T A ACCAC T ACCACACAC ACCACACACA	TATTC G TACTT T A A A A A A A A A A A A A	TTCCA TTTTT F F C TTGAC TTGAC TTGAC TTGAC S S S S S S S S S S S S S S S S S S S	ACAGI ICCIC L C C C C C C C C C C C C C	CAACE V A GGAAG G S CAAGC Q A TACTG Y C GCTG7 A V AAGCT K L AAGCT K L AAGCT R A Q CTGG2 L E TGCA2	ACACO EATT F GCCTT L CAGA D STGAU E STGAU E STGAU S S S S S S S S S S S S S S S S S S S	GAGC TGTA V GATC I TGAC D GGAA E AGAT D GTTA R CAGA R TGTT V t CAAT	AGAC CCTT P I CCAGC Q (CAGCC C S I ATAG I I TTTGG L F ATCA I K t CCCCC P I AACA	CCTCC TIGGI V CAGGA BAGCI BAGCI SATAT D I SATAT D I SATAT D I SATAT D SATAT D SATAT D SATAT D SATAT D SATAT D SACA S	ATCAC GAAGC K I TTATC Y I GAGTC S J TGAAG E A C CATTA L S C ACCTTI L S 8 CAGGG R G CCCTAC L F GTACA	CTCA CAGC A A AAAAT CAGAAAT CTGTA CTGTA CTCTT L CAGAAT CAGAAT CAGAAT CAGAAT	GTGC a ACCTI P GCTIC L ACCCC P TTTIC L TTGCTI A CAAAA K AAGTI S a CAAGS AAGTI S ACCCC A ACCCC A ACCCC A ACCCC A ACCCI A ACCCI A ACCCC A ACCCC A ACCCC A ACCCC A ACCCC A ACCCC A ACCCC A ACCCC A ACCCC A ACCCC A ACCCC A ACCCC A ACCCCC A ACCCCC A ACCCCC A ACCCCC A ACCCCC A ACCCCC A ACCCCC A ACCCCC A ACCCCC A ACCCCC A ACCCCC A ACCCCC A ACCCCC A ACCCCC A ACCCCC A ACCCCC A ACCCCC A A ACCCCC A A ACCCCC A A ACCCCC A A ACCCCC A A ACCCCC A ACCCCC A A ACCCCC A A ACCCCC A A A ACCCCC A A A ACCCCC A A ACCCCC A A A A A A A A A A ACCCC A A A A A A A A A A ACCCC A A A A A A A A A A A A A	CAAGO P I CCCTX P I CACAM T I CCCAC P I CACAM R F N CCCAC L F CCCTCC	ATAC AAAGC X I AAAGC X I AAAGC X I AAAGC J F AGAJ X AACC V F C C C X AAACC I X AACC I I X AAGC I I X AAGC I X I X I X I X I X I X I X I X I X I	AGGA CAGC. 2 Q A AGATA D T T GATA D T CCAAN CCAA	CATT AAGA D CAAT I CAAA N CAAA N CAAA E AGGA F V ICAA K IGCA H AGAT	GCTX TTCI S AAAAK K TTTTX T AAACI T GAAA/ K GAAA/ K CCTI L AAAGC	C TGT ACCCC P GGAT D GGCT P AGCAG G C C ACTG Q C ACTG Q C ACTG Q C ACTG Q	GAGA AAGT K F GGAZ G T ACT T C C C C C C C C C C C C C C C C C C	AAA FTT F ACA F IGT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C TTT C T T T C T C T C T C T C T C T C T C C ACA C T T C C ACA C ACA C T T C C C ACA C ACA C T C C C T T C C C C	-1 90 177 267 357 447 537 627 717 807

Fig. 1. (Continued overleaf).

AATTCCAATTCCTCCTCCCAA		9 aaaaxmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmm	Састассаасал	acmacmaccmacmaccamaa	807
N P I L L A K	CHVNA	FSCLKT	L P V G T Y	Y	091
ca c	8	g c	t	g c	
TTACATGTGTCAAAACTCA	AGCATGGTAACA	AAAGATCATTCTTATTAC	TGTTTACATGTTTATATC	TTCTCCCCACTAAAGCAATTTTC	987
C CCTGTATACCCAGAACCTI	g Tactacgttgaa	g ATGGGATTCTCCTCTGAA	C - ATAATTATTTCACAGCAG	g c t CCATCCTATTTAGGAGTTACAAT	1077
GCCACCCAAACTITCAGAA	g a c ATCTTTTATT-AT	C - G G G CG TCAAGATATA-AAATGTA	i a t CACATTCATATACTACCA	t t t t tg ACATTGAATTCAACAATTACTCA	1165
ttgag t	t c c	tt		- gt	
TTACAAAATCCTGTAAAAA	CCTGCAAGTGTI	PACTICACACACTITACCA	GTTTCCAGAATAATCAAT	GTATTTAACTGTGATAAACCAAA	1255
ACATTTCTTGAGTAAGCT	g a AAGTTTTTACTCO	a tt - CAGTCAAAATGGTTTTTGG	ся с Эсстгааатдаааттттт	cttcaaaactgcat CAGTTAAAA	1331
ggaaaaaaatagcatatg	aaagaaagttat	gtttagagacagtatttc	tccaagtttaacttcactg	aagtttacttaaaaacttggttt	
nfnneese e	tt a	ta			1331
		<u>'9</u> a		g c a	
IGIGAAGAAI	TTTT-ATCAAGC	1 TAGCATACCAAGCAACA	AAGTGAATGGAACATTTC	AAGTATGTATCAAATCATCTTTTT	1412
t	C	t c	AAGTGAATGGAACATTTC a	AAGTATGTATCAAAATCATCTTTTT	1412
t ACATTTTTCTTTCAAAAT	C AATCTAAATGCC	TTAGCATACCAAGCAACA t c ATTATATATTAAAATTGA	AAGTGAATGGAACATTTC 8 ATGTATGCATATAAACAC	AAGTĀTGTATCAAATCATCTTTTT 8 ATTCAACTGTATATGTTACATAAC	1412 1502
t АСАТТТТТСТТТСААААТ t а	c AATCTAAATGCC	TTAGCATACCAAGCAACA t c ATTATATATATAAAATTGA tat	aagtgaatggaacatttc 8 atgtatgcatataaacac CC	ААСТАТСТАТСТАЛАТСАТСТТТТТ а АТТСААСТСТАТАТСТТАСАТААС а	1412 1502
t АСАТТТГТСТТСААААТ t a ТАТТТТТ- АСАСТТТАТА	c caatctaaatgcc g t ttttttaattcaa	ITAGCATACCAAGCAACA t c ATTATATATATTAAAATTGA tat ATATATAATATAATAT	AAGTGAATGGAACATTTC a ATGTATGCATATAAACAC CC AGACATTTTGTTCAGCTG	AAGTÄTGTATCAAATCATCTTTT A ATTCAACTGTATATGTTACATAAC a IGCTCATAGGAGTCATTCTTGCCA	1412 1502 1588
t АСАТТТТТСТТТСААААТ t a ТАТТТТТ-АСАСТТТАТА CC	с с аатстааатссс g t тттттааттсаа са с	ПТАGCАТАССААССААСА t с аттатататтааааттса tat атататаатататат. c	AAGTGAATGGAACATTTC a ATGTATGCATATAAACAC CC AGACATTTTGTTCAGCTG g att a	AAGTÄTGTATCAAATCATCTTTT A ATTCAACTGTATATGTTACATAAC a IGCTCATAGGAGTCATTCTTGCCA C aagac ag	1412 1502 1588
t АСАТТТТТСТТТСААААТ t a ТАТТТТТ-АСАСТТТАТА сс СТТССАТТТТААТАСААС	С ЗААТСТАААТССС д t ТТТТТТААТТСАА Са с АСТСАСТАТССА	TTAGCATACCAAGCAACA t c ATTATATATTAAAATTGA tat ATATATAATATAATAT. C GTAGGGTATGACTCAAGC	aagtgaatggaacatttc a Atgtatgcatataaacac cc Agacattttgttcagctg g att a Acttatttttcttta-	AAGTÄTGTATCAAATCATCTTTT ATTCAACTGTATATGTTACATAAC a IGCTCATAGGAGTCATTCTTGCCA c aagac ag -TTATGACATACTGAGCC	1412 1502 1588 1668
t асатттттсттсаааат t a таттттт-асастттата сс сттссаттттаатасаас с	с аатстааатосс g t тттттааттсаа са с аотсаотатоса	ТТАССАТАССААССААСА t c аттатататтааааттса tat атататаатататат. c стассстатсаастсаасс	AAGTGAATGGAACATTTC a ATGTATGCATATAAACAC cc AGACATTTTGTTCAGCTG g att a ACTTATTTTTCTTTA- g a c g caaaaaaaa	AAGTÄTGTATCAAATCATCTTTTT A ATTCAACTGTATATGTTACATAAC a IGCTCATAGGAGTCATTCTTGCCA C aagac ag -TTATGACATACTGAGCC Iaaaaaaaaaaaa	1412 1502 1588 1668
t ACATTTFTCTTTCAAAAAT t a TATTTTT-ACACTTTATA CC CTTCCATTTTAATACAAC C TTTFTTCCACTGCTTTCTC	С	ТТАССАТАССААССААСА t c аттатататтааааттса tat атататаатататат. c стассоатааастсаасса атсасаатааасттаасал	AAGTGAATGGAACATTTC a ATGTATGCATATAAACAC CC AGACATTTTGTTCAGCTG g att a ACTTATTTTTCTTTA- g a c g caaaaaaaa ACAGATTCAGTTAATGCTC	AAGTÄTGTATCAAATCATCTTTT A ATTCAACTGTATATGTTACATAAC a IGCTCATAGGAGTCATTCTTGCCA c aagac ag -TTATGACATACTGAGCC laaaaaaaaaaa TTACCTGTAAATGAATACATGTT	1412 1502 1588 1668
t ACATTTTTCTTTCAAAAT t a TATTTTT-ACACTTTATA CC CTTCCATTTTAATACAAC C TTTTTCCACTGCTTTCTC CAACTAGAAGTCAGATCG	тттт-алсаасс с затстааатссс g t тттттааттсаа са с астсастатсса ттсасассаааса	ТТАGCАТАССААGCААСА t с аттатататтааааттса tat атататаатататат. c стассотатаастсаасс атсасаатааасттаасал сстстстатасссаатастти	AAGTGAATGGAACATTTC a ATGTATGCATATAAACAC CC AGACATTTTGTTCAGCTG g att a ACTTATTTTTCTTTA- g a c g caaaaaaaa ACAGATTCAGTTAATGCTC ACACAATATACAGAAACAC	AAGTÄTGTATCAAATCATCTTTT a ATTCAACTGTATATGTTACATAAC a IGCTCATAGGAGTCATTCTTGCCA c aagac ag -T TATGACAT - ACTGAGCC aaaaaaaaaaa TTTACCTGTAAATGAATACATGTT FAACATATCTTTAGAATTCTAACT	1412 1502 1588 1668 1758 1848
t ACATTTTTCTTTCAAAAT t a TATTTTT-ACACTTTATA CC CTTCCATTTTAATACAAC c TTTTTTCCACTGCTTTCTC CAACTAGAAGTCAGATCG TAATACTATTCTGTTAAC	ТТТТ-АТСААСС С ЗАТСТАААТССС Я t ТТТТТТААТТСАА Са с АGTCAGTATGCA ТТGACAGGAAACA ААGAAACAAATA СТСАААСААGTT2	ТТАССАТАССААССААСА t с аттатататтааааттса tat атататаатататат. c стассотатааастсаасс атсасаатааасттаасал алсаасаатааттаасал аасаасаатааттаасал	AAGTGAATGGAACATTTC a ATGTATGCATATAAACAC CC AGACATTTTGTTCAGCTG g att a ACTTATTTTTCTTTA- g a c g caaaaaaaa ACAGATTCAGTTAATGCTX ACACACATATACAGAAACAC	ААСТАТСТАТАААТСАТСТТТТ а АТТСААСТСТАТАТСАТААС а IGCTCATAGGAGTCATTCTTGCCA c aagac ag -TTATGACATACTGAGCC Naaaaaaaaaaa CTTACCTCTAAATGAATACATCTT FAACATATCTTTAGAATACATCTT	1412 1502 1588 1668 1758 1848 1938
t ACATTTTTCTTTCAAAAT t a TATTTTT-ACACTTTATA CC CTTCCATTTTAATACAAC C TTTTTTCCACTGCTTTCTC CAACTAGAAGTCAGATCCC TAATACTATTCTGTTAAC AGACCTCAAAGCACATTC	тттт-алсаасс с затстаалтсса д t тттттааттсаа са с адтсадтатсса ттдасаддааса алдааасаадти атааттаттаси	ТТАССАТАССААССААСА t с аттатататтааааттса tat атататаатататат. c GTAGGGTATGACTCAAGC атсасаатаасттаасал атсасаатаасттаасал алсаадааттаттасаст адаааасастаттасат	AAGTGAATGGAACATTTC a ATGTATGCATATAAACAC CC AGACATTTTGTTCAGCTG g att a ACTTATTTTTCTTTA- g a c g caaaaaaaaa ACAGATTCAGTTAATGCTX ACACACATATACAGAAACAC TCAACTCCTTTAAACCAG	AAGTÄTGTATCAAATCATCTTTT a ATTCAACTGTATATGTTACATAAC a IGCTCATAGGAGTCATTCTTGCCA c aagac ag -TTATGACATACTGAGCC aaaaaaaaaaaa CTTACCTGTAAATGAATACATGTT FAACATATCTTTAGAATACATGTT TAACATATCTTTAGAATACATGTT TAACATATCTTTAGAATACATGTT CTTGAAAAGAAAGTGTGTCCC	1412 1502 1588 1668 1758 1848 1938 2028
t ACATTTTTCTTTCAAAAT t a TATTTTT-ACACTTTATA CC CTTCCATTTTAATACAAC C TTTTTTCCACTGCTTTTCCC CAACTAGAAGTCAGATCCC TAATACTATTCTGTTAAC AGACCTCAAAGCACATTC TCAAATACCAACCACAAG	тттт-алсаасс с затстааатсса g t тттттааттсаа са с адтсадтатсса ттдасаддаасааата стсааасаадтя атаатттаттаст	ТТАССАТАССААССААСА t c ATTATATATTAAAATTGA tat ATATATAATATAAATTGA C GTAGGGTATGACTCAAGC ATGACAATAAAGTTAACAA CCTCTGTACCCAATACTTA AACAAGAATTATTTCACT AACAAGACAATATGCCCAA	AAGTGAATGGAACATTTC a ATGTATGCATATAAACAC CC AGACATTTTGTTCAGCTG g att a ACTTATTTTTCTTTA- g a c g caaaaaaaaa ACAGATTCAGTTAATGCTX ACACACTCCTTTTAAACCAG TTACTTCACTACACACACTTX AAAAGTGCTAATTACCGTX	ААСТАТСТАТСАЛАТСАТСТТТТ а АТТСААСТСТАТАТСАТААС а IGCTCATAGGAGTCATTCTTGCCA с aagac ag -TТАТGACATАСТGAGCC наааааааааааа СТТАССТСТАЛААТСААТСТТ ГААСАТАТСТТТАGААТАСАТСТТ ГААСАТАТСТТТАGААТАСАТСТТ ГААСАТАТСТТТАGААТССТСАССТТ СССТСССТПСАСАСТТАТСТСС ЗАСТААСААСАТССТААССТТ	1412 1502 1588 1668 1758 1848 1938 2028 2118
t ACATTTTTCTTTCAAAAT t a TATTTTT-ACACTTTAAA CC CTTCCATTTTAATACAAC C TTTTTTCCACTGCTTTTCTC CAACTAGAAGTCAGATCC TAATACTATTCTGTTAAC AGACCTCAAAGCACATTC TCAAATACCAACCACAAG GTAGCTAGTGTAGATTTC	тттт-алсаасс с затстааатсса g t тттттааттсаа са с адтсадтатсса ттдасадааасааата аддааасааата атаатттаттаст сттааастсттс таатттцагасааа	ТТАССАТАССААССААСА t с аттатататтааааттса tat ататататаатаатат. c GTAGGGTATGACTCAAGC GTAGGGTATGACTCAAGC атсасааатастта аасаадааттатттсас: аасаадааттатттсас: Таасадасаааатсоссааа атастаадасаааатсоссааа атастаадасаааатсоссааа	AAGTGAATGGAACATTTC a ATGTATGCATATAAACAC CC AGACATTTTGTTCAGCTG g att a ACTTATTTTTCTTTA- g a c g caaaaaaaaa ACAGAATATACAGTAATGCTX ACACAATATACAGAAACAC ICCACTTCACTACACACTTX AAAAGTGCTAATTACCGTX AGTAATACACTGCCATCAAA	ААСТАТСТАТАСТАЛАТСАТСТТТТ а АТТСААСТСТАТАТСТТСАТААС а IGCTCATAGGAGTCATTCTTGCCA с aagac ag -TТАТСАСАТАСТСАССС аааааааааааааа УТТАССТСЯТАААТСААТС-ТТ ГААСАТАТСТТТАСААТАСТАТТ ГААСАТАТСТТТАСААТАСТАСТ ТГССТСССТТСАССАСТТАТСТС С ЗАСТААСААСАТССААСТТТ ГААСАТАССТСААТСААСТСТ ССТСССТСААСАСТСААССТТ ААСААТАСТСААТААТААТАТТТСС	1412 1502 1588 1668 1758 1848 1938 2028 2118 2208
t ACATTITITCTTTCAAAAT t a TATTITTT-ACACTTTATAA CC CTTCCATTITTAATACAAC C TTTTTTCCACTGCTTTTCTC CAACTAGAAGTCAGATTCC TAATACTACAAGCACATTCC TCAAATACCAACCACAAG GTAGCTAGTGTAGATTTCC TAATATGTCTATCTTCTTTATT	тттт-алсаасс с затстааатсса д t тттттааттсаа са с адтсадтатсса ттдасаддаааса аддааасааата аддааасааата аддааасааата атдастасассосо сттааастасасосос стасотого	ТТАGCATACCAAGCAAGCA t c ATTATATATATTAAAATTGA tat ATATATAATAATATA TAT c GTAGGGTATGACTCAAGC GTAGGGTATGACTCAAGC ATGACAATAAAGTTAACAA CCTCTGTACCCCAATACTT AAACAAGAATTATTCACT AAACAAGAATTATTCCAT TAACAGACAAATGCCCAAA ATAGTTAAGCACAAATGCCCAAA ATAGTTAAGCACAAATGCCCAAA ATAGTTAAGCACAAATGCCCAAA	AAGTGAATGGAACATTTC a ATGTATGCATATAAACAC CC AGACATTTTGTTCAGCTG g att a ACTTATTTTTCTTTA- g a c g caaaaaaaaa ACAGATTCAGTTAATGCTX ACACACAATATACAGAAACAA ITACTTCACTACAGAAACAA ITACTTCACTACAGAAACAA ACAGAGTGCTAATTACCGTX AAAAGTGCTAATTACCGTX ACAGAACTGCCAATCACCGTX ACAGTACCAGCACTGCCATCAAA DTCTAGAGAAATTATCCTX	AAGTATGTATCAAATCATCTTTTT A ATTCAACTGTATATGTTACATAAC A IGCTCATAGGAGTCATTCTTGCCA C aagac ag TTATGACATACTGAGCC Aaaaaaaaaaaa STTACCTGTAAATGATACATGTT FAACATATCTTTAGAATACATGTT TGTTGAAAAGAGAAATGATGTGTGC STCCTGCCTTGTCAGCTTATCTCC GACTAAGAACATCGATCAAACTTT AAGAATACTGATAAAAATATTTTG ACGTTTAACATACTGATGATGCTCTGCC	1412 1502 1588 1668 1758 1848 1938 2028 2118 2028 2118 2298
t ACATTTTTCCTTTCAAAAT t a TATTTTT-ACACTTTATA CC CTTCCATTTTAATACAAC C TTTTTTCCACTGCTTTCTC CAACTAGAAGTCAGATCG TAATACTAAAGCACAATC AGACCTCAAAGCACAAAGG GTAGCTAAGTAGCACACAAGG GTAGCTAGTGTAGATTTC TATTATGTCTTATCTTA	тттт-алсаасс с аатстааатссс g t тттттааттсаа са с астсадтатсса ттдасаддаасааата сттааасаадтя атаатттаттас; сттааастаттаттас; стаадстасассссс стасдстасассссс стасдстасастт аасстятасаст	ТТАGCATACCAAGCAACA t c ATTATATATATAAAATTGA tat ATTATATAATATAAAATTGA tat ATATATAATATAAAATTGA C GTAGGGTATGACTCAAGC ATGACAAATAAAGTTAACAA CCTCTGTACCCAAATACTTA AGAAAACAGTATTACTTAA AGAAAACAGTAATTACCCAAA TAGTTAAGCACAAATTACC CATTGAGCATAGCCACGAC	AAGTGAATGGAACATTTC a ATGTATGCATATAAACAC CC AGACATTTTGTTCAGCTG g att a ACTTATTTTTCTTTA- g a c g caaaaaaaaa ACAGATTCAGTTAATGCTA ACACACAATATACAGAAACA TTACTTCACTACAGACACTTA AAAAGTGCTAATTACCGTA ACGTATCACTGCCATCAAA CGTATCACTGCCATCAAATTACCGTA ACGTATCACTGCCATCAAATTACCGTA ACGTATCACTGCCATCAAATTACCGTA ACGTATCACTGCCATCAAATTACCGTA ACGTATCACTGCCATCTAATTACCTAAA	AAGTÄTGTATCAAATCATCTTTT 8 ATTCAACTGTATATGTTACATAAC 8 IGCTCATAGGAGTCATTCTTGCCA C aagac ag -TTATGACATACTGAGCC Baaaaaaaaaa STTACCTGTAAATGAATACATGTT TAACATATCTTTAGAATACATGTT TACCTGCCTTGTCAGCTAACT TCCTGCCTTGTCAGCAGAATACTTCC SACTAAGAACATCGTACTAACTTT AAGAATACTGATAAAATATTTTG AGGTTTAACATACTGATACAAATACTTCAAAA	1412 1502 1588 1668 1758 1848 1938 2028 2118 2028 2118 2298 2388 2388
t ACATTTTTCCTTTCAAAAT t a TATTTTT-ACACTTTATA CC CTTCCATTTTAATACAAC C TTTTTTCCACTGCTTTCTC CAACTAGAAGTCAGATCG TAATACTAAAGCACATAC AGACCTCAAAGCACATAC GTAGCTAAGGCACATAC GTAGCTAGTGTAGATTTC TATTATGTCTTACTTATT CACTATAAGTGCTGATAT	с алтсталатосс g t ттттталтсал са с застсадтатоса тталсадаласалата сталасалата аталттаттас; стталастатостос талттостадал тастассосо стасадаастосто тастатосто тасастасососо стасастасасосо тасастасасососо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасососо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасасосо тасастасососо тасасососо тасасососо тасасососо тасасосососо тасасососососо тасасососососососо тасасосососососососососососососососососо	ТТАGCАТАССААGCAACA t c АТТАТАТАТТААААТТGA tat АТАТАТААТАТАААТТGA С GTAGGGTATGACTCAAGC ATGACAATAAAGTTAACAA CCTCTGTACCCAATACTTA AGAAAACAGTATTACTTA AGAAAACAGTATTACTTA AGAAAACAGTAATTACCAAT TAACAGACAAATGCCCAAT ATGATTAAGCACAAATACC STCAAGCACAACTAACGAGT	адатсаатсааасатттс а АТСТАТССАТАТАААСАС СС АСАСАТТТТСЭТСАССТС д att a АССТТАТТТТТСТТТА- д a c g сааааааааа АСАСАТТСАСТТАААТСАСАС ССАССАСТТТААССАСАСТ ТГАСТТСАСТАСАСАСТТ ААААGТССАТТТААССАСАСТ АСАСАСАСТАСТСАСТАСАСАСТТ ААААGТССАТТТААССАСАСТ ЭТТТССТТТССТТСАСТАСАТТАА СССТАТСАТТТССАТТТАА СССТАТСА	AAGTATGTATCAAATCATCTTTTT A ATTCAACTGTATATGTTACATAAC A TCCTCATAGGAGTCATTCTTGCCA C aagac ag TTATGACATACTGAGCC Aaaaaaaaaaaa TTACCTGTAAATGATACATGTT TCACATATCTTTAGAATCCATGTT TCCTGCCTTGTCAGCTTATCTACC TCCTGCCTTGTCAGCTTATCTCC CACTAAGAACACTGATAAAATATTTTG ACGTTTAACATACTGATACATGTTAAAA AATGCTTACTGTATGATATGATAATGTAA	1412 1502 1588 1668 1758 1848 1938 2028 2118 208 2298 2388 2478 2533

Fig. 1. (Continued).

burgh et al., 1997). Mimecan is transcribed into several mRNAs of different sizes that give rise to the same amino acid sequence (Funderburgh et al., 1997; Tasheva et al., 1997).

KSPGs increase in chick cornea during embryonic development (Hart, 1976; Funderburgh et al., 1986; Takahashi et al., 1999). Lumican, keratocan and mimecan mRNA levels are highest during early development, but decline as the cornea becomes transparent (Dunlevy et al., 2000). In mouse embryos, keratocan mRNA is detected in neural crest cells migrating toward the cornea and sclera at E13.5 (embryonic Day 13.5), then becomes restricted to corneal stromal cells by E18.5 (Liu et al., 1998). However, very little is known about when neural crest cells begin to make the other corneal KSPG core proteins, or whether there are regional differences across the cornea in core protein synthesis, polymerization, accumulation, distribution or turnover. Quail/chick chimeric corneas provide an excellent experimental system in which to test the functional roles of individual proteoglycans in specific embryonic regions (e.g. by transfecting a quail neural crest graft with a quail-specific antisense RNA before transplantation into a chick embryo host). To obtain quail-specific molecular probes for such studies, we have cloned and now report the cDNA and deduced amino acid sequences for quail keratocan and mimecan, their comparison with the corresponding chick sequences, and their expression in quail cornea and other tissues. (Lumican and another corneal proteoglycan, decorin, are presented in Corpuz et al., 2000) Our results confirm the presence of an avian-specific mimecan splice variant that gives rise to a unique amino acid sequence (Dunlevy et al., 2000), and demonstrate that in birds, as in mice, keratocan appears to be expressed only in the cornea.

2. Materials and methods

All the experimental procedures used here have been described previously (Corpuz et al., 2000). The quail (*Coturnix coturnix japonica*) cornea/sclera cDNA library was screened with chick keratocan and mimecan cDNA probes (Dunlevy et al., 1998, 2000). Three clones for each of the two KSPGs were sequenced. The quail primers used for the generation of DNA templates are: keratocan, sense: 5'-GCTAACA-CACTGCAGCTGT-3' (position 640), antisense: 5'-TCTTACAGCACAAGAGTGC-3' (position 1116); and mimecan, sense: 5'-GATACAGATATTGT-CACGGG-3' (position 100), antisense: 5'-ATTC-CAGTGCGTTATGTCC-3' (position 664). The polymerase chain reactions (PCRs) generated 477-bp and 565-bp PCR products for keratocan and mimecan, respectively, which were used to synthesize antisense riboprobes for ribonuclease protection assay (RPA) studies.

In conducting RPA experiments, we attempted to use both quail β -actin and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as internal controls. However, we found their mRNA expression to be quite variable from tissue to tissue (Fig. 2). We therefore used 28S and 18S ribosomal RNAs as loading controls, for ethidium bromide staining reveals their uniform expression across all quail tissues. Preparation of β -actin and GAPDH antisense riboprobes has been described (Corpuz et al., 2000).

3. Results and discussion

The cDNA and deduced amino acid sequences for quail keratocan and mimecan, including comparisons with published chick sequences (Dunlevy et al., 1998, 2000), are shown in Fig. 1. The quail keratocan clone (Fig. 1a) is 3108 bp long and encodes a protein containing 353 amino acids. At the nucleotide level, the quail coding region is 98% identical to the chick sequence (Dunlevy et al., 1998), resulting in a single conservative amino acid difference at residue 253. In contrast, quail and chick 3'UTRs are only 86% identical. The three quail keratocan clones sequenced are

nt 750 -

500 -400 -

300

identical in the 5'UTR and translation regions, but are variable in the 3'UTR (data for two other clones not shown). Variable 3'UTRs were also found in chick (Dunlevy et al., 1998).

The quail mimecan clone (Fig. 1b) is 3066 bp long and the derived protein contains 293 amino acids. Quail mimecan shares 96% nucleotide identity with the chick sequence in the coding region, but only 76%in the 3'UTR (Dunlevy et al., 2000). This results in 97% identity at the protein level: one amino acid difference (residue 7) in the putative signal peptide region, six differences in the amino-terminal region of the mature protein, and two amino acid differences toward the carboxy-terminal end. Most significantly, the 58 amino acids (residues 28-85) immediately following the signal peptide in quail are 94% identical to the corresponding 59-amino acid region in chick [previously reported as containing 60 amino acids in Dunlevy et al. (2000)], whereas this region shows no homology to corresponding bovine, human and mouse mimecan regions. The nucleotide sequence in this avian-unique cDNA (quail nucleotides 82-255; chick nucleotides 313-489) is 46% (quail) and 47% (chick) identical to bovine intron 4 nucleotides 13907-14096 (Accession No. AF105150), suggesting that birds have novel splice sites in their mimecan gene that replace the 3' end of mammalian exon 3 and all of exon 4 with a new avian exon 4 from intron sequences located upstream of exon 5 (Dunlevy et al., 2000). This results in an avian-specific amino acid sequence that is distinct from the mammalian sequence.

RPA analysis demonstrates that keratocan and lumican mRNAs are expressed most highly in Day 16

Mimecan

Lumican

- β-ACTIN

28S

Keratocan



us er ekeletal ma

De = = Co = = = - Gapdh

-

embryonic quail cornea, whereas mimecan mRNA is significantly less expressed (Fig. 2). Quail keratocan expression appears to be limited exclusively to the cornea. No detectable message is observed in the cartilaginous quail sclera, in agreement with data from embryonic Day 18.5 and adult mouse sclera (Liu et al., 1998). However, keratocan mRNA is found in the non-cartilaginous fibrous bovine sclera (Corpuz et al., 1996). Conversely, lumican and mimecan mRNAs are expressed in multiple quail tissues, with lumican generally more highly expressed than mimecan. Both transcripts are low or undetectable in quail brain, liver and sternum. Our quail data agree with previous studies on chick cornea, which showed a much lower level of mimecan mRNA compared to lumican and keratocan mRNAs (Dunlevy et al., 2000), but not with results in bovine cornea, which showed higher levels of both mimecan and lumican mRNAs compared to keratocan mRNA (Funderburgh et al., 1997). This study is the first to assess the avian expression of keratocan and mimecan transcripts in multiple tissues. It shows that in most quail tissues four mimecan transcripts are detectable by RPA: one major and one minor transcript, as well as two minimally detectable transcripts that migrate with and just below the lumican transcript. Northern blotting and RPA studies have also identified multiple mimecan transcripts in bovine cornea and various other tissues (Funderburgh et al., 1997; Tasheva et al., 1997).

Acknowledgements

This work has been supported by National Institutes of Health grants EY08104 to J.R.H. and EY00952 to G.W.C.).

References

Blochberger, T.C., Vergnes, J.-P., Hemper, J., Hassell, J.R., 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.

- Corpuz, L.M., Funderburgh, J.L., Funderburgh, M.L., Bottomley, G.S., Prakash, S., Conrad, G.W., 1996. Molecular cloning and tissue distribution of keratocan. Bovine corneal keratan sulfate proteoglycan 37A. J. Biol. Chem. 271, 9759–9763.
- Corpuz, L.M., Dunlevy, J.R., Hassell, J.R., Conrad, A.H., Conrad, G.W., 2000. Molecular cloning and relative tissue expression of decorin and lumican in embryonic quail cornea. Matrix Biol. 19, 693–697.
- Dunlevy, J.R., Neame, P.J., Vergnes, J.-P., Hassell, J.R., 1998. Identification of the N-linked oligosaccharide sites in chick corneal lumican and keratocan that receive keratan sulfate. J. Biol. Chem. 273, 9615–9621.
- Dunlevy, J.R., Beales, M.P., Berryhill, B.L., Cornuet, P.K., Hassell, J.R., 2000. Expression of the keratan sulfate proteoglycans lumican, keratocan and osteoglycin/mimecan during chick corneal development. Exp. Eye Res. 70, 349–362.
- Funderburgh, J.L., Caterson, B., Conrad, G.W., 1986. Keratan sulfate proteoglycan during embryonic development of the chicken cornea. Dev. Biol. 116, 267–277.
- Funderburgh, J.L., Corpuz, L.M., Roth, M.R., Funderburgh, M.L., Tasheva, E.S., Conrad, G.W., 1997. Mimecan, the 25-kDA keratan sulfate proteoglycan, is a product of the gene producing osteoglycin. J. Biol. Chem. 272, 28089–28095.
- Hart, G.W., 1976. Biosynthesis of glycosaminoglycans during corneal development. J. Biol. Chem. 251, 6513–6521.
- Iozzo, R.V., 1997. The family of the small leucine-rich proteoglycans: key regulators of matrix assembly and cellular growth. Crit. Rev. Biochem. Mol. Biol. 32, 141–174.
- Liu, C.-Y., Shiraishi, A., Kao, C.W.-C., Converse, R.L., Funderburgh, J.L., Corpuz, L.M., Conrad, G.W., Kao, W.W.-Y., 1998. The cloning of mouse keratocan cDNA and genomic DNA and the characterization of its expression during eye development. J. Biol. Chem. 273, 22584–22588.
- Madisen, L., Neubauer, M., Plowman, G., Rosen, D., Segarini, P., Dasch, J., Thompson, A., Ziman, J., Bentz, H., Purchio, A.F., 1990. Molecular cloning of a novel bone-forming compound: osteoinductive factor. DNA Cell Biol. 9, 303–309.
- Takahashi, I., Nakamura, Y., Hamada, Y., Nakazawa, K., 1999. Immunohistochemical analysis of proteoglycan synthesis during early development of the chick cornea. J. Biochem. 126, 804–814.
- Tasheva, E.S., Corpuz, L.M., Funderburgh, J.L., Conrad, G.W., 1997. Differential splicing and alternative polyadenylation generate multiple mimecan mRNA transcripts. J. Biol. Chem. 272, 32551–32556.