Gene: MSH6 - Sequence: NG_007111.1 Transcript: NM_000179.2 - Protein: NP_000170.1 Date : March 2, 2015

1^{st} line: Base numbering. Full stops for intronic $+/-$ 5, 10, 15 2^{nd} line: Base sequence. lower case Introns, upper case Exons 3^{rd} line: Amino acid sequence. Printed on FIRST base of codon 4^{th} line: Amino acid numbering. Numbered on 1^{st} and increments of 10														
Exon 1 Start: 4936	End: 534	7 Length	n: 411											
	getgtteeeg	cttccgctcd	cagagaggca	 .gggctttccg										
agcctgctagcccgcggccgc	· · ·	• cgggtcggag	gtgttccggc	· · ·										
	gctcagcagt	teccegegeg	ggggcccagg	 regteggegge										
agggegggeeecteacegeeag	gegtgeeage	cccgccccta	 acccaccagt	 gtgccagccc										
	icacccaada	geggggeetg	ggcgcgcacc	 geeegegeae										
-149 -139 GGCGAGGCGCCTGTTGATTGGC	•	•	•	•										
-89 -79 CCCAGATTTCCCGCCAGCAGGA														
-29 -19 GACAGAACGGTTGGGCCTTGCC		1 TATGTCGCGA M S R 1		21 CCTGTACAGCT L Y S F										
31 41 TCTTCCCCAAGTCTCCGGCGCT F P K S P A L 11	51 GAGTGATGC S D A	CAACAAGGC	CTCGGCCAGG	81 GCCTCACGCG A S R E										

91	101	111	121	131	141
				TTCCCCAGGCGG	
G G R 31	A A A	A P G	A S P 41	S P G G	D A A W
151	161	171	181	191	1201
GGAGCGAGGC				CTCCGCGTCACC	GCCCAAGGCGA
		PRP			
51	0 1 0	. 1 10 1	161	0 11 0 1	1 10 11 10
131			101		
211	221	231	241	251	
•				TGCTGCCCCAC	CNCat sacaaa
	G G I	ı K K S		A A P T	S
71			81		
				• •	• •
gtgggggtgg	ggtcgaag	gcgggggca	tagcggcggg	gcgcttggaacc	cggcgagggga
ggctcgcaca	ıgggggttg	ggggggtgc	acggcctggc	cctgggctcgga	ggaggcggggc
cgcagagttg	gcttgaat	gagtgcagg	ggtcgagtct	ggagcatttggg	ggtgtagcttg
taaacagggt	сааааааа	agagget.gt.	асаааааааа	ggctgcagggga	аасасаааааа
	9 9 - 9 9 - 9	, e. 5 e. 5 5 e e 5 e	5 5 5 5 - 5 - 5	5500500	5 5 - 5 5 5 5
ttcaaacctt	ttaaaaaa	addadacdc	atcccaccaa	· gtgggggtgctg	aac
2229990000	9 9 4 9 9 9	assasacsc.	5 ccc gccag	2-22222252566	5 5 C

Exon 2 Start: 12781 End: 12977 Length: 196	
ttcctgccatcagcattataccaaaattctgccatggtttttaaactttgattctgaga	a
	t
	С
	a
	g
261 271 281 291 301 311 TTGTGACTTCTCACCAGGAGATTTGGTTTGGGCCAAGATGGAGGGTTACCCCTGGTGGC C D F S P G D L V W A K M E G Y P W W P 91 101	
321 331 341 351 361 371 TTGTCTGGTTTACAACCACCCCTTTGATGGAACATTCATCCGCGAGAAAGGGAAATCAG C L V Y N H P F D G T F I R E K G K S V 111	
381 391 401 411 421 431 CCGTGTTCATGTACAGTTTTTTGATGACAGCCCAACAAGGGGCTGGGTTAGCAAAAGGC R V H V Q F F D D S P T R G W V S K R L	
131 141	
441 451	g
	g
	g

Exon 3 Start: 17748 End: 17917 Length: 169														
gattacaggcgtgagcctctgcacccggcccttattgtttataaatacatttctttc														
461 471 481 491 501 511 GTTCAAAATCAAAGGAAGCCCAGAAGGGAGGTCATTTTTACAGTGCAAAGCCTGAAATAC														
S K S K E A Q K G G H F Y S A K P E I L 161 171														
TGAGAGCAATGCAACGTGCAGATGAAGCCTTAAATAAAGACAAGATTAAGAGGCTTGAAT R A M Q R A D E A L N K D K I K R L E L 181 191														
581 591 601 611 621														
TGGCAGTTTGTGATGAGCCCTCAGAGCCAGAAGAGAGAAGAAGAGATGGAGGTGGAGGTGGAGGACG A V C D E P S E P E E E E M E 201														
aacaagataccttgttttatatatgtgtgtgtatatgtattatttat														
catacttctgtagttccctggactgtaggataagttaggttacttagaatctcaacagct														
agcatcgtttttacttaggttttcaagcctactggcagggtaagcaagaggtagtaccat														
tttggtaagaagtagagactagggacagtaaagatggagtaatatat														

EXC	on 4	St	art	: 2	0465	5	End	d:	230	09	⊥€	eng	th:	25	44			
• aaa	acgtgt	tta	ata	cac	ctad	CCC	caco	cga	ata	ttg	tago	ctt	ggg	cgt	agc	cta	acc	tatc
tca	agacgt	gct	cag	aac			atgt							aga	tca	tct	aaca	acaa
ago	cctatt	tta	taa	taa	ggaa	att	gcct	at	ctc	atg	taat	ctc	atc	gaa	tac	tgt	acta	aaaa
ato	gaaaaa	cag	tgg	ctg	cac		taco	cat	tat	aaa	gtc	aaa	aaa	tca	taa	gtt	• gaa	ctgt
ctt	tacatt	atg	gtt	ttc	caaa	att	ttga	att	• tgt	ttt	taaa	ata	ctc	ttt	cct	tgc	ctg	gcag
GT <i>P</i> V	631 AGGCAC G T	AAC T			AAC	AGA'					AGA:	ΓΑΑ' N	rga.		TGA			AGAG
V	G T 211	1	1	V	Т	D	K	5	Ε	E 22	D 21	IN	Ε	Ι	Ε	S	Е	Ε
GAA	691 AGTACA	.GCC		701 GAC	ACA?		711 ATCI		GCG.	7: AAG		CCG		731 AAT	AAA		74: ACG	
Ε	V Q 231	P	K	Τ	Q	G	S	R	R	S 2		R	Q	I	K	K	R	R
	751			761			771			7				791			180	
GT(CATATC I S	AGA D	TTC S	TGA E	GAG' S	IGA: D	CATI I		TGG G	CTC: S		IGT(V		ATT' F	TAA K	GCC P	AGA(D	CACT T
V	251	D	D	تا	5	D	Τ.	G	G	2		V	Ľ	Ľ	11	L	D	1
	811			821			831			8				851			86	
AA(GGAGGA E E	AGG. G	AAG S		TGA: D	ГGA. Е	AAT <i>I</i> I	AAG S	CAG' S	TGG/ G		GGG G	GGA' D		ГGА Е	GAG S	TGAZ E	AGGC G
V	271	G	5	S	ע	Ŀ	1	5	۵	2		G	ע	۵	L	5	Ŀ	G
	871			881			891	L		191	01		1	911			192	1
	GAACAG																	
L	N S	Р	V	K	V	A	R	K	R	K		M	V	Τ	G	N	G	S
	291									130	Ul							

CT:	931 941 CTTAAAAGGAAAAGCTCTAGGA																		
L	K R 311	K	S	S	R	K	Ε	Τ	Р	S A 321	Τ	K	Q	А	Τ	S	I		
TCZ S	991 ATCAGA S E							AGC		1021 CTCTGC S A	CCCC'	TCA	AAA						
~	331	_			-	_			-	341	-	×		~	_	Š	£		
GC	1051 CCACGT									1081 TCGCCC									
A	H V 351	S	G	G	G	D	D	S	S	R P 361	Т	V	W	Y	Н	Ε	T		
ТТ	1111 AGAATG									1141 TGAGC									
L	E W 371	L	K	E	E	K	R		D	E H 381	-	-	R			Н	Р		
GA	1171 TTTGA									1201									
D	F D 391	A					V			D F 401			S	С	Т	P			
ΑТ(1231 GAGGAA									1261 CTTTG									
M	R K 411	W	W	Q		K				F D 421						K			
GG	1291 GAAATT									1321 TCTTAT									
G	K F 431	Y	E	L		Н			A							G			
СТ	1351 ATTCAT									1381									
	F M 451																		
TC:										1441									
	AGATTC D S 471																		

| 1471 | 1481 | 1491 | 1501 | 1511 | 1521 GAAATGATGGAGGCACGATGTAGAAAGATGGCACATATATCCAAGTATGATAGAGTGGTG E M M E A R C R K M A H I S K Y D R V V |491 |501 |1531 |1541 |1551 |1561 |1571 |1581 $\tt AGGAGGGAGATCTGTAGGATCATTACCAAGGGTACACAGACTTACAGTGTGCTGGAAGGT$ R R E I C R I I T K G T Q T Y S V L E G 1511 1521 |1601 |1611 |1621 |1631 11591 11641 GATCCCTCTGAGAACTACAGTAAGTATCTTCTTAGCCTCAAAGAAAAAGAGGAAGATTCT D P S E N Y S K Y L L S L K E K E E D S |531 |541 |1651 |1661 |1671 |1681 |1691 ${\tt TCTGGCCATACTCGTGCATATGGTGTGTGTTGTTGATACTTCACTGGGAAAGTTTTTC}$ S G H T R A Y G V C F V D T S L G K F F 1551 |561 |1711 |1721 |1731 |1741 |1751 |1761 ATAGGTCAGTTTTCAGATGATCGCCATTGTTCGAGATTTAGGACTCTAGTGGCACACTAT I G Q F S D D R H C S R F R T L V A H Y |571 1581 11771 |1781 |1791 |1801 |1811 |1821 CCCCCAGTACAAGTTTTATTTGAAAAAGGAAATCTCTCAAAGGAAACTAAAACAATTCTA P P V Q V L F E K G N L S K E T K T I L 1591 1601 |1871 |1881 |1831 |1841 |1851 |1861 AAGAGTTCATTGTCCTGTTCTCTCAGGAAGGTCTGATACCCGGCTCCCAGTTTTGGGAT K S S L S C S L Q E G L I P G S Q F W D |611 |621 |1891 |1901 |1911 |1921 |1931 |1941 GCATCCAAAACTTTGAGAACTCTCCTTGAGGAAGAATATTTTAGGGAAAAGCTAAGTGAT A S K T L R T L L E E E Y F R E K L S D |631 |641 11961 11971 |1981 |1991 11951 ${\tt GGCATTGGGGTGATGTTACCCCAGGTGCTTAAAGGTATGACTTCAGAGTCTGATTCCATT}$ G I G V M L P Q V L K G M T S E S D S I 1651 1661

|2011 |2021 |2031 |2041 |2051 |2061 GGGTTGACACCAGGAGAAAAGTGAATTGGCCCTCTCTGCTCTAGGTGGTTGTCTTC G L T P G E K S E L A L S A L G G C V F |671 |681 |2071 |2081 |2091 |2101 |2111 |2121 TACCTCAAAAAATGCCTTATTGATCAGGAGCTTTTATCAATGGCTAATTTTGAAGAATAT Y L K K C L I D Q E L L S M A N F E E Y 1691 1701 |2141 |2151 |2161 |2171 |2181 12131 ATTCCCTTGGATTCTGACACAGTCAGCACTACAAGATCTGGTGCTATCTTCACCAAAGCC I P L D S D T V S T T R S G A I F T K A 1711 721 |2191 |2201 |2211 |2221 |2231 |2241 ${\tt TATCAACGAATGGTGCTAGATGCAGTGACATTAAACAACTTGGAGATTTTTCTGAATGGA}$ Y Q R M V L D A V T L N N L E I F L N G 1731 741 12251 |2261 |2271 12281 |2291 |2301 ${\tt ACAAATGGTTCTACTGAAGGAACCCTACTAGAGAGGGTTGATACTTGCCATACTCCTTTT}$ T N G S T E G T L L E R V D T C H T P F |751 761 12311 |2321 |2331 |2341 |2351 |2361 GGTAAGCGGCTCCTAAAGCAATGGCTTTGTGCCCCACTCTGTAACCATTATGCTATTAAT G K R L L K Q W L C A P L C N H Y A I N 1771 1781 |2371 |2411 |2421 |2381 |2391 12401 GATCGTCTAGATGCCATAGAAGACCTCATGGTTGTGCCTGACAAAATCTCCGAAGTTGTA D R L D A I E D L M V V P D K I S E V V 1791 801 |2431 |2441 |2451 |2461 |2471 |2481 GAGCTTCTAAAGAAGCTTCCAGATCTTGAGAGGCTACTCAGTAAAATTCATAATGTTGGG E L L K K L P D L E R L L S K I H N V G 811 821 12501 12511 |2521 |2531 |2541 12491 ${\tt TCTCCCTGAAGAGTCAGAACCACCCAGACAGCAGGGCTATAATGTATGAAGAAACTACA}$ S P L K S Q N H P D S R A I M Y E E T T

1841

1831

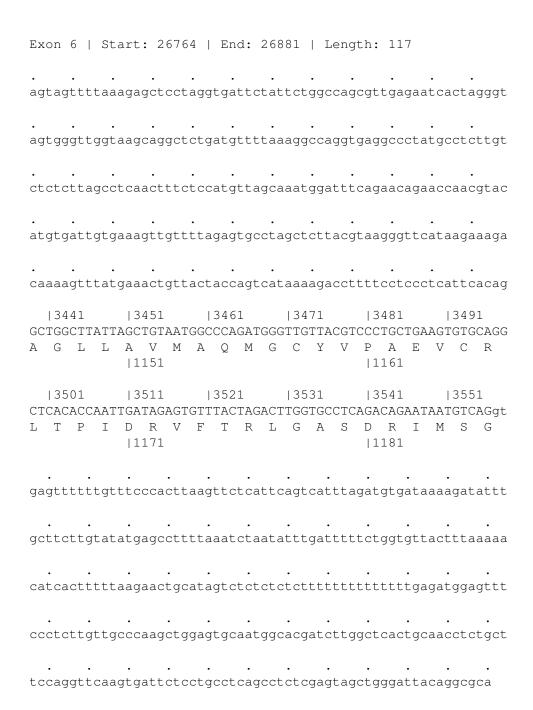
|2551 |2561 |2571 |2581 |2591 |2601 TACAGCAAGAAGAAGATTATTGATTTTCTTTCTGCTCTGGAAGGATTCAAAGTAATGTGT Y S K K K I I D F L S A L E G F K V M C 1851 861 |2611 |2621 |2631 |2641 |2651 |2661 AAAATTATAGGGATCATGGAAGAAGTTGCTGATGGTTTTAAGTCTAAAATCCTTAAGCAG KIIGIMEEVADGFKSKILKQ 1871 1881 |2681 |2691 |2701 |2711 12721 GTCATCTCTGCAGACAAAAAATCCTGAAGGTCGTTTTCCTGATTTGACTGTAGAATTG V I S L Q T K N P E G R F P D L T V E L 1891 1901 |2731 |2741 |2751 |2761 |2771 |2781 ${\tt AACCGATGGGATACAGCCTTTGACCATGAAAAGGCTCGAAAGACTGGACTTATTACTCCC}$ N R W D T A F D H E K A R K T G L I T P 1911 1921 12791 12801 |2811 12821 |2831 |2841 AAAGCAGGCTTTGACTCTGATTATGACCAAGCTCTTGCTGACATAAGAGAAAATGAACAG K A G F D S D Y D Q A L A D I R E N E Q |931 1941 12851 |2861 |2871 |2881 |2891 |2901 AGCCTCCTGGAATACCTAGAGAAACAGCGCAACAGAATTGGCTGTAGGACCATAGTCTAT S L L E Y L E K Q R N R I G C R T I V Y 1951 1961 |2951 |2961 12911 |2921 |2931 |2941 $\tt TGGGGGATTGGTAGGAACCGTTACCAGCTGGAAATTCCTGAGAATTTCACCACTCGCAAT$ W G I G R N R Y Q L E I P E N F T T R N 1971 |981 |2971 |2981 |2991 |3001 |3011 |3021 $\tt TTGCCAGAAGAATACGAGTTGAAATCTACCAAGAAGGGCTGTAAACGATACTGGACCAAA$ L P E E Y E L K S T K K G C K R Y W T K 1991 |1001 13041 13051 |3061 |3071 |3081 13031 ACTATTGAAAAGAAGTTGGCTAATCTCATAAATGCTGAAGAACGGAGGGATGTATCATTG T I E K K L A N L I N A E E R R D V S L 11011 11021

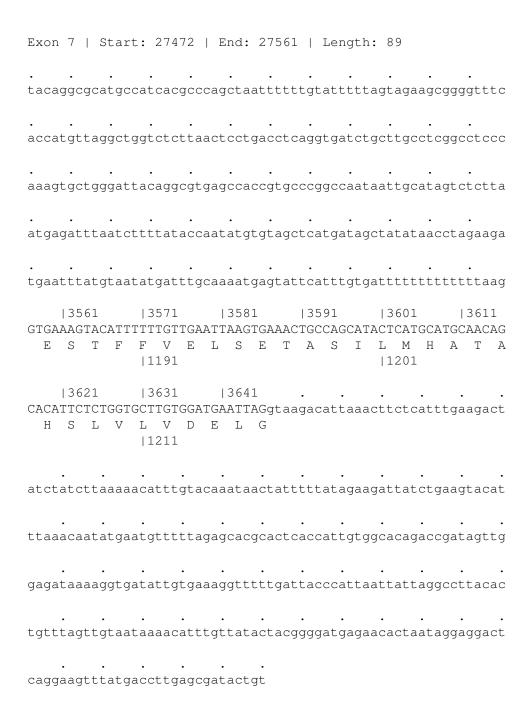
3091	1	3101		3111		3121		313	1	314	1
AAGGACTGC											TCT
K D C 1031											
3151 GCTGTAGAG							caaqo	cttati	tctca	aactt	tga
A V E 1051										J c	- 9 -
taagtagtg	gctgtt	tgcca	ıgctgt	catatt	tatcc	ctaaaa	aataa	igtaa	taagg	tatat	.atg
gtacatatt						tcctga			cccac	agcaa	ıttt
aagttactt	• :gaaac			ccttac		ccttt	ggcct	· ttct	tcagt	ttttt	.ttt
tttttttt	ttttt	ttgag	Jacato	ggtctt	igctc	tgttg:	· cccaç	· ggcta	gaata	tggtg	gaca
caaccatgg	gctact	gcago	cctcga	·							

```
Exon 5 | Start: 25274 | End: 25539 | Length: 265
gttggtcaggctggtcctgaactgctgacctcgtgaactctgcccgcctaggcctcctga
agtgctgggattacaggcgtgagccaccgtgcctggcctctgctctatcttttagctttc\\
\verb|ccttggcacttctatggtccagatgttagagggtaagtattttgatgggggagatcgttg|\\
gactgtaattgaaagttatgtcttataatgaaatgtgttatataaagaagacctataaaa
cacttaggctgataaaacccccaaacgatgaagcctcacttttaccctctcttttaacag
              |3191
                      |3201
                               |3211
      |3181
                                        |3221
\tt ATGTTTTACTGTGCCTGGCTAACTATAGTCGAGGGGGTGATGGTCCTATGTGTCGCCCAG
 V L L C L A N Y S R G G D G P M C R P V
      11061
                                11071
      |3241 |3251 |3261
                                |3271 |3281 |3291
TAATTCTGTTGCCGGAAGATACCCCCCCCTTCTTAGAGCTTAAAGGATCACGCCATCCTT
 I L L P E D T P P F L E L K G S R H P C
      1081
                                 |1091
      |3301 |3311 |3321 |3331 |3341 |3351
GCATTACGAAGACTTTTTTTGGAGATGATTTTATTCCTAATGACATTCTAATAGGCTGTG
 I T K T F F G D D F I P N D I L I G C E
      |1101
                                |1111
      |3361 |3371 |3381
                                |3391
                                        |3401 |3411
\tt AGGAAGAGGAGCAGGAAAATGGCAAAGCCTATTGTGTGCTTGTTACTGGACCAAATATGG
 E E E Q E N G K A Y C V L V T G P N M G
      |1121
                                |1131
      |3421 |3431
                      . . . . . .
GGGGCAAGTCTACGCTTATGAGACAGqtaactqattcttaaaqttttqttatcaqaaaqt
 G K S T L M R O
      11141
  . . .
                        . . . . . .
```

 $\verb|catttgtgacattaggaataacatacttaggtgatcattttccaaacacagttacataaa| \\$

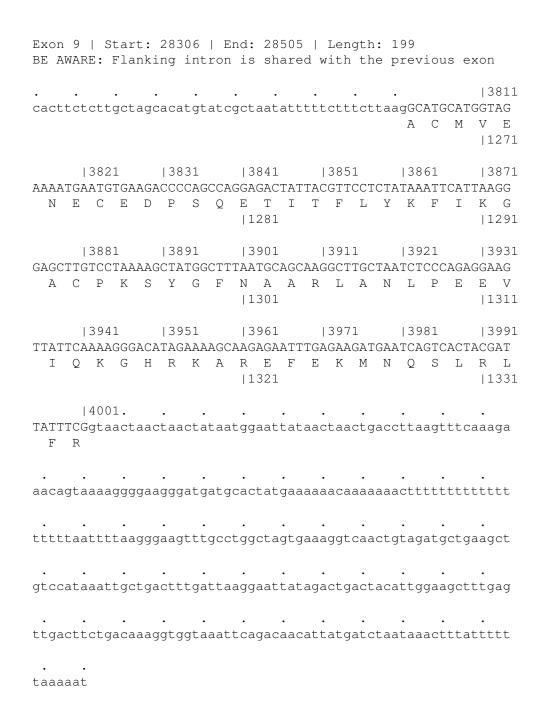
agtca	gccac	rtgact	taata	qqaaq	· caaaq	Iggaaa	ttact	ccctc	qtqtta	ıtaaaa	ttgag				
2	, ,	, ,		,,,,,				-	, 3		3 3				
· aatta	· tattt	agctg	aaaca	· tcgat		ıatgtt	• aaggg	gaat <i>a</i>	· atatgt	taaaa	• agggg				
								•							
aagga	aaggaggtcagtcattcaggtcatgaggccctttgacttgaattcatttcctcagaaggt														
	+ a++c				•										

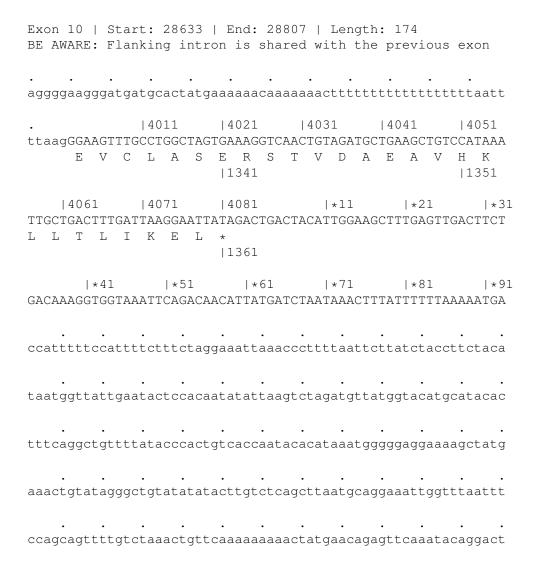




				t: 28 nking											ing	exo	n
att	ag	gcct	taca	ctgtt	tagt	tgt	aat	• aaa	aca	ttt	gtta	atac	tacç	1ggga	atga	• agaa	ca
cta	ıat	agga	· iggac	tcagg	gaagt	tta	tga	cct	tga	gcg	ata	ctgt	attt	tctt	taa	• naag	aa
acc	ctc	acto	· ccca	tgggc	tgct							aaac	• aagg	godta	attt	tata	.ga
ato	ıct	ttta	· igacg	tggat	gtac	ctaa	ccg	atg	ttg	ctt	ttc	tgtc	ctag	gcatt	ttt	.gtt	tt
aat	tc	cttt	:tttg	tttta	iatto	cctt	tga	gtt	act	tcc	tta	tgca	tatt	ttad	cttt	caac	ag
GAA		3651 GGTA		366 AACAT	1 TTG <i>F</i>											370 AACT	
F	}	G I	A	T F		G	Τ	А	Ι	A	N		V V 123	7 K 31	Ε	L	Α
	1	3711	-	372	1	1	373	1		37	41		375	51	I	376	1
CTO	SAG	ACTA	ATAAA.	ATGTC									TCAI	TAG	[AG	AAGA	ТТ
E		T I	K	C F		L	F	S	Τ	Н	Y		S I 125	. V 51	Ε	D	Y
		3771			1					38							
ATI	СТ	CAAA	ATGT	TGCTG	TGC	SCCT	AGG	ACA'	TAT	Ggt	atg	tgca	aatt	gttt	ttt	tcc	ac
S	5	Q N	1 V	A V		L	G	Н	М								

 ${\tt aaattcggttttttgagaggg}$





LRG Parser: Version: 1.1, Version Date: 11/02/2015

Reader: Version: 1, Version Date: 11/02/2015 Writer: Version: 1, Version Date: 11/02/2015 Control: Version: 1, Version Date: 11/02/2015