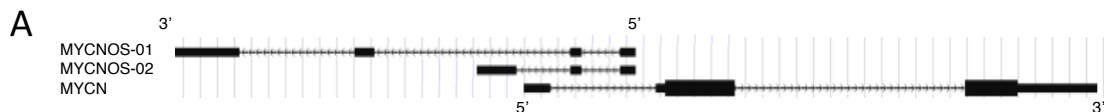


Figure S1



B

5'3' Frame 1
 RGWWRGSATLETAISFTQGTAWGRGLFLAAEF **Stop**LSRARRQPHSQRCGAGCSGEAPGSPTQLP
G Met GHPLEKNPSLWGRGTGSPTAQTQTDLLEPREFLFAEAITPPAP StopSPPSPPAEFSPSILSFPV
 QHSPPFV**Met IHAWEYLFSAHFRDRTVEDRSLPH Stop**PVTWVSFAPEGILTAQQCTCI **Stop**GWY
A Met KALSNCWHLDIGLLSLQKHEKYISFLYKLP SLCY StopVIATQNELRQVPIYYCV **Stop**APGWVW
 NRGKLYSRKELTVDQGRPTSE **Met D A S I Stop**SPTARIELYTAFG **Stop**LR **Stop**FLREENILVYKSP
 L Y **Stop**DTHTQRNNTI H **Met L R R A Q I Met T G Stop**ESNSL **Stop**Met **V F P D C E L K I C W V G Q A W W L T P V I L**
A L Q E S E E G R W L E P R I L R P A W A T W StopNPVSTKNIKISWAWWVRACGPSYLG **Stop**GRKIT **Stop**AQ
 GDQGCSPSSHCTPAW **Met T E Stop**DLVSKEKKKDLLSLSSGTRAQFLGSSHL PQNVYIFTD **Stop**I
I StopK

5'3' Frame 2
 GGGGEAPQLWKLPPFHSKALPGGGGCSWLQNSSSHEHADNRTRSGVGPAAQGGKQALRPSYREW
 GTLWRRTPAWGGDAPALRQLKHRQIF **Stop**SRGNFFSQKPLPPPPNRHQALLNSRLASCPFLSSTH
 LHL **Stop** **Stop**S **Met P G S I C L V L T S V T L W K T G S V S L T N L Stop**PGSVLLQRESLRHSSVPVYRDGTP **Stop**RP
 SAIAAGTLILD FSARN **Met R N T F L F F I N Y P V C A I R L Stop**QHK **Met N Stop**DKYPSITVSEPLGGFGTEV
 S **Stop**GTAARSSLWTRREGDPLKW **Met P Q S N H Q L L G Stop**SYTQHLVYN **Stop**GDF **Stop**ERRIF **Stop**FTN
 LLFTKTHTHKETTHTYT **Stop**EGPKS **Stop**LARRATASRFFFQIVN **Stop**RSAGLARHGGSRLL **Stop**SWH
 FKSLRRADGLSPGF **Stop**DRPGQHGETLSLQKI **Stop**KLARGGCVVVPATWEAEAGRSPEPREIKAA
 VCHHHTTALQPG **Stop**QSETLSQKKKKIY **Stop**VCLQGQEH SFS **Stop**DLHTSLK **Met C T F L L I K Stop**FKN

5'3' Frame 3
 GVVARLRNFGNCFIHTRHCLGEGAVPGCRILALTSTQTTALAAVWGRLLRGSRLSDPATGN GAP
 FGEEPQPGVGHRLSDSSNDRSSRAEGISFRRSHYSPRPLIATKPSC **Stop**ILA **Stop**HPVLSCPALTSI
 CNDPCLGVFV **Stop**CSLP **Stop**HCGRQAQSPSLTCDLGQFC SRGNPYGTAVYLYIG **Met V R H E G P Q Q L L**
A P StopYWTSQPPET **Stop**EIHFFSL **Stop**ITQSVLLGYSNTK **Stop**TETSTHLLLCLSPWVGLEQR **Stop**VKV
 QPQGAHCGPGKETHI **Stop**NGCLNLITNC **Stop**DRVIHSIWFIKVFIFERGEYFSLQISSLLRHTHTKKQH
 THTHVEKGNHDWLGEQQPLDGF SRL **Stop**IKDLLGWPG **Met V A H A C N P G T S R V Stop**GG
 Q **Met A Stop**AQDFETGLGN **Met V K P C L Y K K Y K N Stop**LGVVGA CLWSQLLGR LRQEDHLS PGRSRLQC
 AIITPLHSSLDDRVRPCLKRKKKRSTESVFRDKSTVSPRIFTPPSKCVHFY **Stop**LNNLKI

C

<u>No. of ATG from 5'end</u>	<u>Reliability</u>	<u>Frame</u>	<u>Identity to Kozak rule A/GXXATGG</u>	<u>Start (bp)</u>	<u>Finish (bp)</u>	<u>ORF Length (aa)</u>	<u>Stop codon found?</u>	<u>Sequence</u>
1	0.10	1	GXXATGG	49	180	44	Yes	MGHPLEKNPSLWGRGTGSPTAQTQTDLLEPREFLFAEAITPPAP
2	0.06	1	GXXATGa	259	333	25	Yes	MIHAWEYLFSAHFRDRTVEDRSLPH
3	0.06	2	tXXATGc	266	340	25	Yes	MPGSICLVLTSTVTLWKTGVS LTNL
5	0.06	1	GXXATGa	415	519	35	Yes	MKALSNCWHLDIGLLSLQKHEKYISFLYKLP SLCY
10	0.05	1	cXXATGt	820	849	10	Yes	MLRRAQIMTG