# Spot Translation Checks

Each spot check table has four rows

1. Original sequence
2. Hits to uniprot\_sprot (if any – used to find consensus frame (if any))
3. big\_blast\_translate.py translation(s)
4. Geneious translation(s)
5. Blast of translation(s) back against Uniprot

Contents

[With blast hit frame 2 1](#_Toc387679099)

[With blast hit frame -2 2](#_Toc387679100)

[With blast hit frame -3 3](#_Toc387679101)

[All 6 frames because blast hits in more than one frame 5](#_Toc387679102)

[All 6 frames because no blast hits 7](#_Toc387679103)

[With blast hit frame -2 8](#_Toc387679104)

## With blast hit frame 2

|  |
| --- |
| >140304CS1901028000001 CS1901028000001 DNA=(0%N 19%C 19%G 27%A 35%T ) Expr= 2=CEIE(2) Refseq=gi|27894290|ref|NM\_173074.1| Homo sapiens phosphatidylinositol glycan, class F (PIGF), transcript variant 2, mRN(eval= .00E+00) NRProtein=ref|NP\_001070542.1| hypothetical protein LOC768015 [Bos taurus] gb|AAI11282.1| Hypothetical protein MGC133893 [Bos taurus(eval= 3.00E-76)  GAGGGCAGTGAGTGGCACCTGATGGTGGGAGGTCGTAATCCTCCCTTCCATTCCGCGGGTGGGAGAGATAGATAGCCACT  CAAGAAAAACACCATGAAAGATACTGACATCAAGAAAACTACTGTATACCCATCTTTTATGCATATTTTCAATTATCCTA  AGCATCTTCATTCCATCATTCTTCTTGGAGAACTTCTCAATATTGGAGACACACTTGACATGGTTATGCATCTGTTCTGT  TTCTGTAACTGCCGTCAATTTAGTATTGTATTTAGTACTGAAACCAAATGTGTCTTCTAAAAGAAGTTCATTATCATACA  AGGTAACCAGATTTTTGAAATGCTGTATCTACTTTCTTACATCTTGTTTCTTCTTTCATGTAATTTTTGTTCTGTATGGC  GCACCACTAATAGAGTTGGCACTGGAAACATTTTCACTTGCAGTTATTTTGTCCACTTTTACTACTGTACCTTGTTTATG  TTTGTTAGGACCAAACTTCAAAGCCTGGCTAAGAATTTTCAGTAGAAATGGAGTTACGTCCGTTTGGGAGAATAGTCTTC  AGATCACTACAATTTCTAGTTTTGTAGGAATATGGC |
| 140304CS1901028000001 113 595 2 Q07326 Phosphatidylinositol-glycan biosynthesis class F protein OS=Homo sapiens GN=PIGF PE=1 SV=1  140304CS1901028000001 113 595 2 O09101 Phosphatidylinositol-glycan biosynthesis class F protein OS=Mus musculus GN=Pigf PE=1 SV=1 |
| >140304CS1901028000001\_F2\_98 Q07326 Phosphatidylinositol-glycan biosynthesis class F protein OS=Homo sapiens GN=PIGF PE=1 SV=1 : translation of 140304CS1901028000001 (translation method=using blast hit consensus frame 2) ; frame=2 ; start=98 ; end=595 ; molweight=18899.1 ; aromaticity=0.1506 ; instability\_index=29.988 ; gravy=0.9886 ; isoelectric\_point=9.1297 ; secondary\_structure\_fraction=( 0.5000, 0.2108, 0.2289)  KILTSRKLLYTHLLCIFSIILSIFIPSFFLENFSILETHLTWLCICSVSVTAVNLVLYLV  LKPNVSSKRSSLSYKVTRFLKCCIYFLTSCFFFHVIFVLYGAPLIELALETFSLAVILST  FTTVPCLCLLGPNFKAWLRIFSRNGVTSVWENSLQITTISSFVGIW |
|  |
| Database: UniProt only Swissprot  542,503 sequences; 192,888,369 total letters  Query= 140304CS1901028000001\_F2\_98  Length=166  Score E  Sequences producing significant alignments: (Bits) Value  sp|Q07326|PIGF\_HUMAN Phosphatidylinositol-glycan biosynthesis cl... 282 2e-95  sp|O09101|PIGF\_MOUSE Phosphatidylinositol-glycan biosynthesis cl... 265 2e-88  >sp|Q07326|PIGF\_HUMAN Phosphatidylinositol-glycan biosynthesis class F protein OS=Homo  sapiens GN=PIGF PE=1 SV=1  Length=219  Score = 282 bits (721), Expect = 2e-95, Method: Compositional matrix adjust.  Identities = 143/161 (89%), Positives = 150/161 (93%), Gaps = 0/161 (0%)  Query 6 RKLLYTHLLCIFSIILSIFIPSFFLENFSILETHLTWLCICSVSVTAVNLVLYLVLKPNV 65  ++LLYTHLLCIFSIILS+FIPS FLENFSILETHLTWLCICS VTAVNLVLYLV+KPN  Sbjct 7 KRLLYTHLLCIFSIILSVFIPSLFLENFSILETHLTWLCICSGFVTAVNLVLYLVVKPNT 66  Query 66 SSKRSSLSYKVTRFLKCCIYFLTSCFFFHVIFVLYGAPLIELALETFSLAVILSTFTTVP 125  SSKRSSLS+KVT FLKCCIYFL SCF FHVIFVLYGAPLIELALETF AVILSTFTTVP  Sbjct 67 SSKRSSLSHKVTGFLKCCIYFLMSCFSFHVIFVLYGAPLIELALETFLFAVILSTFTTVP 126  Query 126 CLCLLGPNFKAWLRIFSRNGVTSVWENSLQITTISSFVGIW 166  CLCLLGPN KAWLR+FSRNGVTS+WENSLQITTISSFVG W  Sbjct 127 CLCLLGPNLKAWLRVFSRNGVTSIWENSLQITTISSFVGAW 167 |

## With blast hit frame -2

|  |
| --- |
| >140304CS1901890300001 CS1901890300001 DNA=(0%N 31%C 37%G 15%A 17%T ) Expr= 2=CEVB(1) CEFE(1) (no refseq hit) NRProtein=ref|NP\_001070550.1| hypothetical protein LOC768023 [Bos taurus] gb|AAI11651.1| Hypothetical protein MGC137027 [Bos taurus(eval= 1.00E-41)  GCTGGGGACGTGGGGGAAGGTAGGCGCGGGAGGGGACGGGCGCTCACCCGAAGGCAGGCTGCGCTCCCGGTGGCCCATAG  CATCCGTGCCAGCGGGAATAAGCCTCACGGTACGGGCTGTCCGGAGCCCGTGGGAGCGTGTACCAAGCTTCCCGGGAGTC  CGAGTTGGTCAGGCCCGTGTAACATAGCTGGCCGGCCGCATCGCGTCCCATGGGCGTGTACTTCCAGCGCGTGGCCTGCT  TAATGGCCGGAGTCCAGCGGGGCCCCTCCTGGGACAGGTAGTCATCGCTGTGGACGGGGGAGGAGAGAGACTGAGCTTGC  AGCCCCCCTGCCCTCCCTGGGCCCCTCTATCCTCCTTGCTGGCATCGGGTG |
| 140304CS1901890300001 298 77 -2 Q2M2T2 Uncharacterized protein C19orf71 homolog OS=Bos taurus PE=2 SV=1  140304CS1901890300001 310 77 -2 A6NCJ1 Uncharacterized protein C19orf71 OS=Homo sapiens GN=C19orf71 PE=4 SV=2  140304CS1901890300001 310 77 -2 A5LFW8 Uncharacterized protein C19orf71 homolog OS=Macaca fascicularis GN=QtsA-14935 PE=2 SV=1  140304CS1901890300001 295 77 -2 A6H6Q4 Uncharacterized protein C19orf71 homolog OS=Mus musculus PE=2 SV=1 |
| >140304CS1901890300001\_R2\_346 Q2M2T2 Uncharacterized protein C19orf71 homolog OS=Bos taurus PE=2 SV=1 : translation of 140304CS1901890300001 (translation method=using blast hit consensus frame -2) ; frame=-2 ; start=346 ; end=47 ; molweight=10987.6 ; aromaticity=0.12 ; instability\_index=60.914 ; gravy=-0.927 ; isoelectric\_point=9.0671 ; secondary\_structure\_fraction=( 0.2000, 0.3300, 0.2100)  RGPGRAGGLQAQSLSSPVHSDDYLSQEGPRWTPAIKQATRWKYTPMGRDAAGQLCYTGLT  NSDSREAWYTLPRAPDSPYREAYSRWHGCYGPPGAQPAFG |
|  |
| Query= 140304CS1901890300001\_R2\_346  Length=100  Score E  Sequences producing significant alignments: (Bits) Value  sp|Q2M2T2|CS071\_BOVIN Uncharacterized protein C19orf71 homolog O... 149 1e-44  sp|A6NCJ1|CS071\_HUMAN Uncharacterized protein C19orf71 OS=Homo s... 130 3e-37  sp|A5LFW8|CS071\_MACFA Uncharacterized protein C19orf71 homolog O... 130 4e-37  sp|A6H6Q4|CS071\_MOUSE Uncharacterized protein C19orf71 homolog O... 117 3e-32  >sp|Q2M2T2|CS071\_BOVIN Uncharacterized protein C19orf71 homolog OS=Bos taurus PE=2 SV=1  Length=208  Score = 149 bits (376), Expect = 1e-44, Method: Compositional matrix adjust.  Identities = 71/75 (95%), Positives = 74/75 (99%), Gaps = 0/75 (0%)  Query 17 PVHSDDYLSQEGPRWTPAIKQATRWKYTPMGRDAAGQLCYTGLTNSDSREAWYTLPRAPD 76  P++SDDYLSQEGPRWTPAIKQATRWKYTPMGRDAAGQL YTGLTNSDSREAWYTLPRAPD  Sbjct 25 PLYSDDYLSQEGPRWTPAIKQATRWKYTPMGRDAAGQLWYTGLTNSDSREAWYTLPRAPD 84  Query 77 SPYREAYSRWHGCYG 91  SPYREAY+RWHGCYG  Sbjct 85 SPYREAYARWHGCYG 99 |

## With blast hit frame -3

|  |
| --- |
| >140304CS1901914000001 CS1901914000001 DNA=(0%N 30%C 33%G 16%A 21%T ) Expr= 2=CEVB(2) Refseq=gi|39780551|ref|NM\_018052.3| Homo sapiens Vac14 homolog (S. cerevisiae) (VAC14), mRN(eval= 3.00E-83) NRProtein=ref|NP\_001075913.1| hypothetical protein LOC533202 [Bos taurus] gb|AAI33601.1| Unknown (protein for MGC:148468) [Bos taurus(eval= 8.00E-70)  GCCGCAGGCCAGGTGGAAGAAGCAATTTCTGCCAGCACCTCCAGATCCTTTAGGATAACCTCGTCAGACTCGTCCGACAG  CGTCTGCAGGAGCACGGGGAACAGGCTGTCCGTGTGCCGCGACATCTTGCGGGGTGTCTTGATGTACAGGTGGTAGAGCC  ACTTGAGGACGGCGATCCTGGTCATCATGCCGATGGCCGTGTCGCTGAGGTGGCAGTTCAGCACCTGCACGATCCCGTCC  AGGTGGAGGGTGACTGGGGCCCTCTCTGCGCTGGCTGGAGTGAAGACACTGATGCCGCTACTGAAGCTAGAGTCGCAGGA  ACCATCAGGCCCTCCGCTGGCTGCCCCCTCCTGCGTGGCTGCGCAGTCCTCAAGGCTGGGGCCTGCCTGCTTCTGCACCG  CCGGCCTGGGCTCATCGGGCTCGTCATCTTCGGGGGTGACCAGCTTCATCAGGCTCTGGTTGCAGACGTTGGCTACCTCC  TTGATGTTCCTCTTGCGGTCGTCGTAGGCCAGGCAGGGCAGGACGGCGGTCA |
| # Query: 140304CS1901914000001 CS1901914000001 DNA=(0%N 30%C 33%G 16%A 21%T ) Expr= 2=CEVB(2) Refseq=gi|39780551|ref|NM\_018052.3| Homo sapiens Vac14 homolog (S. cerevisiae) (VAC14), mRN(eval= 3.00E-83) NRProtein=ref|NP\_001075913.1| hypothetical protein LOC533202 [Bos taurus] gb|AAI33601.1| Unknown (protein for MGC:148468) [Bos taurus(eval= 8.00E-70)  # Database: uniprot\_sprot.fa  # Fields: query id, q. start, q. end, query frame, subject acc., subject title  # 7 hits found  140304CS1901914000001 530 15 -3 A2VE70 Protein VAC14 homolog OS=Bos taurus GN=VAC14 PE=2 SV=1  140304CS1901914000001 530 15 -3 Q08AM6 Protein VAC14 homolog OS=Homo sapiens GN=VAC14 PE=1 SV=1  140304CS1901914000001 530 15 -3 Q80W92 Protein VAC14 homolog OS=Rattus norvegicus GN=Vac14 PE=1 SV=1  140304CS1901914000001 530 15 -3 Q80WQ2 Protein VAC14 homolog OS=Mus musculus GN=Vac14 PE=1 SV=1  140304CS1901914000001 530 15 -3 Q5ZIW5 Protein VAC14 homolog OS=Gallus gallus GN=VAC14 PE=2 SV=1  140304CS1901914000001 530 15 -3 Q66L58 Protein VAC14 homolog OS=Danio rerio GN=vac14 PE=2 SV=1  140304CS1901914000001 530 15 -3 Q68F38 Protein VAC14 homolog OS=Xenopus laevis GN=vac14 PE=2 SV=1 |
| >140304CS1901914000001\_R3\_530 A2VE70 Protein VAC14 homolog OS=Bos taurus GN=VAC14 PE=2 SV=1 : translation of 140304CS1901914000001 (translation method=using blast hit consensus frame -3) ; frame=-3 ; start=530 ; end=3 ; molweight=19051.82 ; aromaticity=0.0455 ; instability\_index=50.9324 ; gravy=-0.1449 ; isoelectric\_point=4.7342 ; secondary\_structure\_fraction=( 0.2727, 0.2330, 0.2898)  TAVLPCLAYDDRKRNIKEVANVCNQSLMKLVTPEDDEPDEPRPAVQKQAGPSLEDCAATQ  EGAASGGPDGSCDSSFSSGISVFTPASAERAPVTLHLDGIVQVLNCHLSDTAIGMMTRIA  VLKWLYHLYIKTPRKMSRHTDSLFPVLLQTLSDESDEVILKDLEVLAEIASSTWPA |
|  |
| Query= 140304CS1901914000001\_R3\_530  Length=176  Score E  Sequences producing significant alignments: (Bits) Value  sp|A2VE70|VAC14\_BOVIN Protein VAC14 homolog OS=Bos taurus GN=VAC... 345 5e-113  sp|Q08AM6|VAC14\_HUMAN Protein VAC14 homolog OS=Homo sapiens GN=V... 307 1e-98  sp|Q80W92|VAC14\_RAT Protein VAC14 homolog OS=Rattus norvegicus G... 301 3e-96  sp|Q80WQ2|VAC14\_MOUSE Protein VAC14 homolog OS=Mus musculus GN=V... 300 6e-96  sp|Q5ZIW5|VAC14\_CHICK Protein VAC14 homolog OS=Gallus gallus GN=... 249 2e-76  >sp|A2VE70|VAC14\_BOVIN Protein VAC14 homolog OS=Bos taurus GN=VAC14 PE=2 SV=1  Length=783  Score = 345 bits (885), Expect = 5e-113, Method: Compositional matrix adjust.  Identities = 168/172 (98%), Positives = 169/172 (98%), Gaps = 0/172 (0%)  Query 1 TAVLPCLAYDDRKRNIKEVANVCNQSLMKLVTPEDDEPDEPRPAVQKQAGPSLEDCAATQ 60  TAVLPCLAYDDRKRNIKEVA+VCNQSLMKLVTPEDDEPDEPRP VQKQAGPS EDCAA Q  Sbjct 302 TAVLPCLAYDDRKRNIKEVASVCNQSLMKLVTPEDDEPDEPRPVVQKQAGPSPEDCAAKQ 361  Query 61 EGAASGGPDGSCDSSFSSGISVFTPASAERAPVTLHLDGIVQVLNCHLSDTAIGMMTRIA 120  EGAASGGPDGSCDSSFSSGISVFTPASAERAPVTLHLDGIVQVLNCHLSDTAIGMMTRIA  Sbjct 362 EGAASGGPDGSCDSSFSSGISVFTPASAERAPVTLHLDGIVQVLNCHLSDTAIGMMTRIA 421  Query 121 VLKWLYHLYIKTPRKMSRHTDSLFPVLLQTLSDESDEVILKDLEVLAEIASS 172  VLKWLYHLYIKTPRKMSRHTDSLFPVLLQTLSDESDEVILKDLEVLAEIASS  Sbjct 422 VLKWLYHLYIKTPRKMSRHTDSLFPVLLQTLSDESDEVILKDLEVLAEIASS 473 |

## All 6 frames because blast hits in more than one frame

|  |
| --- |
| >140304CS1901892700001 CS1901892700001 DNA=(0%N 29%C 36%G 19%A 16%T ) Expr= 3=CEVB(3) Refseq=gi|89077243|ref|XM\_942373.1| PREDICTED: Homo sapiens similar to Rab GTPase binding effector protein 2 (Rabaptin-5beta) (LOC652743), mRN(eval= 2.00E-64) NRProtein=ref|XP\_884362.1| PREDICTED: similar to Rab GTPase-binding effector protein 2 (Rabaptin-5beta) isoform 6 [Bos taurus(eval= 6.00E-27)  GAGCTCACAGACCTGCTCTCGGAGCAGCGGGCAAAGATGCTGCGGCTGCAGGCGGAGCTGGAGACCAGTGAGCAGGTGCA  GCGGGATTTCGTACGACTGTCCCAGGCCCTGCAGGTGCGTCTGGAACGGATCCGCCAGGCAAGAGAGCCTGGAGCAAGTG  CGCGGCATCATAGACGAGGCGCCCCTTCGGGACGTCAGGGACATCAAGGACACCTGAGGGGCCGGCGCCTACCCACCCTG  GGGGAGACCGCCTTTCTCCCCTCCTGAACTGTGAAGCTGAGGCTTGCGGGGGGTGGTCTCAGGATGGAAGTCCGTCACCC  TTTCCAAGGCTGGGGT |
| # Query: 140304CS1901892700001 CS1901892700001 DNA=(0%N 29%C 36%G 19%A 16%T ) Expr= 3=CEVB(3) Refseq=gi|89077243|ref|XM\_942373.1| PREDICTED: Homo sapiens similar to Rab GTPase binding effector protein 2 (Rabaptin-5beta) (LOC652743), mRN(eval= 2.00E-64) NRProtein=ref|XP\_884362.1| PREDICTED: similar to Rab GTPase-binding effector protein 2 (Rabaptin-5beta) isoform 6 [Bos taurus(eval= 6.00E-27)  # Database: uniprot\_sprot.fa  # Fields: query id, q. start, q. end, query frame, subject acc., subject title  # 13 hits found  140304CS1901892700001 1 141 1 A4FUG8 Rab GTPase-binding effector protein 2 OS=Bos taurus GN=RABEP2 PE=2 SV=1  140304CS1901892700001 146 214 2 A4FUG8 Rab GTPase-binding effector protein 2 OS=Bos taurus GN=RABEP2 PE=2 SV=1  140304CS1901892700001 1 141 1 Q91WG2 Rab GTPase-binding effector protein 2 OS=Mus musculus GN=Rabep2 PE=2 SV=3  140304CS1901892700001 137 214 2 Q91WG2 Rab GTPase-binding effector protein 2 OS=Mus musculus GN=Rabep2 PE=2 SV=3  140304CS1901892700001 1 141 1 Q9H5N1 Rab GTPase-binding effector protein 2 OS=Homo sapiens GN=RABEP2 PE=1 SV=2  140304CS1901892700001 137 214 2 Q9H5N1 Rab GTPase-binding effector protein 2 OS=Homo sapiens GN=RABEP2 PE=1 SV=2  140304CS1901892700001 1 141 1 Q62835 Rab GTPase-binding effector protein 2 OS=Rattus norvegicus GN=Rabep2 PE=2 SV=1  140304CS1901892700001 137 214 2 Q62835 Rab GTPase-binding effector protein 2 OS=Rattus norvegicus GN=Rabep2 PE=2 SV=1  140304CS1901892700001 1 141 1 Q5RCR6 Rab GTPase-binding effector protein 2 OS=Pongo abelii GN=RABEP2 PE=2 SV=2  140304CS1901892700001 137 214 2 Q5RCR6 Rab GTPase-binding effector protein 2 OS=Pongo abelii GN=RABEP2 PE=2 SV=2  140304CS1901892700001 13 141 1 O35551 Rab GTPase-binding effector protein 1 OS=Mus musculus GN=Rabep1 PE=1 SV=2  140304CS1901892700001 13 141 1 O35550 Rab GTPase-binding effector protein 1 OS=Rattus norvegicus GN=Rabep1 PE=1 SV=1  140304CS1901892700001 13 141 1 Q15276 Rab GTPase-binding effector protein 1 OS=Homo sapiens GN=RABEP1 PE=1 SV=2 |
| >140304CS1901892700001\_F1\_1 translation of 140304CS1901892700001 (translation method=6 frames because blast hits in different frames) ; frame=1 ; start=1 ; end=264 ; molweight=10043.99 ; aromaticity=0.0227 ; instability\_index=61.75 ; gravy=-0.8989 ; isoelectric\_point=11.5646 ; secondary\_structure\_fraction=( 0.2159, 0.1932, 0.3409)  ELTDLLSEQRAKMLRLQAELETSEQVQRDFVRLSQALQVRLERIRQAREPGASARHHRRG  APSGRQGHQGHLRGRRLPTLGETAFLPS  >140304CS1901892700001\_F1\_268 translation of 140304CS1901892700001 (translation method=6 frames because blast hits in different frames) ; frame=1 ; start=268 ; end=336 ; molweight=2565.95 ; aromaticity=0.087 ; instability\_index=48.4522 ; gravy=-0.2826 ; isoelectric\_point=11.7114 ; secondary\_structure\_fraction=( 0.3043, 0.2609, 0.2609)  TVKLRLAGGGLRMEVRHPFQGWG  >140304CS1901892700001\_F2\_2 translation of 140304CS1901892700001 (translation method=6 frames because blast hits in different frames) ; frame=2 ; start=2 ; end=214 ; molweight=8021.69 ; aromaticity=0.0423 ; instability\_index=62.8028 ; gravy=-0.8324 ; isoelectric\_point=8.9631 ; secondary\_structure\_fraction=( 0.1831, 0.2958, 0.0986)  SSQTCSRSSGQRCCGCRRSWRPVSRCSGISYDCPRPCRCVWNGSARQESLEQVRGIIDEA  PLRDVRDIKDT  >140304CS1901892700001\_F3\_3 translation of 140304CS1901892700001 (translation method=6 frames because blast hits in different frames) ; frame=3 ; start=3 ; end=68 ; molweight=1932.95 ; aromaticity=0.0 ; instability\_index=17.2273 ; gravy=-0.2591 ; isoelectric\_point=6.7916 ; secondary\_structure\_fraction=( 0.0455, 0.2727, 0.4545)  AHRPALGAAGKDAAAAGGAGDQ  >140304CS1901892700001\_F3\_72 translation of 140304CS1901892700001 (translation method=6 frames because blast hits in different frames) ; frame=3 ; start=72 ; end=170 ; molweight=3101.26 ; aromaticity=0.0606 ; instability\_index=26.7939 ; gravy=-0.3212 ; isoelectric\_point=10.054 ; secondary\_structure\_fraction=( 0.0909, 0.3939, 0.2424)  AGAAGFRTTVPGPAGASGTDPPGKRAWSKCAAS  >140304CS1901892700001\_F3\_174 translation of 140304CS1901892700001 (translation method=6 frames because blast hits in different frames) ; frame=3 ; start=174 ; end=335 ; molweight=5563.82 ; aromaticity=0.0556 ; instability\_index=91.7074 ; gravy=-0.8556 ; isoelectric\_point=6.4376 ; secondary\_structure\_fraction=( 0.1296, 0.4630, 0.1852)  TRRPFGTSGTSRTPEGPAPTHPGGDRLSPLLNCEAEACGGWSQDGSPSPFPRLG  >140304CS1901892700001\_R1\_336 translation of 140304CS1901892700001 (translation method=6 frames because blast hits in different frames) ; frame=-1 ; start=336 ; end=232 ; molweight=3682.91 ; aromaticity=0.0286 ; instability\_index=38.48 ; gravy=-1.1086 ; isoelectric\_point=6.6268 ; secondary\_structure\_fraction=( 0.1429, 0.3714, 0.2000)  TPALERVTDFHPETTPRKPQLHSSGGEKGGLPQGG  >140304CS1901892700001\_R1\_228 translation of 140304CS1901892700001 (translation method=6 frames because blast hits in different frames) ; frame=-1 ; start=228 ; end=1 ; molweight=8425.63 ; aromaticity=0.0263 ; instability\_index=83.2605 ; gravy=-0.0329 ; isoelectric\_point=8.5611 ; secondary\_structure\_fraction=( 0.2763, 0.2500, 0.3421)  APAPQVSLMSLTSRRGASSMMPRTCSRLSCLADPFQTHLQGLGQSYEIPLHLLTGLQLRL  QPQHLCPLLREQVCEL  >140304CS1901892700001\_R3\_298 translation of 140304CS1901892700001 (translation method=6 frames because blast hits in different frames) ; frame=-3 ; start=298 ; end=8 ; molweight=9671.4 ; aromaticity=0.0412 ; instability\_index=74.0557 ; gravy=-0.2227 ; isoelectric\_point=9.9496 ; secondary\_structure\_fraction=( 0.2062, 0.3814, 0.2887)  DHPPQASASQFRRGERRSPPGWVGAGPSGVLDVPDVPKGRLVYDAAHLLQALLPGGSVPD  APAGPGTVVRNPAAPAHWSPAPPAAAASLPAAPRAGL  >140304CS1901892700001\_R2\_314 translation of 140304CS1901892700001 (translation method=6 frames because blast hits in different frames) ; frame=-2 ; start=314 ; end=207 ; molweight=4042.59 ; aromaticity=0.0278 ; instability\_index=102.7972 ; gravy=-0.6806 ; isoelectric\_point=12.2233 ; secondary\_structure\_fraction=( 0.2222, 0.3333, 0.1667)  RTSILRPPPASLSFTVQEGRKAVSPRVGRRRPLRCP  >140304CS1901892700001\_R2\_167 translation of 140304CS1901892700001 (translation method=6 frames because blast hits in different frames) ; frame=-2 ; start=167 ; end=3 ; molweight=6113.71 ; aromaticity=0.0545 ; instability\_index=96.6364 ; gravy=-0.3655 ; isoelectric\_point=11.3644 ; secondary\_structure\_fraction=( 0.1818, 0.2909, 0.1818)  CRALAPGSLAWRIRSRRTCRAWDSRTKSRCTCSLVSSSACSRSIFARCSESRSVS |
|  |
| Query= 140304CS1901892700001\_F1\_1 translation of 140304CS1901892700001  (translation method=6 frames because blast hits in different frames)  ; frame=1 ; start=1 ; end=264 ; molweight=10043.99 ;  aromaticity=0.0227 ; instability\_index=61.75 ; gravy=-0.8989 ;  isoelectric\_point=11.5646 ; secondary\_structure\_fraction=( 0.2159,  0.1932, 0.3409)  Length=88  Score E  Sequences producing significant alignments: (Bits) Value  sp|A4FUG8|RABE2\_BOVIN Rab GTPase-binding effector protein 2 OS=B... 87.0 8e-20  sp|Q91WG2|RABE2\_MOUSE Rab GTPase-binding effector protein 2 OS=M... 84.7 5e-19  sp|Q62835|RABE2\_RAT Rab GTPase-binding effector protein 2 OS=Rat... 82.8 2e-18  sp|Q9H5N1|RABE2\_HUMAN Rab GTPase-binding effector protein 2 OS=H... 82.8 3e-18  sp|Q5RCR6|RABE2\_PONAB Rab GTPase-binding effector protein 2 OS=P... 82.8 3e-18  >sp|A4FUG8|RABE2\_BOVIN Rab GTPase-binding effector protein 2 OS=Bos taurus GN=RABEP2  PE=2 SV=1  Length=585  Score = 87.0 bits (214), Expect = 8e-20, Method: Compositional matrix adjust.  Identities = 46/47 (98%), Positives = 47/47 (100%), Gaps = 0/47 (0%)  Query 1 ELTDLLSEQRAKMLRLQAELETSEQVQRDFVRLSQALQVRLERIRQA 47  +LTDLLSEQRAKMLRLQAELETSEQVQRDFVRLSQALQVRLERIRQA  Sbjct 515 QLTDLLSEQRAKMLRLQAELETSEQVQRDFVRLSQALQVRLERIRQA 561 |

## All 6 frames because no blast hits

|  |
| --- |
| >Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333  TTATTCTTTATTATTTATGCCTAACATGTGCTAGTGGGGTGCTGGGGACAGAGCAATGAA  CACAGCAGACAGAACTCCTGCTCTCTTTGTGCTTTTATGTAATTGTCATATGTGTAATGA  GGGAAAAGTTCTAGAGTTGTGACAGCATTATAAACGGGAGTCCTAATTTAGATTCGTGGT  CAGAGAAGGATCCCTGAGGAAGTGACATTTAAGCTTAGAACTGAAGGATCAGTTAGGATG  GGCCTGGAGGTTTTAGAGAGAGAAAGCAGCATCTATGCAGTCCTTAAATGTACAGAAGTT  TGGCATTTTCTAGAAATTGAAAGGCTATGTGGG |
| # Query: Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333  # Database: uniprot\_sprot.fa  # 0 hits found  # BLASTX 2.2.28+ |
| >Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333\_F1\_25 translation of Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333 (translation method=6 frames because no blast hits) ; frame=1 ; start=25  ; end=99 ; molweight=2794.07 ; aromaticity=0.08 ; instability\_index=54.536 ; gravy=-0.132 ; isoelectric\_point=6.8934 ; secondary\_structure\_fraction=( 0.2400, 0.3200, 0.2000)  HVLVGCWGQSNEHSRQNSCSLCAFM  >Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333\_F1\_172 translation of Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333 (translation method=6 frames because no blast hits) ; frame=1 ; start=17  2 ; end=333 ; molweight=6364.21 ; aromaticity=0.0741 ; instability\_index=64.7815 ; gravy=-0.2093 ; isoelectric\_point=6.7877 ; secondary\_structure\_fraction=( 0.3519, 0.1481, 0.3148)  IRGQRRIPEEVTFKLRTEGSVRMGLEVLERESSIYAVLKCTEVWHFLEIERLCG  >Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333\_F2\_35 translation of Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333 (translation method=6 frames because no blast hits) ; frame=2 ; start=35  ; end=139 ; molweight=3850.35 ; aromaticity=0.0571 ; instability\_index=4.3086 ; gravy=0.0857 ; isoelectric\_point=5.4833 ; secondary\_structure\_fraction=( 0.2571, 0.2000, 0.3714)  WGAGDRAMNTADRTPALFVLLCNCHMCNEGKVLEL  >Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333\_F3\_60 translation of Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333 (translation method=6 frames because no blast hits) ; frame=3 ; start=60  ; end=131 ; molweight=2877.45 ; aromaticity=0.125 ; instability\_index=26.2546 ; gravy=0.975 ; isoelectric\_point=5.8066 ; secondary\_structure\_fraction=( 0.5000, 0.0417, 0.2917)  TQQTELLLSLCFYVIVICVMREKF  >Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333\_F3\_135 translation of Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333 (translation method=6 frames because no blast hits) ; frame=3 ; start=13  5 ; end=194 ; molweight=2164.19 ; aromaticity=0.05 ; instability\_index=78.775 ; gravy=-0.905 ; isoelectric\_point=3.8387 ; secondary\_structure\_fraction=( 0.2000, 0.5000, 0.1000)  SCDSIINGSPNLDSWSEKDP  >Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333\_R1\_228 translation of Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333 (translation method=6 frames because no blast hits) ; frame=-1 ; start=2  28 ; end=166 ; molweight=2408.67 ; aromaticity=0.0952 ; instability\_index=32.381 ; gravy=-0.4238 ; isoelectric\_point=9.3034 ; secondary\_structure\_fraction=( 0.2857, 0.3333, 0.1905)  SFSSKLKCHFLRDPSLTTNLN  >Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333\_R1\_105 translation of Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333 (translation method=6 frames because no blast hits) ; frame=-1 ; start=1  05 ; end=34 ; molweight=2528.88 ; aromaticity=0.0417 ; instability\_index=67.3917 ; gravy=0.225 ; isoelectric\_point=11.1663 ; secondary\_structure\_fraction=( 0.2917, 0.2500, 0.2917)  QLHKSTKRAGVLSAVFIALSPAPH  >Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333\_R3\_208 translation of Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333 (translation method=6 frames because no blast hits) ; frame=-3 ; start=2  08 ; end=134 ; molweight=2763.13 ; aromaticity=0.08 ; instability\_index=70.004 ; gravy=0.152 ; isoelectric\_point=5.3027 ; secondary\_structure\_fraction=( 0.3200, 0.3600, 0.3200)  MSLPQGSFSDHESKLGLPFIMLSQL  >Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333\_R3\_94 translation of Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333 (translation method=6 frames because no blast hits) ; frame=-3 ; start=94  ; end=14 ; molweight=3013.4 ; aromaticity=0.0 ; instability\_index=80.3074 ; gravy=-0.5556 ; isoelectric\_point=9.5019 ; secondary\_structure\_fraction=( 0.1852, 0.2593, 0.1111)  KHKESRSSVCCVHCSVPSTPLAHVRHK  >Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333\_R2\_272 translation of Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333 (translation method=6 frames because no blast hits) ; frame=-2 ; start=2  72 ; end=183 ; molweight=3156.66 ; aromaticity=0.0 ; instability\_index=42.5267 ; gravy=0.7667 ; isoelectric\_point=8.498 ; secondary\_structure\_fraction=( 0.4000, 0.3667, 0.3000)  MLLSLSKTSRPILTDPSVLSLNVTSSGILL  >Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333\_R2\_167 translation of Locus\_278996\_Transcript\_1/1\_Confidence\_1.000\_Length\_333 (translation method=6 frames because no blast hits) ; frame=-2 ; start=1  67 ; end=24 ; molweight=5622.27 ; aromaticity=0.125 ; instability\_index=44.4896 ; gravy=-0.1833 ; isoelectric\_point=6.0268 ; secondary\_structure\_fraction=( 0.3333, 0.1667, 0.2292)  IRTPVYNAVTTLELFPHYTYDNYIKAQREQEFCLLCSLLCPQHPTSTC |
|  |
| No hits for any of the translations |

## With blast hit frame -2

|  |
| --- |
| >010608CEEF015155HT\_CS19 010608CEEF015155HT DNA=(0%N 13%C 17%G 25%A 44%T ) Expr= 1=CEEF(1) (no refseq hit) NRProtein=ref|XP\_537276.2| PREDICTED: similar to LINE-1 reverse transcriptase homolog isoform 1 [Canis familiaris(eval= 4.00E-20)  GGTCATGCTTTATCTTTGCCATATTAGTATAAAAAATTCTGAATTATTTTCCCATTGAGGGTCATAATTCATCCCTTCAG  TCAATTAAGTCTTCCAATCCATGCATAAGAGATTTTAAATTTTGTTTAGGTCTTTAATTTTGAAAGAAGTGTTTCATAGT  TTTGTGCATACAAGTTGTGTGCTTCCTTTATTCCTAAGTATTTTACTCCATTGTTTGCTATAGTAAATGGAATTTTTTCC  AATTCATTTTCTTTTTGGATGTTCATTGCTAGTGTGTAGAAATACAACTGATTTTTGCCTGTTGATTTTGTATCTTGAAA  CTGCTGAATTTATTGATTTGCATTAACAGTTTTTAAGAAAATGGATTCTTTAGGGTTTTAGAGCTAGAAGATTTTGTCAT  TTGTAAACAGAGATAATTTTACTTCTTGCTTTTGATTTAGACATTTTTTTTTCTTGCCTAATTGCCCTGGCTCTAATTTC  CAGTAGTGAAGTGAACATCTTATTTCTGATCTTAGGGGAAGAGCTTTCAGTTTTTAACCACTGAGTATGTTTGTGGGTTT  TTCATAAATGGCCTTTTTCATTTGGAGTAAGTTCCTTTCTATTCTTAGTTTATTGAATGTTTTTGTATCATAAAGGGGTA  TTTCATCAGTTGAGATGATCATGTAGTTTGATGTACTCTTT |
| 010608CEEF015155HT\_CS19 668 33 -2 O00370 LINE-1 retrotransposable element ORF2 protein OS=Homo sapiens PE=1 SV=1  010608CEEF015155HT\_CS19 668 33 -2 P08548 LINE-1 reverse transcriptase homolog OS=Nycticebus coucang PE=1 SV=1  010608CEEF015155HT\_CS19 668 33 -2 P11369 LINE-1 retrotransposable element ORF2 protein OS=Mus musculus GN=Pol PE=1 SV=2 |
| >010608CEEF015155HT\_CS19\_R2\_329 O00370 LINE-1 retrotransposable element ORF2 protein OS=Homo sapiens PE=1 SV=1 : translation of 010608CEEF015155HT\_CS19 (translation method=using blast hit consensus frame -2) ; frame=-2 ; start=329 ; end=129 ; molweight=7800.67 ; aromaticity=0.1194 ; instability\_index=16.8404 ; gravy=-0.6179 ; isoelectric\_point=8.9377 ; secondary\_structure\_fraction=( 0.3284, 0.1940, 0.2687)  IQQFQDTKSTGKNQLYFYTLAMNIQKENELEKIPFTIANNGVKYLGIKEAHNLYAQNYET  LLSKLKT |
|  |
| Query= 010608CEEF015155HT\_CS19\_R2\_329 O00370 LINE-1 retrotransposable  element ORF2 protein OS=Homo sapiens PE=1 SV=1 : translation of  010608CEEF015155HT\_CS19 (translation method=using blast hit  consensus frame -2) ; frame=-2 ; start=329 ; end=129 ;  molweight=7800.67 ; aromaticity=0.1194 ; instability\_index=16.8404 ;  gravy=-0.6179 ; isoelectric\_point=8.9377 ;  secondary\_structure\_fraction=( 0.3284, 0.1940, 0.2687)  Length=67  Score E  Sequences producing significant alignments: (Bits) Value  sp|P11369|LORF2\_MOUSE LINE-1 retrotransposable element ORF2 prot... 38.9 0.002  sp|O00370|LORF2\_HUMAN LINE-1 retrotransposable element ORF2 prot... 38.1 0.004  sp|P08548|LIN1\_NYCCO LINE-1 reverse transcriptase homolog OS=Nyc... 36.2 0.016  sp|Q7T297|F172A\_DANRE Protein FAM172A OS=Danio rerio GN=fam172a ... 29.3 2.0  sp|Q6DHK1|NNRE\_DANRE NAD(P)H-hydrate epimerase OS=Danio rerio GN... 28.9 3.2 |