

Figure S2

A

Exon 16

Y R A N T Y C S

Mmu ...agATACCGAGCCAACACCTACTGCTCC...

Hgla...agCTGT**TG**AGCCAACACCTATTGCTCC...

Fdam...agTTATCGAGCCAACACCTACTGCTCC...

Y R A N T Y C S

Exon 20

N L R W M A P E V F T

Mmu ...agAACCTGCGCTGGATGGCCCCGAGGTGTTTACA...

Hgla...agAACCTGCGCTGGATGGCCCGGG-GGTGTTTACG

Fdam...agAACCTGGGCTGGATGGCGCCGGAGGCGTTTACA...

N L G W M A P E A F T

Exon 21

I S S L L M R G W N A C

Mmu ...ATCTCATCTCTGCTGATGCGGGGCTGGAATGCATGT...

Hgla...ATATCCTCTCTGCTGATG**TG**AGGTTGGAATGCGTGT...

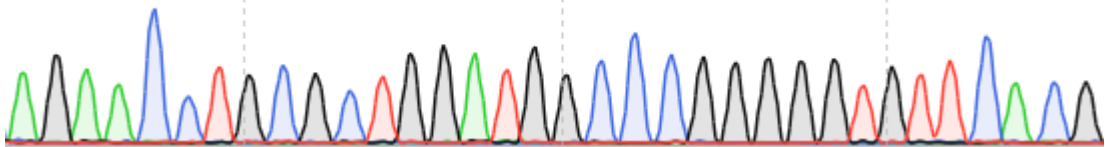
Fdam...ATATCTTCTCTGCTTATGCAATGTTGGAATGTATGT...

I S S L L M Q C W N V C

B

H. glaber exon 20

A G A A C C T G C G C T G G A T G G C C C G G G G G T G T T C A C G



H. glaber exon 21

A T A T C C T C T C T G C T G A T G T G A G G T T G G A A T G C G T G T

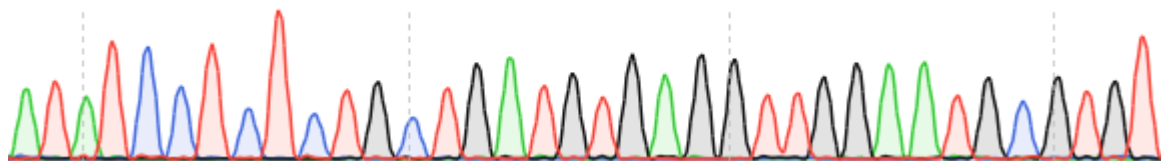


Figure S2 continued

C

ATGGGAAATTATAAATCTAGACCAACACAAACTTGTACTGATGAGTGGAAGAAGAAAAGTCAATGAATCTT
ATGTTATTACAATAGAAAAGATCAGAAGATGACCCGAGGATCAAGGAAAAAGAATTTACAGAGATGAGACA
TATTTTTGGCTCTGATGAAGCCTTCAGTAAAGTCAACCTAAATTACCGCACTGAAAGTGGACTATCTCTA
CTGCATTTATGTTGCATTTGTGGAGGCAACCAGTCTCATATTCGAGCCCTTATGTTGAAAGGACTCTGCC
CATCTCTACTGACAAGAAATGGATTACAGCGTTGCACCTAGCAGTTTATAAGGACAGTGCAGAACTGAT
CACTTCCTTCCTCCACAGTGGGGGTGATGTGCAGCAGGTGGGGTTCGGTGGCCTCACTGCCTCCACATT
GCTGTCATCGTGGTACCTGGAGGCCCTGATGTGTTACTCCAACACGGAGCTAATGTCAATGTTCAAG
ATGCTGTTTTTTTTACTCCATTGCACATTACAGTGTACTATGGGCATGAGCAGGTAACCCATCTCCTTTT
GAAATTTGGTGTGCTGATGTAAATGTAAGTGGTGAAGTTGGGGATAGGCCTCTCCATCTGGCATCTGCAAAG
GGATTCCTCAACATTGCAAAACTCCTGATGGAAGATGGGAATAAAGCAGATGTGAATGCTCAGGATAATG
AAGACCATGACCCACTGCATTTCTGTTCTCATTTTGGCCATCACGAGATAATGAAGTATCTGCTCCAAAG
TGACTCGGAAGTTCAGCCTCATGTTGTTAATATCTATGAAGACACTCCTTTCAGTTCGGCATGCTACAAT
GGCAAATTTGAAGTTCGTAAGGAGATCATTCAAATAACTGGAACAGAAAAGTTTGACTAAGGAAAACATCT
TTAGTGAGACAGCTTTTTCAGTGTCTGTACCTATGGCAAGAGCATTGATCTGGTTAAATTTCTTCTTGA
TCAGAATGTGGTAAACATCAGTCACCAAGGAAGAGATGGGCACACAGGATTACACTCTGCTTGCTACCAT
GGTCACATTACCTGGTTCAGTTCCTACTGGACAATGGGGCTGATATGAATCTAGTTGCTTGTGATCCCA
GCAGGTCTAGTGGTGAAAAAAATGAGCAGACATGTTTGTGATGTGGGCTTATGAAAAAGGACAGGATGCCAT
TGTCACACTTCTGAAGCATTATAAGAGACCCAGGATGAACTGCCCTGTAATGAATATTCTCGTCTGGGA
GGAGATGGCTCCTATGTGTCTGTTCCATCCCCTTTGGGGAAGATTAAAAACATAACAAAAGAGAAGGTGG
ATGTTCTCCTCCTGAGGGCCAGATTGCCTTCACATTTCCATCTTCAGCTCTCGGAAATTGAGTTCATGA
GATTATCGGCTCAGGTTCTTTTGGGAAAGTATACAAAGGACGATGCAGAAAATAAATAGTGGCTATAAAA
CACTGTTGAGCCAACACCTATTGCTCCAAATCAGATATGGATATGTTTTGCCGAGAGGTGTCCATTCTCT
GCCAGTCAATCATCCCTGTGTCAATTCAGTTTGTGGGCACTTGCTTAAATGATCCCAGCCAGTTTGGCAT
TGTCACCCAGTACATCTCAGGAGGTTCTCTGTTCTCCCTCCTTCATGAGCAGAAAAGGGCCCTTGATTTG
GAGTCTAAGTTAATTATTGCAGTAGATATTGCCAAAGGCATGGAGTACCTTCACCACCTGACCCAGCCAA
TTATACACCGTGACTTGAACAGTCACAATATTCTTCTCTATGAGGACGGGCATGCTGTGGTGGCCGATTT
TGGAGAATCAAGATTTCTACAATCTCTGGATGAAGACAACATGACAAAGCAACCTGGGAACCTGCGCTGG
ATGGCCCGGGGGTGTTCACGCAGTGCGCCGCTACACCCTAAGGCCGAAGTCTTCAGCTGCTCGCTGTG
TCTGGGGGAGCTGCTGACTGGTGTGAGATTCCCTTCGCTCATCTCAAGCAGTGGCAGCGGCAGCAGACATGG
CCTCCACCACATCAGGCATCCCATTGGATATTCCATTCCCAAGCCATATCCTCTCTGCTGATGTGAGG
TTGGAATGCGTGTCTGAAGGAAGACCTGAGTTTTCCAAAGTGTACCAAGTTAGAAGAATGTCTCTGC
AACATTGAGCTGATGTCTCCGGCATCGAGTAACAGCAGTGGGTCTCTTTCACCTTCCTCTTCTTCTGATT
GCCTGGCAAGCCAGGGAGCCCCTGGCCAGAGCACGTGGCAGCTTTACGGAGCCGTTTTGAGTTGGAGTGT
GCTCTCAACGCAAGGTCTATGCTGCCTGGCCCCAAAGCACTGGACCAAGTTCTCCTCAGGGGCTGTCTT
TGGAGGAGATGTGCAGGAGTCTCCAGTGTCCACCATCGACAAATACGGCTACATGTCTGATCCCCTGAG
CCCCATGCATTTCTTCTTGGCCGGGATAGTGGCAGCTTCAAGGACAGCAGCTGA

MGNYKSRPTQTCTDEWKKKVNESYVITIERSEDDPRIKEKEFTEMRHFSGSDEAFSKVNLNRYRTESGLSL
LHLCCICGGNQSHIRALMLKGLCPSLLTRNGFTALHLAVYKDSAELITSFLHSGGDVQQVGFGLTALHI
AVIAGHLEAPDVLQHGANNVNVQDAVFFTPLHITVYYGHEQVTHLLKFGADVNSGEVGDRLHLASAK
GFFNIAKLLMEDGNKADVNAQDNEDHDPLHFCSHFGHHEIMKYLLQSDSEVQPHVNNIYEDTPLHLACYN
GKFEVAKEIIQITGTESLTKENIFSETAFHSACTYGKSIDLVKFLLDQNVVNIHQGRDGHGTGLHSACYH
GHIHLVQFLLDNGADMNLVACDPSRSSGKEQNTCLMWAYEKQDAIVTLLKHYKRPQDELPCNEYSRPG
GDGSYVSVSPPLGKIKNITKEKVDVLLLRARLPSHFHLQLSEIEFHEIIGSGSFGKVYKGRNRNKIVA
HC*

Figure S2 continued

D

Identities: 431/490 (88%)

Positives: 453/490 (92%)

Gaps: 1/490 (0%)

| | | | |
|------|-----|--|-----|
| Hgla | 1 | MGNYSRPTQTCTDEWKKKVNESYVITIERSEDDPRIKEKEFTEMRHIFGSDEAFSKVNL | 60 |
| | | MGNYSRPTQTC+DEWKKKV+ESY I IER EDD +IKE EF E+RHIFGSDEAFS+V+L | |
| Mmu | 1 | MGNYSRPTQTCSEWKKKVSESYAIIIERLEDDLQIKENEFQELRHIFGSDEAFSEVSL | 60 |
| Hgla | 61 | NYRTEGLSLLHLCCICGGNQSHIRALMLKGLCPSLLTRNGFTALHLAVYKDSAEELITSF | 120 |
| | | NYRTE GLSLLHLCC CCGN+SHIRALMLKGL PS LTRNGF ALHLAVYKDS ELITS | |
| Mmu | 61 | NYRTERGLSLLHLCCACCGNKSHIRALMLKGLRPSRLTRNGFPALHLAVYKDSLEELITSL | 120 |
| Hgla | 121 | LHSGGDVQQVGFGLTALHIAVIAGHLEAPDVLLQHGANVNVQDAVFFFTPLHITVYYGHE | 180 |
| | | LHSG DVQQ G+GGLTALHIA IAGH EA +VLLQHGANVNVQDAVFFFTPLHI YYGHE | |
| Mmu | 121 | LHSGADVQQAGYGGTALHIAAIAAGHPEAVEVLLQHGANVNVQDAVFFFTPLHIAAAYGHE | 180 |
| Hgla | 181 | QVTHLLLKFGADVNVSGEVGDRPLHLASAKGFFNIAKLLMEDGNKADVNAQDNEDHDPLH | 240 |
| | | QVT +LLKFGADVNVSGEVGDRPLHLASAKGFFNI KLL+E GNKADVNAQDNEDH PLH | |
| Mmu | 181 | QVTSVLLKFGADVNVSGEVGDRPLHLASAKGFFNIVKLLVE-GNKADVNAQDNEDHVPLH | 239 |
| Hgla | 241 | FCSHFGHHEIMKYLLQSDSEVQPHVNIYEDTPLHLACYNGKFEVAKEIIQITGTESLTK | 300 |
| | | FCS FGHH I+ YLLQSD EVQPHV+NIY DTPLHLACYNG FEVAKEI+ +GTESLTK | |
| Mmu | 240 | FCSRFGHHNIVSYLLQSDLEVQPHVINIYGDTPHLHLACYNGNFEVAKEIVHVVTGTESLTK | 299 |
| Hgla | 301 | ENIFSETAFHSACTYGKSIDLVKFLLDQNVVNIHQGRDGHTGLHSACYHGH IHLVQFLL | 360 |
| | | ENIFSETAFHSACTYGK+IDLVKFLLDQN VNI+H+GRDGHTGLHSACYHGH I LVQFLL | |
| Mmu | 300 | ENIFSETAFHSACTYGKNIDLVKFLLDQNAVNIHNRGRDGHTGLHSACYHGHIRLVQFLL | 359 |
| Hgla | 361 | DNGADMNLVACDPSRSSGEKNEQTCLMWAYEKGDIVTLLKHYKRPQDELPCNEYSRPG | 420 |
| | | DNGADMNLVACDPSRSSGEK+EQTCLMWAYEKGDIVTLLKHYKRPQDELPCNEYS+PG | |
| Mmu | 360 | DNGADMNLVACDPSRSSGEKDEQTCLMWAYEKGHDAIVTLLKHYKRPQDELPCNEYSQPG | 419 |
| Hgla | 421 | GDGSYVSVPSPLGKIKNITKEKVDVLLLRARLPSHFHLQLSEIEFHEIIGSGSFGKVYKG | 480 |
| | | GDGSYVSVPSPLGKIK++TKEK DVLLLR LPS FHLQLSEIEFHEIIGSGSFGKVYKG | |
| Mmu | 420 | GDGSYVSVPSPLGKIKSMTKEKADVLLLRARLPSRFHLQLSEIEFHEIIGSGSFGKVYKG | 479 |
| Hgla | 481 | RCRNKIVAIAK 490 | |
| | | RCRNKIVAIAK | |
| Mmu | 480 | RCRNKIVAIAK 489 | |

Figure S2 continued

E

ATGGGAAATTATAAATCTAGACCAACACAAACTTGTACTGATGAGTGGAAGAAGAAAGTCAATGAATCTT
ATGTTATTACAATAGAAAGATCAGAAGATGACCCGAGGATCAAGGAAAAAGAATTTACAGAGATGAGACA
TATTTTTGGGCCCTGATGTGTTACTCCAACACGGAGCTAATGTCAGTGTTCAAGATGCTGTTTTTTTCA
CTCCATTGCACATTACAGTGTACTATGGGCATGAGCAGGTAACCCATCTCCTTTTTGAAATTTGGTGCTGA
TGTAATGTAAGTGGTGAAGTTGGGGATAGGCCTCTCCATCTGGCATCTGCAAAGGGATTCTTCAACATT
GCAAACTCCTGATGGAAGATGGGAATAAAGCAGATGTGAATGCTCAGGATAATGAAGACCATGACCCAC
TGCATTTCTGTTCTCATTTTGGCCATCACGAGATAATGAAGTATCTGCTCCAAAGTGACTIONCGGAAGTTCA
GCCTCATGTTGTTAATATCTATGAAGACACTCCTTTGCACTTGGCATGCTACAATGGCAAATTTGAAGTT
GCTAAGGAGATCATTCAAATAACTGGAACAGAAAGTTTACTAAGGAAAACATCTTTAGTGAGACAGCTT
TTCACAGTGCTTGTACCTATGGCAAGAGCATTGATCTGGTTAAATTTCTTCTTGATCAGAATGTGGTAAA
CATCAGTCACCAAGGAAGAGATGGGCACACAGGATTACACTCTGCTTGCTACCATGGTCACATTCACCTG
GTTTCAGTTCTTACTGGACAATGGGGCTGATATGAATCTAGTTGCTTGTGATCCCAGCAGGTCTAGTGGTG
AAAAAATGAGCAGACATGTTTGATGTGGGCTTATGAAAAGGACAGGATGCCATTGTCACACTTCTGAA
GCATTATAAGAGACCCCAGGATGAACTGCCCTGTAATGAATATTCTCGTCCTGGAGGAGGTGCCCACTTT
TGCATTCACCTTTTCATTGTTGTGATGCATGGCACCTAAAGAGAAGGTGGATGTTCTCCTCCTGAGGGCCA
GATTGCCTTCACATTTCCATCTTCAGCTCTCGGAAATTGAGTTCATGAGATTATCGGCTCAGGTAACCT
AAAACACCACAGGTCTGCTTTAGTTTCATGAACCCGAGATAATCTGTGTTTTATTTATTTTAAACAATTTT
AGGTTCTTTTTGGGAAAGTATACAAAGGACGATGCAGAAATAAAATAGTGGCTATAAAACACTGTTGA

MGNYSRPTQTCTDEWKKKVNESYVITIERSEDDPRIKEKEFTEMRHIFGPLMCYSNTELMVFKMLFFS
LHCTLQCTMGMSR*

Figure S2 continued

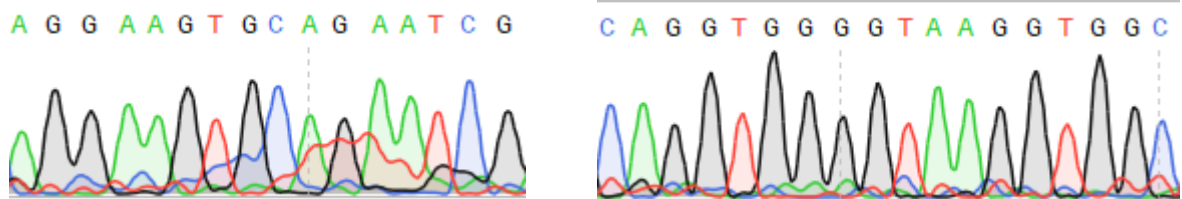
F

Exon 5

```
          D S L E L I T S L L H S G A D V Q
Mmu ...agGATAGCCTGGAACCTTACTTCACTGTTGCATAGCGGAGCAGATGTTTCAG
Fdam...agGA-AGTGCAGAATCGATCACTTCCCTGCTGCACAGTGGGACTGATGAGCAG
Hgla...agGACAGTGCAGAACTGATCACTTCCCTCCACAGTGGGGGTGATGTGCAG
          D S A E L I T S F L H S G G D V Q

          Q A G Y G G L T A L H I A A I A
Mmu      CAAGCGGGATATGGCGGCCTCACAGCCCTCCACATCGCTGCAATAGCT...
Fdam     CAGGTGGGGTAAGGTGGCTTCTCTGCGCTCCACACTGCTGTGCATAGCC...
Hgla     CAGGTGGGGTTCGGTGGCCTCACTGCACCTCCACATTGCTGTGCATCGCT...
          Q V G F G G L T A L H I A V I A
```

F. damarensis exon 5



Exon 15

```
          S F G K V Y K G R C R N K I V A I
Mmu ...agGTTCTTTTGGG AAAGTCTATAAAGGGCGATGCAGAAATAAAATAGTGGCAATC...
Fdam...agGTTCTTTTAGGGGAAGGTAAATAAAGGATGATGCAGGAATAAGATAGCGGCTATA...
Hgla...agGTTCTTTTGGG AAAGTATACAAAGGACGATGCAGAAATAAAATAGTGGCTATA...
          S F G K V Y K G R C R N K I V A I
```

F. damarensis exon 15

