

Sequences alignment:

>[Human] **KNOP1**, NM_001348527.2 Homo sapiens lysine rich nucleolar protein 1 (KNOP1), transcript variant 1, mRNA

> - 6677 nucleotides

```
1      gactcgaagtggacttccggggtcacggcgggagctgggctctcacgtggaggcggggaaatt
61     tcgcccaccgggtgagtgccagggcgcggttcgaggcgggcccggcgcggggtgtggagattg
121    caccagaggaagtgcgcggcggggaagatggggaaatgcgggtccaccacatccctgaat
1      M·G·K·C·G·P·P·T·S·L·N·

181    ccgaccctccagtcggtgtgcgcccgcctccggcctgaggccgcccgcctccagaaagttt
12     ·P·T·P·P·V·G·V·R·P·L·R·P·E·A·A·A·S·R·K·F·

241    tcccgcctcagagccggccctgggaccccacgtaatccctcccagtcctggtgatgaggag
32     ·S·R·L·R·A·G·P·G·T·P·R·N·P·S·Q·S·V·D·A·E·

301    gcttgggcccattgggagccttggggaatgatcaccaagacacacaaagtagaccttggg
52     ·A·W·A·H·G·S·L·G·G·M·I·T·K·T·H·K·V·D·L·G·
Ex1 | Ex2
Start of
Isoform c and B

361    ctcccagagaagaaaaagaagaagaagtggtcaaagaaccagagactcgataactcagtt
72     ·L·P·E·K·K·K·K·K·K·V·V·K·E·P·E·T·R·Y·S·V·

421    ttaaacaatgatgattactttgctgatgtttctcctttaagagctacatccccctctaag
92     ·L·N·N·D·D·Y·F·A·D·V·S·P·L·R·A·T·S·P·S·K·

481    agtgtggcccatgggacggcacctgagatgcctctagtgaagaaaaagaagaaaaag
112    ·S·V·A·H·G·Q·A·P·E·M·P·L·V·K·K·K·K·K·K·K·K·

541    aagggtgtcagcaccctttgagaggagcatgtagaacctgagaccacgctgacctgctaga
132    ·K·G·V·S·T·L·C·E·E·H·V·E·P·E·T·T·L·P·A·R·

601    cggacagagaagtcacccagcctcaggaagcaggtgtttggccacttggagttcctcagt
152    ·R·T·E·K·S·P·S·L·R·K·Q·V·F·G·H·L·E·F·L·S·

661    ggggaaaaagaaaaataagaagtcacctctagccatgtcccatgcctctggggtgaaaacc
172    ·G·E·K·K·N·K·K·S·P·L·A·M·S·H·A·S·G·V·K·T·

721    tccccagaccctagacaggggtgagggaaaccagagttggcaagaagctcaaaaaacac
192    ·S·P·D·P·R·Q·G·E·E·E·T·R·V·G·K·K·L·K·K·H·
Ex2 | Ex3

781    aagaaggaaaaaaagggggcccaggacccccacagccttctcgggtccaggacccttggttc
212    ·K·K·E·K·K·G·A·Q·D·P·T·A·F·S·V·Q·D·P·W·F·

841    tgtgaggccagggaggccagggatggtggggacacttgctcagtggggaagaaggatgag
232    ·C·E·A·R·E·A·R·D·V·G·D·T·C·S·V·G·K·K·D·E·

901    gaacaggcagccttggggcagaaacggaagcggaagagccccagagaacacaatgggaag
252    ·E·Q·A·A·L·G·Q·K·R·K·R·K·S·P·R·E·H·N·G·K·

961    gtgaagaagaaaaaaaaaatccaccaggaggagatgcctcccaggccactccaagccc
272    ·V·K·K·K·K·K·I·H·Q·E·G·D·A·L·P·G·H·S·K·P·
```

1021 tccaggtccatggagagcagccctaggaaggaagtaaaaagaagccagtc aaagttgag
292 ·S·R·S·M·E·S·S·P·R·K·G·S·K·K·K·P·V·K·V·E·

1081 gctccggaatacatcccataagtgatgaccctaaggcctccgcaaagaaaagatgaag
312 ·A·P·E·Y·I·P·I·S·D·D·P·K·A·S·A·K·K·K·M·K·

1141 tccaaaaagaaggtagagcagccagtc atcgaggagccagctctgaaaaggaagaaaaag
332 ·S·K·K·K·V·E·Q·P·V·I·E·E·P·A·L·K·R·K·K·K·

1201 aagaagaggaaagagagtggtggtagcaggagacccttggaaggaggaaacagacacggac
352 ·K·K·R·K·E·S·G·V·A·G·D·P·W·K·E·E·T·D·T·D·

Ex3 | Ex4
Region DUF5594

1261 ttagaggtggtggttggaaaaaaaggcaacatggatgaggcgccacatagaccagggtgagg
372 L·E·V·V·L·E·K·K·G·N·M·D·E·A·H·I·D·Q·V·R

Ex4 | Ex5

1321 cgaaaggccttgcaagaagagatcgatcgcgagtcaggcaaaacggaagcttctgaaacc
392 R·K·A·L·Q·E·E·I·D·R·E·S·G·K·T·E·A·S·E·T·

Ex5 | Ex6

1381 aggaagtggacgggaacccagtttggccagtggtggatactgctggttttgagaacgaggac
412 ·R·K·W·T·G·T·Q·F·G·Q·W·D·T·A·G·F·E·N·E·D·

Region SMAP

1441 caaaaactgaaatctcagacttatgggtggcttcaaaaacctgtccccttcggttcagc
432 Q·K·L·K·F·L·R·L·M·G·G·F·K·N·L·S·P·S·F·S·

1501 cgccccgcagcagcagattgcaaggccaacatggcctcggcaagaaggcggctgacagc
452 R·P·A·S·T·I·A·R·P·N·M·A·L·G·K·K·A·A·D·S·

1561 ctgcagcagaatctgcagcgggactacgaccgggcatgagctggaagtacagccgggga
472 L·Q·Q·N·L·Q·R·D·Y·D·R·A·M·S·W·K·Y·S·R·G·

1621 gccggcctcggcttctccaccgcccccaacaagatcttttacattgacaggaacgcttcc
492 A·G·L·G·F·S·T·A·P·N·K·I·F·Y·I·D·R·N·A·S·

1681 aagtcagtc aagctggaagat taaactctagagttttgtcccccaaaactgccacaatt
512 ·K·S·V·K·L·E·D·*

1741 gctttgattattccatttatgctggagattacaaatTTTTTTTgtgaaaaaatcagatct
1801 tgggtgaggacctcgagcagtaagatataaaataactcccataagcttagcgttccagtaat
1861 ggaacactaggcataaatggttttattcagttgtgcaaatgaaagccatctgacagttggc
1921 tcacattgaacacctgtggagattaaggacgaggacaactatattgatgggcttggatga
1981 actggggcagggcagctcatatttcggggagccaggagaacgagtgagtgctaaaacctcc
2041 tgttttctgtgttaaaccattccgtccctgtttgagacatcagtatgtacagttactttt
2101 gttgagtgtttagcaggtactagggacatactagtgtttcccttaattgtatttaactctc
2161 ataattatgaaatgggtgctattattagccccatcttatagatgaggcaactgaggttca
2221 gggataaagt aataaa attgcctggggtcaccagccacta agtgtcagggtgttgactt
2281 ttgtaccogaagccctaagttcactattcgccactctgaatgtccccttttagggaatttc
2341 caccagaatcctcgttggggattgaaatgtcttttagatggaggaaaaagttttatgacaa
2401 gtctgcatctctgataaaaagtggagtgaatgaggaaacggagaatcgcaaactccttttc
2461 ctctccttttctttccctgtcatagagcagctgttaggcagagaggtgtctgagttgta
2521 ccaaacacatgtgactgagctgctgctgctgggaaatgtgcctctaaaccagtcctgcct
2581 tttccactgggtgtcccttcggccctaagtgtgaggtaggattttttgtgctttgactgga
2641 ttctgaacaaggcatttttaggaagagtggttagtcccgaaatgaatcaatggcagatcc
2701 cagggacgtgtgtgtgtgtttttgagataggctctcactctgtcaccaggttgagtgta
2761 gtggcatatcacggcttgttgcaaccactacctatccagcttaagtgatcctcccacctc
2821 agcctcccagtagctgggactgtgggtgtgcaccacaacaaccagctagtgtttaaaat
2881 tttttgtagagatggggtctcactatattgtccaggctagtcccaaactcctggactcaa
2941 ctcatgctcctgcctcagcccccaagagctgggatcacaggtgtgagccactgtgcccg

polyA signal sequence
major polyA site

3001 cccaggggatgtgtctgatgagattcttttttatcacaaggaagagaatcccaggtgact
3061 tctgtaggtggatattgaatgccaggaggagatgaacgtgtggagatgacacgggaagtt
3121 gcctcatcggctgtggggacagcgtttccaagcctggactctacagcaagcctcaaagag
3181 ctccaggtccagtgggcccaagtcgcatgtgaccatctgtgaccacctctttccactcac
3241 ctagctttttgcatccaaaatacaggttctgactgggacattttcacgatcgtttggcc
3301 tgaacatgcctcctacagacatcccagcccaggcaggcctcgatggccagaaaagttcagt
3361 ctcacgtggcaggaaggcctgattagggcctggttcaacagctcgccaaggaagcattct
3421 ctccatcatctccagtggttggcttcatcaaggctagtacaaagatggcagtaacagttcc
3481 aggcagcatatcaagaccctagaacattccagaggaagaaaagcctttttgtgtctcaag
3541 gaaacttccagaagcctccattgctgggcaatgagtcattctgccatagctgcacccgcc
3601 cctgtcatgggaaggctgggtgattcccttcagtcagaagagtgatgctgggtgaagtgg
3661 ccattggtgtccacagctgcccattgggagaggtgaggcaggcacttcaggcagtttccc
3721 tggctctgaaccctgatccgcaagcattggccaggtcaatagttggacaagtatgtgcta
3781 caaggacctacttggggccacttccctcaggacgtgaacttgtcctgcctttcttctctgg
3841 attcaaccttgagatacataatggggagaggcgtggaaagcccaagaggcagtgttctgg
3901 atttgagtgagtgaccttctgggacacctcctcctttgatggagaaaacaaatccctcctgac
3961 cttgagcaacactacaggaatgtgctggtgtaaacaggaaccagcgttctgtttctttt
4021 tttttgagatggagtctcacggctcgcgcaggctggagtggtggcgtgatctcggctca
4081 ttgtaacctctgcttctgggttcaagcgattcttgtgctcagctcctgagttagctgg
4141 aatcggcggtttactactcatagtttttatcttgccaagaggctcaccctgacgttggcct
4201 gccttagctctgttaaaaaggttgtgctctgcccaggtgagggtgactcacgcctgtaatcc
4261 cagcactttgggaggccaaggtggggcggatcacgaggtcaagagttcgagaccagcctga
4321 ccaacatggtgaaaccccgctctctactaaaaatacaaaaattagctgggtgggtgtggtg
4381 tcacgggcttgtaatcccggctactcaggaggctgacaggagaatcgcttgaacccggga
4441 ggtggagggttgtagttagctgggtgacctgcacactctagccccgggcaacagaa
4501 tgagactctgtctcaaaaaacaaacaaaaaaaaactgtctttgggtgattgtggagagggtg
4561 atcatatgtccgctgtgttctcccatccaagcctgggtgactggtagctctcatgaggg
4621 tgtgatggtcaggccttggccctgggtgtaaatctgtctgtaaataggctgtgactttcta
4681 actcaacatcaccactggcttaggagttccatttgcacaggggggagctttattgcttttc
4741 tatatacataagaactttacgtctcaccagtgagggttggggcagccaagacaccacaatg
4801 ctgggtgtcttcagaatggacttaacttctgctccttttctgcttctgcccagtgagatcacc
4861 atctctcagatgagcaggacccccacctctggcagctcttcttggggagaatcactctga
4921 gctcccctcctttgtaaatataacattgggtttagcaccaccctccatcacggctgagtg
4981 agccatgattgcaccactgtgttccagcctgggtgacagagcaagacccccattccaaaaa
5041 aaaaaaaaaatttgccaaatagacacagtcctcacaagaaggttggcagaataacaccc
5101 ctctcaccacatctcctcagaacctgtgtatttggttagggttacgaaacaaaaggggaatt
5161 aagtttgcataatcagctgaccttaaaacaggaagaggatcccgtgttctgggtgggccc
5221 gaggaatcacaagatccctaaatgtggaagaggaaggcaaaaagagtcacatcaaag
5281 agacttaagaaagacttgactagacactgctgggtttgaaatgggaaggggaaggggtcag
5341 aagccaaggaatgtgtgtgctgtctagaagctggaaaaggcaaggaagtagattctcccc
5401 tagagcctcccagaaggagcgtagccctgctgatgccttaatttttagctcagtgagacca
5461 ggccgacttaagacctatagaaccgtcagatgacaacattatgttttaagccgctacaat
5521 tgtggcaatttgttaaaagtaacaatagaagacctaatgtaagtaagtctcatgtggaagt
5581 tgccatggttttgttttctttcttttttttttttttttttttttttttttttttttttttt
5641 gcccaggctggagtgacagaggcgtgatctcagctcactgcaacctccacctccaggggttc
5701 aagcgattctcctgctcagcctcctgagaagctgggttacagggtgagcaccaccacac
5761 ccagctaatttttttttgtatttttagtagagatggagtttcaccatgttggccaggtg
5821 gtcttgatctcttggcctcatgatccacctacctcagcctcccagaagtgttgggtattaca
5881 ggcagtagccaccgcaccagcccccatttttttttttttttttttttttttttttttttttt
5941 ttattgctccatccttgctaattgatagacttcatctcatatagaccagtggttttcagaga
6001 ccggcttattgattcagtaaaagggctgggggaaaactgcacacacaaggacatataaaatc
6061 taagcctctgctgggcccacatgacaggattcagaagcacctttgatataatgggtgccc
6121 gacctcatgtgaaagtcgaagttcagggtgtaattgtgagagcggtaagaattcactggcct
6181 aggcttgggaaactaatgctgcccaccacgttccagcctcagaggagaagaaatgacttgaa
6241 gaacataatttgtgaaatggccaaatgtttttcatttttatagcatcacacagtaaaagta
6301 gataaaaataattttaataaaaatacttctgtggggcagggtgagggttctagaccagc
6361 tcccagcactttgggaggccgaggcggcggatcacctgaggtcaggagttctagaccagc

6421 ctgaccaacatggagaaacccccatctctactaaaaatacaaaattagctgggtgtggtgg
6481 cgcactcctgttatcccagctactcaggaggctgaggcaggagaattgcttagactcggg
6541 aggcggagggttgggggtgagccgagatcacgccattgtactccagcctgggcaacaagagt
6601 gaaactgtctcaaaaaaagaaaaaaaaacagaatTTTgtgaataaaaggccacttggga polyA signal sequence
6661 gtatctgagccaccca major polyA site