a

```
-222  TGTATTTTTGGCAGATGATGTCCTGTCAAGGAAGCTGCAATTGCAAGCCCAGATGTTGGACGTGCACGCTCGCCCTCCCTATTAAACCCGCCCTCCTCCGATGCTCCCTCT
-111  CCTCCTCCCCCCTCGCCGCCGCCGCCGCCACCTTCTCCGGCAGCTCACCGTCGTTGTAGTAGCTCTGGTGGTGGTGGTGGTGACGACGACGGCTACGGCGGCGGCGACGGC
   1  ATGCTCCGGTCACTCCACTCCTCGTCGTCTTCTGACACGGATAACAACAGCGGTGGCTGCAAGAACAATGGCGGCGGCGGTGGCGAGGCCGCCGCCGCCGTTGAGGGTGGC
      M  L  R  S  L  H  S  S  S  S  S  D  T  D  N  N  S  G  G  C  K  N  N  G  G  G  G  G  E  A  A  A  A  V  E  G  G
 112  GGTGATCAGCGTGCCGTTGCGGCGGCGGCGCCGTCGACGCGGGACTTGCTGCTGGCGTGCGCGGACCTGCTGCAGAGGGGGGACCTGCCGGCGGCGAGGCGAGCGGCGGAG
      G  D  Q  R  A  V  A  A  A  A  P  S  T  R  D  L  L  L  A  C  A  D  L  L  Q  R  G  D  L  P  A  A  R  R  A  A  E
 223  ATCGTCTTGGCGGCGGCGGCGTCGCCGCGGGGCGACGCGGCGGACCGCCTGGCGTACCACTTCGCCCGCGCGCTGGCGCTCCGGGTGGACGCCAAGGCCGGCCATGGCCAC
      I  V  L  A  A  A  A  S  P  R  G  D  A  A  D  R  L  A  Y  H  F  A  R  A  L  A  L  R  V  D  A  K  A  G  H  G  H
 334  GTCGTCGTCGGCGGCGGCGCGGCGCGGCCGGCGTCGTCCGGGGCGTACCTGGCGTTCAACCAGATCGCGCCGTTCCTGCGTTTCGCGCACCTCACGGCGAACCAGGCCATC
      V  V  V  G  G  G  A  A  R  P  A  S  S  G  A  Y  L  A  F  N  Q  I  A  P  F  L  R  F  A  H  L  T  A  N  Q  A  I
 445  CTCGAGGCCGTCGACGGCGCGCGCCGCGTCCACATCCTCGACCTCGACGCCGTCCACGGCGTGCAGTGGCCGCCGCTGCTGCAGGCCATCGCCGAGCGCGCGGACCCGGCG
      L  E  A  V  D  G  A  R  R  V  H  I  L  D  L  D  A  V  H  G  V  Q  W  P  P  L  L  Q  A  I  A  E  R  A  D  P  A
 556  CTCGGCCCGCCCGAGGTCCGCGTCACCGGCGCCGGCGCCGACCGCGACACCCTCCTCCGCACCGGCAACCGCCTCCGCGCCTTCGCCCGCTCCATCCACCTCCCCTTCCAC
      L  G  P  P  E  V  R  V  T  G  A  G  A  D  R  D  T  L  L  R  T  G  N  R  L  R  A  F  A  R  S  I  H  L  P  F  H
 667  TTCACCCCGCTCCTCCTCTCCTGCGCCACCACGGCGCCCCACCACGTCGCCGGCACGAGCACCGGCGCCGCCGCCGCCGCCTCGACCGCCGCCGCCGCCACTGGCCTCGAG
      F  T  P  L  L  L  S  C  A  T  T  A  P  H  H  V  A  G  T  S  T  G  A  A  A  A  A  S  T  A  A  A  A  T  G  L  E
 778  TTTCACCCGGACGAGACGCTCGCCGTGAACTGCGTCATGTTCTTGCACAACCTGGCCGGCCACGACGAGCTCGCCGCGTTCTTGAAGTGGGTGAAGGCCATGTCGCCCGCC
      F  H  P  D  E  T  L  A  V  N  C  V  M  F  L  H  N  L  A  G  H  D  E  L  A  A  F  L  K  W  V  K  A  M  S  P  A
 889  GTGGTGACCATCGCCGAGAGGGAAGCAGGCGGCGGCGGCGGCGGCGGCGACCACATCGACCGACCTGCCGCGGCGGGTCGGGGTGGCCATGGATCACTACTCGGCGGTGTTC
      V  V  T  I  A  E  R  E  A  G  G  G  G  G  G  G  D  H  I  D  D  L  P  R  R  V  G  V  A  M  D  H  Y  S  A  V  F
1000  GAGGCGCTGGAGGCGACGGTGCCGCCGGGGAGCCGGGAGCGCCTCGCCGTGGAGCAGGAGGTGCTGGGCCGGGAGATCGAGGCCGCGGTGGGGCCCTCCGGCGGCCGGTGG
      E  A  L  E  A  T  V  P  P  G  S  R  E  R  L  A  V  E  Q  E  V  L  G  R  E  I  E  A  A  V  G  P  S  G  G  R  W
1111  TGGCGCGGCATCGAGCGGTGGGGCGGCGCCGCCCGCGCCGCGGGGTTCGCGGCGCGGCCGCTCAGCGCGTTCGCCGTGTCGCAGGCGCGGCTGCTGCTCCGGCTGCACTAC
      W  R  G  I  E  R  W  G  G  A  A  R  A  A  G  F  A  A  R  P  L  S  A  F  A  V  S  Q  A  R  L  L  L  R  L  H  Y
1222  CCGTCGGAGGGCTACCTCGTGCAGGAGGCGCGCGGCGCCTGCTTCCTCGGGTGGCAGACGCGCCCGCTGCTCTCCGTCTCAGCGTGGCAGCCGTCGTCGTCGTAGCCGGCG
      P  S  E  G  Y  L  V  Q  E  A  R  G  A  C  F  L  G  W  Q  T  R  P  L  L  S  V  S  A  W  Q  P  S  S  S  .
1333  ACGAAGACGACGATGGCGACGGCGGCGGCGACCACGGGCTAGGGTTTGCCAAGCTCGATTGATCGATCGATCACTACTGCTCCCATATATAAATAGAGAATATATATATAT
1444  ATATTATATATGAACATTTGCATAGATGCAATAATATTGGCTACTCGATCTCTAGTTAGCATGCAAGATCTCATGGAGGTTTTGGGAGGATGTGAACTAGTTGGG
```

b

```
   1  ATGCTCCGGTCACTCCACTCCTCGTCGTCTTCTGACACGGATAACAACAGCGGTGGCTGCAAGAACAATGGCGGCGGCGGTGGCGAGGCCGCCGCCGCCGTTGAGGGTGGC
      M  L  R  S  L  H  S  S  S  S  S  D  T  D  N  N  S  G  G  C  K  N  N  G  G  G  G  G  E  A  A  A  A  V  E  G  G
 112  GGTGATCAGCGTGCCGTTGCGGCGGCGGCGCCGTCGACGCGGGACTTGCTGCTGGCGTGCGCGGACCTGCTGCAGAGGGGGGACCTGCCGGCGGCGAGGCGAGCGGCGGAG
      G  D  Q  R  A  V  A  A  A  A  P  S  T  R  D  L  L  L  A  C  A  D  L  L  Q  R  G  D  L  P  A  A  R  R  A  A  E
 223  ATCGTCTTGGCGGCGGCGGCGTCGCCGCGGGGCGACGCGGCGGACCGCCTGGCGTACCACTTCGCCCGCGCGCTGGCGCTCCGGGTGGACGCCAAGGCCGGCCATGGCCAC
      I  V  L  A  A  A  A  S  P  R  G  D  A  A  D  R  L  A  Y  H  F  A  R  A  L  A  L  R  V  D  A  K  A  G  H  G  H
 334  GTCGTCGTCGGCGGCGGCGCGGCGCGGCCGGCGTCGTCCGGGGCGTACCTGGCGTTCAACCAGATCGCGCCGTTCCTGCGTTTCGCGCACCTCACGGCGAACCAGGCCATC
      V  V  V  G  G  G  A  A  R  P  A  S  S  G  A  Y  L  A  F  N  Q  I  A  P  F  L  R  F  A  H  L  T  A  N  Q  A  I
 445  CTCGAGGCCGTCGACGGCGCGCGCCGCGTCCACATCCTCGACCTCGACGCCGTCCACGGCGTGCAGTGGCCGCCGCTGCTGCAGGCCATCGCCGAGCGCGCGGACCCGGCG
      L  E  A  V  D  G  A  R  R  V  H  I  L  D  L  D  A  V  H  G  V  Q  W  P  P  L  L  Q  A  I  A  E  R  A  D  P  A
 556  CTCGGCCCGCCCGAGGTCCGCGTCACCGGCGCCGGCGCCGACCGCGACACCCTCCTCCGCACCGGCAACCGCCTCCGCGCCTTCGCCCGCTCCATCCACCTCCCCTTCCAC
      L  G  P  P  E  V  R  V  T  G  A  G  A  D  R  D  T  L  L  R  T  G  N  R  L  R  A  F  A  R  S  I  H  L  P  F  H
 667  TTCACCCCGCTCCTCCTCTCCTGCGCCACCACGGCGCCCCACCACGTCGCCGGCACGAGCACCGGCGCCGCCGCCGCCGCCTCGACCGCCGCCGCCGCCACTGGCCTCGAG
      F  T  P  L  L  L  S  C  A  T  T  A  P  H  H  V  A  G  T  S  T  G  A  A  A  A  A  S  T  A  A  A  A  T  G  L  E
 778  TTTCACCCGGACGAGACGCTCGCCGTGAACTGCGTCATGTTCTTGCACAACCTGGCCGGCCACGACGAGCTCGCCGCGTTCTTGAAGTGGGTGAAGGCCATGTCGCCCGCC
      F  H  P  D  E  T  L  A  V  N  C  V  M  F  L  H  N  L  A  G  H  D  E  L  A  A  F  L  K  W  V  K  A  M  S  P  A
 889  GTGGTGACCATCGCCGAGAGGGAAGCAGGCGGCGGCGGCGGCGGCGGCGACCACATCGACTGTTGGAAATGGCCAAGATCGGAGACACACACAAACACGAGAGTGAGTAGC
      V  V  T  I  A  E  R  E  A  G  G  G  G  G  G  G  D  H  I  D  C  W  K  W  P  R  S  E  T  H  T  N  T  R  V  S  S
1000  AGAGTTTGTGAAGCTTGA
      R  V  C  E  A  .
```