Assignment 2 – Task 2

By Andrea Ernst, 588404 and Jonas Kramer, 592764

Human KRAS sequence from: https://www.ncbi.nlm.nih.gov/nuccore/NM_021284.7?report=fasta

Length: 5306

Murine KRAS sequence from https://www.ncbi.nlm.nih.gov/nuccore/NM_004985.5?report=fasta

Length: 4674

Differences between our and Emboss alignment:

The Emboss algorithm produced a much longer alignment core, even much longer than the entire sequences, which are compared.

Core length – ours: 4675

Core length – Emboss: 5614

With defaults Emboss uses a different score matrix:

Match: +5 (+1 ours)
I/D/R: -4 (-1 ours)

This could explain the much longer core length, since extending to have one more match, while having also one more mismatch, is still rewarded. It seems gaps are also favoured by the Emboss algorithm, as there are several long-running sections with just gaps in the pattern.

Total number of gaps in both cores:

Ours: 721
Emboss: 1260

Our local alignment (found just one core):

********AGGCGGCGGCCGCGGCGGCGGAGGCAGCAGCG__GCGGCGGCAG_TG_G_CGGC_GGCGAAGG
TGGCGGCGGCTCGGCCAGTACTCCCGGC_CCC_CGCCATTTCGGACTGGGAGCGAGCGCGGCGCAGGCACTG
AAGGCGGCGGCGGG_GCCAGAGGCTCAGCGGCTCC_CAGGTGCGGGAGAGAGGCCTGCTGAAAATGACTGA
ATATAAACTTGTGGTAGTTGGAGCTGGTGGCGTAGGCAAGAGTGCCTTGACGATACAGCTAATTCAGAATCAT
TTTGTGGACGAATATGATCCAACAATAGAGGATTCCTACAGGAAGCAAGTAGTAATTGATGGAGAAACCTGTC
TCTTGGATATTCTCGACACAGCAGGTCAAGAGGAGTACAGTGCAATGAGGGACCAGTACATGAGGACTGGGG
AGGGCTTTCTTTGTGTATTTGCCATAAATAATACTAAATCATTTGAAGATATTCACCATTATAGAGAACAAATTA
AAAGAGTTAAGGACTCTGAAGATGTACCTATGGTCCTAGTAGGAAATAAATGTGATTTGCCTTCTAGAACAGT
AGACACAAAACAGGCTCAGGACTTAGCAAGAAGTTATGGAATTCCTTTTATTGAAACATCAGCAAAGACAAGA
CAGGGTGTTGATGATGCCTTCTATACATTAGTTCGAGAAATTCGAAAACATAAAGAAAAGATGAGCAAAGATG
GTAAAAGAAGAAAAAGAAGTCAAAGACAAAGTGTGTA_A_TTATGTAAATACAATTTGTACTTTTTTCTTAAG
GCATACT_A_GTACAAGTGGTAATTTTTGTACATTACACTAAATTATTAGCATTTGTTTTAGCATTACCTAAT_TT
TTTT___C__CTGCTCCATGCAGACTGTTAGCTTTTACCTTAAATGCTTATTTTAAAATGA_CAGTGGAAGTTTT_

```
TTTTTCCTCTAAGTGCCAGTATTCCCAGAGTTTTGGTTTTTGAACTAGCAATGCCTGTGAAAAAGAAACTGAATA
CCTAAGATTTCTGTCTTGGGGCTTTTGGTGCATGCAGTTGATTACTT_CTTATTTTTCTTACCAATTGTGAATGTT
GGTGTGAAACAA_ATTAATGAAGCTTTTGAATCATCCCTATTCTGTGTTTTATCTAGTCA_CATAAATGGATTAA
TTA_CTAATTTCAGTTGA_GACCTTCTAATTGGTTTTTACTGAAACATTGAGGGAACACAAATTTATGGGCTTCC
TGATGATGATT_CTTCTAGGCATCATGTCCTATAGTTTGTCATCCCTGATGAAT_GTAAAGTTACACTGTTCACA
AAGG_TTT_TGTCTCCTTTCCACTGCTATTAGTCATGGTCACTCTCCCCAAAATATTATATTTTTTCTATAAAAG
AAAAAAATGGAAAAAAATTACAAGGCAATGGAAACTATT_ATAAGGCCATTTCCTTTTCACAT_T_AGATAAAT
TACTATAAAGACTCCTA_A_TAGCTTTTCCTGTTAAGGCAGACCCAGTATGAAATGGGGATTATTATAGCAACC
ATTTTGGGGCTATATTTACATGCTACTAAATTTTTATAATAATTGAAA_AGATTTTAACA_AGTATAAAAAATTCT
CATAGGAATTAAATGTAG__TCTCCC_TGTGTCAGACTGCTCTTTCATAGTATAACTTTAAA_TCTTTTCTTCAAC
TTGAGTCT_TTGAA_GATAGTTTTAATTCTGCTTGTGACAT_TAAAAGATTATTTGGGCCAGTTATAGCTTATTA
GGTGTTGAAGAGACCA_AGGTTGCAAGGCCAGGCCCTGT__GTGAACCTTTGA_G__CTTT_CATAGAGAGTTT
CACAGCA_TGGACTGTGTCCCCACGGTCATCCAGTGTTGTC_ATGCATTGGTTAGTCAAAATGGGGAGGGACT
AGGGCAG__TTTGGATAGCTCAACA_AGATACAATCTCACTC_TGTGGTGGTCCTGCTGACAAATCAAGAGCAT
TGCTTTTGTTTCTTAAGAA_AACAAACTCTTTTTTAAAAATTACTTTTAAATATTAACT_CAAAAGTTGAGATTTT
GGGGT_GGTGGTGTGC__CAAGACATTAATTTTTTTTTTAAA_CAATGAAGTGA_AA_AAGTTTT_ACAATCTCT
AGGTTTGGCTAGTTCTCTTAACACTGGTTAAATTAACATTGCATAAACACTTTTCAAGTCTGATCCA__TATTTAA
__T_A___A_TGCTTTAA_AATAA__AA_AT_AAAAACAATC_CTTTTGATAAATTTAAAATGTTACTTATTTTAAA
ATAAATGAAGTGAGATGGCATGGTGAGGTGAAAGTATCACTGGACTAGGAAGAAGGTGACTTAGGTTCTAGA
TAGGTGTCTTTTAGGACTCTGATTTTGAGGACATCACT_TACTATCCATTTCTTCATGTTAAAAGAAGTCATCTC
AAACTCTTAGTTTTTTTTTTTTTACAACTATGTAATTTATATTCCATTTACATAAG__GATACACTTATTTGTCAAGC
TCAGCACAATCTGTAAAT_TTTTAACCTATG__TTACACCATCTTCAGTGCCAGTCTTGGGCAAAATTGTGCAAG
AGGTGAAGTTTATATTTGAATATCCATTCTCGT_TTT_AGGACTCTTCTTCCAT_ATTA_G_TGTCATCTT__GCCT
CCCTACCTTCCACATGCCCCATGACTT_GATGCAGTTTTAATACTTGTAATTCCCCTAAC__CATAAGATTTACTG
CTGCTGTGGATATCTCCATGAAGTTTTCCCACT__G__AG_TCACATCAGAAA_TGCCCTACAT__CTTAT_TTCC
TCAGGG_CTCAAGAGAATCTGACAGATACCATAAAGGGATTTGACCTAATCACTAATTTTCAGGTGGTGGCTG
ATGCTTTGAACATCTCTTTGCTGCCCAATCCATTAGCGACAGTAGGATTTTTCAAACCTGGTATGAATAGACAG
AACCCTATCCAGTGGAAGGAGAATTTAATAAAGATAGTGCTGAAAGAATTCCTTAGGTAATCTATAACTAGGA
CTACTCCTGGTAACAGTAATACATTCCATTGTTTTAGTAACCAGAAATCTTCATGCAATGAAAAATACTTTAATT
CATGAAGCTTACTTTTTTTTTTTGGTGTCAGAGTCTCGCTCTTGTCACCCAGGCTGGAATGCAGTGGCGCCATCT
CAGCTCACTGCAACCTCCATCTCCCAGGTTCAAGCGATTCTCGTGCCTCGGCCTCCTGAGTAGCTGGGATTACA
GGCGTGTGCCACTACACTCAACTAATTTTTGTATTTTTAGGAGAGACGGGGTTTCACCCTGTTGGCCAGGCTGG
TCTCGAACTCCTGACCTC_AAGTGATTCACCCACCTTGGCCTCATAAACCTGT_TTTGCAGAACTCATTTATTCAG
_CAAATATTTATTGAGTGCCTACCAGATGCCAGTCACCACACAAGGCACT__GGGTATATGGTATCCCCAAACA
AGAGACATAATCCCGGTCCT_TAGGTAGTGCTAGTGTGGTCTGTAATATCT_TAC_TAAGGCCTTTGGTATACG
ACCCAGAGATAAC_ACGATGCGT_ATTTTAGTTTTGCAAGAAGGGGTTTGG_TCTCTGTGCCAGCTCTATAAT
TGTTTTGCTACGATTCCACTGAAACTCTTCGATCAAGCTACTTTATGTAAATCACTTCATTGTTTTAAAGGAATAA
ACTTGATTATATTGTTTTTTTATTTGGCATAACTGTGATTCTTTTAGGACAATTACTGTACACATTAAGGTGTATG
TCAGATATTCATATTGACCCAAATGTGTAATATTCCAGTTTTCTCTGCATAAGTAATTAAAA_TATACTTAAAAAT
TAATAGTTTTATCTGGGTACAAATAAACAGGTGCCTGAACTAGTTCACAGACAAGGAAACTTCTATGTAAAAAT
CACTATGATTTCTGAATTGCTATGTGAAACT_ACAGATCTTTGGAACACTGTTTAGGTAGGGTGTTAAGACTTA
CACAGTACCTCGTTTCTACACAGAGAA_AGAAATGGCCATACTTCAGGAA_CTGCAGTGCTTATGAGGGGATA
TTTAGGCCTCTTGAATTTTTGATGTAGATGGGCATTTTTTTAAGGTAGTGGTTAATTACCTTTATGTGAACTTTG
AATGGT_TTAACAAAAGATTTGTTTTTGTAGAGATTTTAAAGGGGGA_GAATTCTAGAAATAAAT__GTTACCT
AATTATTACAGCCTTAAAGACAAAAATCCTTGTTGAAGTTTTTTTAAAAA__AA_GCTAAATTACATAGACTTAG
GCATTAACATGTTTGTGG_AAGAATATAGCAGACGTATATTGTATCA_TTTGAGTGAATGTTCCCAAGTAGGCA
_TTCTAGGCTCTATTTAACTGAGTCACACTGCATAGGAATTTAGAACCTAACTTTTATAGGTTATCAAAACTGTT
GTCACCATTGCACAATTTTGTCCTAATATATACATAGAAACTT_TGTGGGGCATGTTAA__GTTACAGTTTGCAC
AAGTTCATCTCATTT_GTATTCCATTGATTTTTT_TTTTCTTCTAAACATTTTTC_T___T_CA_A_A_CAGTATAT
```

A_ACTTTTTT_T_AGGGGATTTTTTTTT__AGACAGCAAAAACT_ATCTGAAGATTTCCATTTGTCAAAAGTAAT
GATTTCTTGATAATTGTGTAGTAATGTTTTTTAGAACCCAGCAGTTACCTTAAAGCTGAATTTATATTTAGTAAC
TTCTGTGTTAATACTGGATAGCATGAATTCTGCATTGAGAAACTGAATAGCTG_T_C_A_TA_____AAAT___G
__AAA_C_T_____TTCTTTCTAAAGAAAGATACTCACATGAGTTCTTGAA_GAATAGTCATAACTAGATTAAGA
TCTGTGTTTAGTTTAATAGTTTGAAGTGCCTGTTTGGGATAATGATAGGTAATTTAGATGAATTTAGGGGAAA
AA_AAAGTTATCTGCAGATATGTTGAGGGCCCA_TCTCTC___CCCC_C_ACACCCCCACAGAGCTAA_CTGGGT
TACAGTGTTTTA__TCCGAAAGTTTCC_AATTCCACTGTCTTG_TGTTTTCATGTTGAAAATACTTTTGCATTTTTC
CTTTGAGTGCCAATTTCTTACTAGTACTATTTCTTAATGTAACATGTTTACCTGGAATGTATTTTAACTATTTTTGT
ATAGTGTAAACTGAAACATGCACATTTTGTACATTGTGCTTTC_TT_TT_____GT___G___G
___G_____A_____CATATGCAGTGTGATCCAGTTGTTTTCCATCATTTGGTTGCGCTGACCT
AGG_AATGTTGGTCATATCAAACATTAAA___AA___TGACCACTCTTTTAATTGAAATTAACTTTTAAATGTTTA
TAGGAGTATGTGCTGTGAAGTGATCTAAAATTTGTAATATTTTTGTCATGAACTGTACTACTCCTAATTATTGTA
ATGTAATAAAAATAGTTACAGTGAC**

********AGGCGGCGGCCGCGGCGGCTGAGGCGGCAGCGCTGTGGCGGCGGCTGAGACGGCAGGGGAAGG
CGGCGGCGGCTCGGCCCGGAGTCCCG_CTCCCGCGCCATTTCGGACCCGGAGCGAGCGCGGCGCGGGC_CTG
AAGGCGGCGGCGGGAGCCTGAGGCGCGGCGGCTCCGC_GGCGCGG_AGAGAGGCCTGCTGAAAATGACTGA
GTATAAACTTGTGGTGGTTGGAGCTGGTGGCGTAGGCAAGAGCGCCTTGACGATACAGCTAATTCAGAATCAC
TTTGTGGATGAGTATGACCCTACGATAGAGGACTCCTACAGGAAACAAGTAGTAATTGATGGAGAAACCTGTC
TCTTGGATATTCTCGACACAGCAGGTCAAGAGGAGTACAGTGCAATGAGGGACCAGTACATGAGAACTGGGG
AGGGCTTTCTTTGTGTATTTGCCATAAATAATACTAAATCATTTGAAGATATTCACCATTATAGAGAACAAATTA
AAAGAGTAAAGGACTCTGAAGATGTGCCTATGGTCCTGGTAGGGAATAAGTGTGATTTGCCTTCTAGAACAGT
AGACACGAAACAGGCTCAGGAGTTAGCAAGGAGTTACGGGATTCCGTTCATTGAGACCTCAGCAAAGACAAG
ACAGGGTGTTGACGATGCCTTCTATACATTAGTCCGAGAAATTCGAAAACATAAAGAAAAGATGAGCAAAGAT
GGGAAGAAGAAGAAGAAGAAGTCAAGGACAA_G_GTGTACAGTTATGTGAATAC__TTTGTACTCTTT_CTTA
AGGCACACTTAAGTAAAAGTG_TGATTTTTGTACATTACACTAAATTATTAGCATTTGTTTTAGCATTACCTAAT
CTTTTTTTTTCTTCTGTTCG_TGCAAACTGTCAGCTTTTATCTCAAATGCTTATTTTAAAA_GAACAGTGGAAACC
TTCTTTTT__TCTAAGTGCCAGTATTCCCTGGGTTTTGGACTTA_AACTAGCAATGCCTGTGGAAGAGA__CTAA
AGACCTGAGACT_CTGTCTTGGGA_TTT_GGTGCATGCAGTTGATTCCTTGCT_AGTTCTCTTACCAACTGTGAA
CACTGATGGGAAGCAGGAT_AATGAAGCTTCCGGACCATCCCTGCTCTGTGTCC_ATCTACTCATCC_AA_TGG
AGTCATTAGC_AGT__CAATCGCCG_C_TTC_ACT_GG_____AC__A__C_T_GAGGGGTCACAGACTTA_GG_
CTCCCTT_TGA_G_TCGCGTCCAG_CGT___GTCCTAGACTTTATCATCTTTCA_GAGGCGTAG_GC_AGACTGT
TCACAAAGGCTTTCTGTAGC_TTTCCACTGCAATTAATCTTGGTCACTC_CCTCAAA_TAGTATATTTTTTCTAGA
AAAGGGGAAAAATGGAAAAAAA__A_A_GGCAATGGAAAATGTTGA_AAT_CCATT_CAGTTTC_CATGTTAG
CTAAATTACTGTAA_GATTCCTATAATAGCTTTTCCTGGTAAGGCAGACCCAGTATGAAATAG___T_A__ATA_
_A_CCATTT_GGG_CTATATTTACATGCTACTAAATTTTTGTAATAATTCAAACA_ACTTTAGCATA_TATAAAAA
GTTCTCATAAGAATTAA__GTACAAT_TCCCCTTTGTCAGATTGTTCTT__ATCCTA_A_CTTTCAAGTCTTTT_TT
GAATTTCTGT_TGTTGAAAG_TAGTTTTAATG__G_TTGTGA_AGCTGAA_GATGATCTGAGACAGTTATAGCTT
GGCAGGTGTTGAGGAGACCAGAG_TTGCAGGGTTGGGCC_T_TACGTGAACCTGTGACGAACGCTAC_T_G_
G_GTTTTGCAGCACTG__CTG___C___AT__TCA___A_TGTTGGCGACGCATTGTTTGGTCAACATAGGG_G_
_A_TAAGG_AGACTTTG_ATGGCTTAGTATA_ATGCATTCTCAC_CATGTAACAGTCCTACTGACAAATCAAGAA
_ATT__T___GTTTAT_AATAATAAAAAA_T_TTTT__AAAAATTTCG___A__TGTTCGCTTCAAG_GTTGAGATT
TTGGGGTAGGAGGC_TACAACAAGAG_TAAATCTT_____AAAGCAA_G__GTTTTAAGAAGGTTTGAAAATG
_C_AGGTTTGACTAGT_CTCTCAACTCTAGCTAAACAAACATTCCC_AAGTACTTCCCAAATCTGAT__AGGTATT
TAAAATTATCTAATGCTTTAAGAATAGTTAACAGGAAAAA_AATCTCCTCAG_TGCACTTAAA__GCAACCC_TT
C_ACA_TCATT____TGAAATG__A_GATG_G___AAA_TATCACTGGACTATGA_G__G___ACT__GG____A_
_T_G_T__CT____G___TCTGATTTTAAGCAAATCACTGT_CTG_C__TTGGTT__T_T_____GAA_TCATCTCAA
AGAC__A_TTAACCTCC____CAGCCGTGTAAC__ATAGT___TT_ACAT__GTTGACACACCTAGTTATCAAGCT
CAGCACAATCTGTAACTGTTTTA_C__ATGGATTA_AC_ATCTTCACTGCCAGTCTTGGGCAAA_TTGTGCAAGA

GGTAAAATTTATATTTCAGTATCCATTCTCCCATTTCAGGACTCCCCT_CCAACATTATGCTGGC_T_TTCAGCCT
GTCT_C_T_C_ACCTGCCC_ATCACTTAG_TGTAGTTTTAATAATT_TC___CCCC_A_CTTCA_AAC_TTTGTTTCC
ACTATGGACAACTTCATGAACTTTGCCCACTAAGGTAGGT_ACATCA_AAGCTGCCCTA__TGGCTT_TCTTCCC
C_GGGACTGAA_A_A_TA__ACAGACACCATAGTGGGATTT_A___AA_C__TAAT____AGATGGT___T__T__
T_____CAG_____G__GCC_A__C_T_A_CAACA__A___TT___CAATC_TC__A__A_T__C___C_T_T_
___TGGAC_____TTCA_T_____TCCTGCTGC_____CC__AGGC___C_A___CT_GG__T_____G____C_
_____C_T_C_A__GT___AG_____GAA_T_TT__T_CAA__AA___T_T_AGTG__TGAA_C__A_____
____GA___CAGAG_CACA_____GTC_C__AG__TGGAA_G__GTGA_GC__T_TAA__TC__T_____TC_ATC
T___AG___C__C_AT_CA__T_CAT_GG__T____A__AG_TG__AT_A__G_____A_T____TC___TA_T
T___GT_TTTA_AT_A_A_A_____TA_CA____GT___CTAA_CAA_T___GAAAAAC__ACTTCGAAGT__TTCA_
__A__T___C___ATAAAGCTGTCTTTTTAAAAAT__TTTATTTACTCAAC_ATTTATTCAGTGCTTGTCATATTCTG
GGAATTACACTAGGCACTCAGGGTGC__GGTGTCCTCAATC_____C_T__T___GG_CCAGT_GGTA_TG_TA
GCATGATCTGTAATACCACTAAATAAGGCATATAGCATATGACTTAGACATAATGAA_ATACATGATTTGAGTT
TTGCAGAGA_GGAGTTTGGGT_T_TGT___A_C___AT__TC_____C__C__TTCC_C_____C_C__C___CA_G
_T__TT_A_GCAAG__A____ATTGTTT____G_____CT_G_TGA_ATCCA_____AT__G_CA_A_CT_T__TT
_____A__A_A_TCA____A_AC_T_A_____C___T_TT_ATAT_A__A_T_T__A_T_TTC_A_TTTT_T
CT__A_AAGGAAC_AGAAGTACCCT_AAAC_T_A_T__TTTT_T_TG___A_AA_T_____GTT_C_TAAACT_GT_
_ACATA_____TTC_AT__AGAA__CA_T_TC_TTT__G____G___GTGAATTTTA_AG_TCTTA__AA_A_
TGCA_AT_TAG__T___AATACTT_CTCATT___TC_TAT_T_CAGAG_GAACAGG__TG____TACTTCAA_AAGC
TGCAGTG_T_AT_A____ATC___AGA__T____A_TTTTTAATG__GA____CAATGTGTTAAAGAAGTGGT_AA
TTACCACTATGTAAA_TTTGAATTGTGTTA_CAC____TTTGGTT____A_ACA____AAAGGGGAAAGAATCCTA
GAAACAAATATGTTATCTAGTTACTGCAGCCTTAAAG_____TCCTTGTTGAAGTT_____AAAAAGCAATGC
TAAGTTACA__GTCATAGGCATTAACATGTTTATGGGAAGGATATAGTAGGCA_A_AT___A_CAATTTGAGTA
AATATTTTCA_GTAGGGAATTTTAGGCTCTACTGA_CTGAGTCACACTGCATAGGAATTTAGATCTTAACTTTTA
TAGGTTATCGAC_CT_TTGCCACCATTGCACAATTTTGTCCTAACATA_A_ATACAAG_TTCTGTGAGGCATGTC
AAAAGTTACAGTTTGCATAAATTCATCTCATTTTGTATTCCACTGATTTTACATTTTCCTC_AAACATACATACAT
ACATACATACAACAC_ACACACTCACACATGAAGGGTTTTTTTTTTGTAGGCAATAAAAATTTAACT_AA__T
TTCCATTTGTTAAAAGTAGTGATTTATTGAGAATTATGCAGTCAT_TTTTT_A_AACCCAAAAGTTAT_TTAAAG
GTGAATTTATACTCAATAACTTCTGTGT_AATACTGGGTAGCATGAATTCTGCATTGAAAAATTGAACAGATAA
TACCAATAGCTGTAAATTCTGTCAAAACATGAAAATTATTTCTAAAGAA_G_TAC__AT_T_AGTT_TTCAAAGA
ACAGTTAT___TAGAATCAGATCTGTGGTTTAGTTCAATAATTTGAAGTGCCTGTTTGGGATGGTGGTAGGCAT
TTTAGATGAATTT_GGG_AAAAATAAAGTT__CTGCAGAAATG_____CC_AGT_T_TCAGACCCCGCTA_AC
CCGCT__GAG_TGGGCTGTGTG_CTGTGTT__AGCTCC_A__GTGCCCCAAT_CC_C_GT_TTCATGTCTTCATGT
TGAAAC_ACTTCTGCATTTTTA_TTTGAGTGCCAATTTCTTACTAGTGCTATTTCTTAGTGTAACATGTTTACCTG
GGATGTATTTTAACTATTTTTGTATAGTGTAAACTGAAACATGCACATTTTGTACATTGTGCTTTCCTTCTTTCCA
TTCCTTTTCTTTCTGTTTTGTTTGTTTGTTTGTTTGTTTGTTTGTTATGGGACATATGCAGTGTGATCCAGTTGTTT
TCCATCCTTTGGTTGCGCTGACCTAGGGAATGTTGGTCATATCAAACATTAAATTTAAAAGTGACCACTCTT__A
ATTAAAATTAACTTTTAAATGTTTATAGGAGTACGTGCTGTGAAGTGATCTGAAATTTGTAATATTTTTGTCATG
AACCGTACTGCTCCTAATCATTGTAATGTAATAAAAATAGTTATGGTGAC**

EMBOSS Alignment (beginning):

```
# Commandline: water
#    -auto
#    -stdout
#    -asequence emboss_water-I20220527-203641-0057-37591492-p1m.asequence
#    -bsequence emboss_water-I20220527-203641-0057-37591492-p1m.bsequence
#    -datafile EDNAFULL
#    -gapopen 10.0
#    -gapextend 0.5
#    -aformat3 pair
#    -snucleotide1
#    -snucleotide2
# Align_format: pair
# Report_file: stdout
########################################

#=======================================
#
# Aligned_sequences: 2
# 1: EMBOSS_001
# 2: EMBOSS_001
# Matrix: EDNAFULL
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 5614
# Identity:    3681/5614 (65.6%)
# Similarity:  3681/5614 (65.6%)
# Gaps:        1260/5614 (22.4%)
# Score: 13069.0
#
#
#=======================================

EMBOSS_001         3 AGGCGGCGGCCGCGGCGGCGGAGGCAGCAGC-----GGCGGCGGCAGTGG     47
                     |||||||||||||||||||.|||||.|||||     |||||||||.|.|.
EMBOSS_001         9 AGGCGGCGGCCGCGGCGGCTGAGGCGGCAGCGCTGTGGCGGCGGCTGAGA     58

EMBOSS_001        48 CGGC-GGCGAAGGTGGCGGCGGCTCGGCCAGTACTCCCG-GCCCCCGCCA     95
                     |||| ||.|||||.||||||||||||||||||||.|.|.|||||  .|||.|||||
EMBOSS_001        59 CGGCAGGGGAAGGCGGCGGCGGCTCGGCCCGGAGTCCCGCTCCCGCGCCA    108

EMBOSS_001        96 TTTCGGACTGGGAGCGAGCGCGGCGCAGGCACTGAAGGCGGCGGCGGG-G    144
                     |||||||||..|||||||||||||||.|||  ||||||||||||||||| |
EMBOSS_001       109 TTTCGGACCCGGAGCGAGCGCGGCGCGGGC-CTGAAGGCGGCGGCGGGAG    157

EMBOSS_001       145 CCAGAGGCTCAGCGGCTCCCAGGTGCGGGAGAGAGGCCTGCTGAAAATGA    194
                     ||.|||||.|.|.|||||||..||.||  |||||||||||||||||||||
EMBOSS_001       158 CCTGAGGCGCGGCGGCTCCGCGGCGC-GGAGAGAGGCCTGCTGAAAATGA    206

EMBOSS_001       195 CTGAATATAAACTTGTGGTAGTTGGAGCTGGTGGCGTAGGCAAGAGTGCC    244
                     |||||.||||||||||||||.|||||||||||||||||||||||||.|||
EMBOSS_001       207 CTGAGTATAAACTTGTGGTGGTTGGAGCTGGTGGCGTAGGCAAGAGCGCC    256

EMBOSS_001       245 TTGACGATACAGCTAATTCAGAATCATTTTGTGGACGAATATGATCCAAC    294
                     |||||||||||||||||||||||||.||||||||||.||.||||.||.||
EMBOSS_001       257 TTGACGATACAGCTAATTCAGAATCACTTTGTGGATGAGTATGACCCTAC    306

EMBOSS_001       295 AATAGAGGATTCCTACAGGAAGCAAGTAGTAATTGATGGAGAAACCTGTC    344
                     .||||||||.||||||||||||.||||||||||||||||||||||||||||
EMBOSS_001       307 GATAGAGGACTCCTACAGGAAACAAGTAGTAATTGATGGAGAAACCTGTC    356

EMBOSS_001       345 TCTTGGATATTCTCGACACAGCAGGTCAAGAGGAGTACAGTGCAATGAGG    394
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
EMBOSS_001       357 TCTTGGATATTCTCGACACAGCAGGTCAAGAGGAGTACAGTGCAATGAGG    406

EMBOSS_001       395 GACCAGTACATGAGGACTGGGGAGGGCTTTCTTTGTGTATTTGCCATAAA    444
                     ||||||||||||||||.|||||||||||||||||||||||||||||||||
EMBOSS_001       407 GACCAGTACATGAGAACTGGGGAGGGCTTTCTTTGTGTATTTGCCATAAA    456

EMBOSS_001       445 TAATACTAAATCATTTGAAGATATTCACCATTATAGAGAACAAATTAAAA    494
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
EMBOSS_001       457 TAATACTAAATCATTTGAAGATATTCACCATTATAGAGAACAAATTAAAA    506
```