

Additional File 2. Key alignments between CBO0798 and CNTs produced by SSEARCH and PSI-BLAST.

SSEARCH alignment between BoNT/A and flagellin CBO0798 (default parameters used with -z 11 flag). E-values are based on a single pairwise alignment of both sequences rather than a database search.

s-w opt: 175 Z-score: 83.9 bits: 26.3 E(): 0.0073
Smith-Waterman score: 175; 21.1% identity (59.2% similar) in 488 aa overlap (676-1113:1-469)

```

      650      660      670      680      690      700
BoNT/A MLYKDDFVGALIFSGAVILLEFIPETIAIPVLGTFALVSYIA--NKVLTVQTIDNALSKRN
      .. ... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
flagel      MNKLSLINNNYFNKLATINSMKLSTNKRKRV
      10      20      30

      710      720      730      740
BoNT/A EKWDEVYKYIVTNWLAKVNTQI---DLIRKKMKEAL-----ENQAEATKAI-----
      . . . : . : : : : : : : : : : : : : : : : : : : : : : : : : :
flagel NVAADSPSEI--GRISRINTSVRSRNIIVVKNIQEGISLLQVKENAVNDLKEMGTRLKELS
      40      50      60      70      80

      750      760      770      780      790      800
BoNT/A INYQYNQYTEEKNINFNIDDLSSKLNESINKAMININKFLNQCSVSYLMNSMIPYGVK
      . : : : . : : : : : : : : : : : : : : : : : : : : : : : : : :
flagel LMYKNNTLNKDDKKNIEYEAKSLLENMSFTIKNTKFNKRNLFNGEKLN-IQGSTYNTDIK
      90      100      110      120      130      140

      810      820      830      840      850
BoNT/A RLEDFDASLKDALLKYYIDNRGT---LI---GQVDRLKDKVNNLTSTDIPFQLSKYVD
      . : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
flagel -LMDFKVSKGDTI--FIGNNKSSNESHTIKVSDNEIDTILKKYNLEKADKIVLYELEMVE
      150      160      170      180      190      200

      860      870      880      890      900      910
BoNT/A NQRLSTFT- EY-IKNIINTSILNRLYESNHLIDLRLYASKINIGSKVN--FDPIDKNQI
      . . . : : : : : : : : : : : : : : : : : : : : : : : : : : :
flagel EHNKTVFTSKYGYKNNANTDIVNVSNANKMFSIN---GTINDNSKVGYLQEKDNKLI
      210      220      230      240      250      260

      920      930      940      950      960      970
BoNT/A QLFNLESSKIEVILKNAIVNSMYENFSTSFWRIPKYPNSISLNNEYTIINC MENNSGW
      . : : : . : : : . : : : : : : : : : : : : : : : : : : : :
flagel GNININDKGNIQL-DAIMEDALKTNLNLSIKTTIINK--NDHTFRGNTSIID-IRDNSGM
      270      280      290      300      310

      980      990      1000      1010      1020
BoNT/A -----KVSLN-YGEIITLQDTQEIQRVVFQYSQMINISDYIN-RWIFVTITNNRLNN-
      . : : : . : : : : : : : : : : : : : : : : : : : : : : :
flagel TYIFKEITVDEFKEIIANDKSNKNKQ-IVFELTNKKCALDINSKDVNLYISDEEYNSF
      320      330      340      350      360      370

      1030      1040      1050      1060      1070
BoNT/A --SKIYINGRLIDQKPISNLGNIHA-----SNNIMFKLDGCRDTH--RYIWKYFNLFD-
      : : . : : : : : : : : : : : : : : : : : : : : : : : : :
flagel DISDILNSPNIIGKVLSELKNYQNYIGIKTNELEYKLNFEQNNQILEETLTKIQSIDI
      380      390      400      410      420      430

      1080      1090      1100      1110      1120      1130
BoNT/A -KELNEKEIKDLYDNQNSGILKDFWGDYLYDKPYMYMLNLYDPNKYVDVNNVGIRGYMY
      : : : : . : : : . : : : : : : : : : : : : : : : : :
flagel AKELVEKSKNEILVN-TNAVLLQS----SLENDKNYILTLLR
      440      450      460      470
```

SSEARCH alignment between NTNH/A and flagellin CBO0798 (default parameters used with -z 11,12,14,15,16 flag). E-values are based on a single pairwise alignment of both sequences rather than a database search.

s-w opt: 151 Z-score: 64.8 bits: 22.7 E(): 0.081
Smith-Waterman score: 151; 23.6% identity (51.4% similar) in 449 aa overlap (158-562:44-471)

```

      130      140      150      160      170      180
NTNH  APKSNKKLNSLISSTIPFPYAGYRETNLYLSEDNKSFYASNIVIFGPGANIVENNTV FYK
      . : : . : : : : : : : : : : : : : : : : : : : :
flagel NKLATINSMKLSTNKRNVNAADSPSEIGRISRINTSVRSRNIIV---KNIQEGISLLQV
      20      30      40      50      60

      190      200      210      220      230      240
NTNH  KEDAENGMGTMTEIWFQPFLLTYKYDEFYIDPA--IEL-IKCLIKSLYFLYGIKPSDDLVI
      : : : : : : : : : : : : : : : : : : : : : : : : : :
flagel KENAVNDLKEMGTRLKELSLMYKNNTLNKDDKKNIEYEAKSLLENMSFT--IK-NTKFN
      70      80      90      100      110      120

      250      260      270      280      290
NTNH  PYRLRSELENIEYSQLN----IVDLLVSGGIDPKFINTDP-----YWFDTNYFSNAKK
      . : : : : : : : : : : : : : : : : : : : : : : :
flagel KNLFNGEKLNIQGSTYNTDIKLMDFKVS KG-DTIFIGNNKSSNESHTIKVSDNEIDTILK
      130      140      150      160      170      180

      300      310      320      330      340
NTNH  VFE----DHRNIYETEIEGNAIGNDIKRLKQKFRININ-DIWELNLNYSKEFSIMMP
      . : : : . : : : . : : : : : : : : : : : : : : :
flagel KYNLEKADKIVLYELEMVEEHNIKT--VFTSKYGYKNNANTDIVNVSNANKMFSINGT
      190      200      210      220      230      240
```

PSI-BLAST detection of *C. butyricum* BoNT/E from flagellin CB00798
(composition-based statistics, *Clostridia*-restricted database)

Score = 37.9 bits (87), Expect = 0.050, Method: Composition-based stats.
Identities = 82/345 (23%), Positives = 136/345 (39%), Gaps = 52/345 (15%)

Query 365 NLYISDEEYNSFDISDILNSPNI-IGEKVLSLKN--YQNYIGIK 406
LYI N D ILN NI ++ +L ++ N Y YIGI+
Sbjct 1004 KLYI---NGNLDKKSIILNLGNIHVSNDILFKIVNCSYTRYIGIR 1045