

Aufgabe 2

Human Gene Sequenz von "Human Hemoglobin subunit alpha"(HBA)

```
MVLSPADKTN VKAAWGKVGA HAGEYGAEAL ERMFLSFPTT KTYFP HFDLS  
HGSAQVKGHG KKVADALTNA VAHVDDMPNA LSALSDLHAH KLRVDPVNFK  
LLSHCLLVTL AAHLPAEFTP AVHASL DKFL ASVSTVLTSK YR
```

Human Gene Sequenz von "Human Hemoglobin subunit beta"(HBB)

```
MVHLTPEEKS AVTALWGKVN VDEVGGEALG RLLVVYPWTQ RFFESFGDLS  
TPDAVMGNPK VKAHGKKVLG AFSDGLAHLN NLKGTFATLS ELHCDKLHVD  
PENFRLLGNV LVCVLAHHFG KEFTPPVQAA YQKVVAGVAN ALAHKYH
```

Aufgabe 3

Global Alignment

Es versucht den End to End Alignment auf einer ganzen Sequenz.

In der Regel, wurde Global Alignment zum Vergleich von homologen Genen verwendet.

Es verwendet Needleman-Wunsch-Algorithmus.

Local Alignment

Es versucht den Ort, an dem die höchste Alignment zwischen beiden Sequenzen gefunden werden kann, zu finden.

Es sucht nach konservierten Mustern in DNA Sequenzen oder Motifs in zwei Sequenzen.

Es verwendet Smith-Waterman-Algorithmus.

Aufgabe 4

a.

```
# Aligned_sequences: 2
# 1: EMBOSS_001
# 2: EMBOSS_001
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 149
# Identity:      65/149 (43.6%)
# Similarity:    90/149 (60.4%)
# Gaps:          9/149 ( 6.0%)
# Score: 292.5
#
#=====

EMBOSS_001      1 MV-LSPADKTNVKAANGKVGAGAHAGEYGAEALERMF LSFPTTKTYFPHF-D      48
|| |:|:|:|:|.|.|||| |..|.|.|||.|:|:|.|.|.|.|.|.|.|.
EMBOSS_001      1 MVHLTPEEKSAVTALWGKV--NVDEVGGEALGRLLVVYPWTQRFFESFGD      48

EMBOSS_001     49 LS-----HGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSDLHAHKLR      93
|| |..|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.
EMBOSS_001     49 LSTPDAMGNPKVKAHGKKVLGAFSDGLAHLNLIKGT FATLSELHCDKLH      98

EMBOSS_001     94 VDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR      142
|||.||:|.|:|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.
EMBOSS_001     99 VDPENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH      147
```

Abb.1. Globales Alignment mit voreingestellten Parametern

```
# Aligned_sequences: 2
# 1: EMBOSS_001
# 2: EMBOSS_001
# Matrix: EBLOSUM30
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 149
# Identity:      65/149 (43.6%)
# Similarity:    102/149 (68.5%)
# Gaps:          9/149 ( 6.0%)
# Score: 401.5
#
#=====

EMBOSS_001      1 MV-LSPADKTNVKAANGKVGAGAHAGEYGAEALERMF LSFPTTKTYFPHF-D      48
|| |:|:|:|:|.|.|||| |..|.|.|||.|:|:|.|.|.|.|.|.|.|.
EMBOSS_001      1 MVHLTPEEKSAVTALWGKVN--VDEVGGEALGRLLVVYPWTQRFFESFGD      48

EMBOSS_001     49 LS-----HGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSDLHAHKLR      93
|| |..|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.
EMBOSS_001     49 LSTPDAMGNPKVKAHGKKVLGAFSDGLAHLNLIKGT FATLSELHCDKLH      98

EMBOSS_001     94 VDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR      142
|||.||:|.|:|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.
EMBOSS_001     99 VDPENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH      147
```

Abb.2. Globales Alignment mit einer anderen Substitution MATRIX


```

HBA_HUMAN  -----VLSPADKTNVKAAWGKVGA--HAGEYGAEALERMFLSFPTTKTYFPHF
HBB_HUMAN  -----VHLTPEEKSAVTALWGKV---NVDEVGGEALGRLLVVYPWTQRFFESF
HBA_HUMAN  -DLS-----HGSAQVKGHGKKVADALTNAVAHV---D--DMPNALSALSDLHAHKL-
HBB_HUMAN  GDLSTPDVAVMGNPKVKAHGKKVLGAFSDGLAHL---D--NLKGTFFATLSELHCDKL-
HBA_HUMAN  -RVDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR-----
HBB_HUMAN  -HVDPENFRLLGNVLCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH-----

```

Abb.5. Alignment von Human Hemoglobin subunit alpha und beta aus Vorlesung

Im Vergleich zu den Sequenz von Vorlesung und aus dem Such Tool sehen wir den Unterschied von Anfang, wo es in der zehnte Stelle der Beta Sequenz eine Insertion gibt und deshalb sind allen Basen eine Stelle vor verschoben. Die Anzahl der Deletion ist sogar mehr in die Sequenz aus Vorlesung als der gedruckte.

b.

Substitution Matrix

Substitution Matrix beschreibt eine relative Rate, mit welcher im Laufe der Evolution eine Aminosäure in eine andere mutiert. Die Verwendung von Substitution Matrix hilft dem Score von bestimmten Sequenzalignment zuordnen, und somit bestimmt, wie gut das Alignment ist. Blosom62 ist eine der häufigste benutzte Matrix, und im Vergleich zu anderen verwandte ProteinSequenz , die maximal identisch dazu sind.

Gap Penalty

Es ist eine Methode, die Score von zwei oder mehr Sequenzen Alignment rechnet. Die Verwendung von Gaps (Lücken) darf den Aligned Sequenzen mehrere "Matches" als mit den wenigen Gaps, wenn die Sequenzen ausgerichtet werden. Minimale Gaps sind wichtig für den Alignment aber Gap Penalties ordnet den Score von bestimmten Sequenzalignment zu.