

Prof. Dr. Knut Reinert  
Dr. Roland Krause  
Matthias Winkelmann  
Patrick Pett

Institut für Informatik  
AG Algorithmische Bioinformatik

## Algorithmische Bioinformatik

### 9. Übungsblatt WS 11/12

Abgabe am 26. Januar 2012

Name:

Matrikelnummer:

Übungsgruppe:

1  2  3

Ich kann Aufgabe  nicht vorrechnen.

---

#### Aufgabe 1: Nussinov-Algorithmus

Der Nussinov-Algorithmus ist einer der wichtigen Algorithmen zur Berechnung von RNA-Sekundärstrukturen.

- a) Schreiben Sie ein Programm, das eine FASTA-Datei mit einer RNA-Sequenz einliest und den Nussinov-Algorithmus auf diese Sequenz anwendet.

Als Nebenbedingung soll gelten, dass Hairpin-Loops mindestens drei ungepaarte Basen enthalten. Die Basenpaare sollen unterschiedlich bewertet werden: 3 Punkte für G-C, 2 für A-U, 1 für G-U. Berechnen Sie die Matrix  $W(i, j)$  für den maximalen Score in der Teilsequenz zwischen  $i$  und  $j$ . Implementieren Sie das *traceback* und finden Sie die Menge der Basenpaare mit optimalem Score.

- b) Geben Sie die Funktion mit dem Kern des Algorithmus im Aufgabenzettel an.
- c) Demonstrieren Sie die praktische Anwendung und suchen Sie sich als Testfall die Sequenz der menschlichen tRNA<sup>1</sup> chr6.trna95. Behandeln Sie eventuelle Ts als Us. Welchen Score erhalten Sie mit dem Nussinov-Algorithmus?
- d) Stellen Sie die Struktur in mindestens drei unterschiedlichen Darstellungen dar.

---

<sup>1</sup><http://gtrnadb.ucsc.edu>

- e) Geben Sie in Ihrer Struktur Beispiele für hairpin, bulge und multiloop.
- f) Vergleichen Sie Ihr Ergebnis mit der Ausgabe des Zuker-Algorithmus. Verwenden Sie dazu die UNAFold-Implementierung<sup>2</sup>
- g) Vergleichen und diskutieren Sie die Ergebnisse. Tipp: Auf der oben genannten Seite ist auch eine Muster-Struktur der tRNA erhältlich. Vergleichen Sie in Bezug zu dieser.

## Aufgabe 2: Structure-Alignments

Gegeben ist folgendes Alignment:

```

1 2 3 4 5
A G U C C
A C U G A
A G U C U
A C U G G

```

- a) Beschreiben Sie kurz das Prinzip von *Structure Alignments*.
- b) Berechnen Sie den *mutual information content* zwischen den Spalten (1, 2), (1, 3), (2, 4), (2, 5), sowohl in der nativen Version, als auch mit der Modifikation von Gorodkin.
- c) Erläutern Sie, wie man dieses Maß verwenden kann, um Struktur-Alignments zu erstellen.

<sup>2</sup><http://dinamelt.bioinfo.rpi.edu/download.php>, getestet unter Ubuntu 10.4. Weiterhin steht ein Web-Server zu Verfügung.