

Biopython – Sequence Objects

Doris Steinbauer und Klaus Leitner

Einleitung und Motivation

Eine der grundlegendsten Arbeiten in der Bioinformatik ist das Analysieren von Sequenzen. Diese ermöglichen es Prozesse in einer Zelle abzubilden. Aus biotechnologischer Sicht ist es essentiell für die Produktion eines rekombinanten Proteins, das Genom einzelner Regionen eines Organismus zu verändern. Modifikationen im Genom können bioinformatisch aufbereitet werden. Das kann das Löschen oder Hinzufügen von Basen beinhalten. Um dies zu ermöglichen, bietet sich unter anderem das Python-Modul Bio.Seq und insbesondere die Klasse Seq an. Das Seq Objekt verhält sich wie ein Python String, und es können Python build-in Methoden für Strings angewendet werden. Darüber hinaus erzeugt Seq ein immutable Objekt und ermöglicht das schnelle und einfache Editieren von biologischen Sequenzen durch spezifische Methoden. Im folgenden Abschnitt werden die Anwendungsmöglichkeiten für das Seq Objekt näher erläutert.

Seq Object

```
from Bio.Seq import Seq
```

Immutable

Wie Strings sind auch Seq Objekte immutable: `my_seq = Seq("AGTACACTGGT")`

Kann zu mutable umgewandelt werden: `mutable_seq = MutableSeq(my_seq)`

String Eigenschaften

Seq Objekte lassen sich wie Strings bearbeiten: `my_seq[0:5]`

Biologische Anwendungen

```
my_seq.reverse_complement()
```

```
my_seq.transcribe(), my_seq.translate()
```

```
my_seq.index("CAC"), my_seq.find("CAC") # Gibt die Zahl 4 aus
```

Zusammenfassung

Die Klasse Seq ermöglicht effizientes Bearbeiten von DNA, RNA und Proteinsequenzen.