**M2: Die Genbank NCBI – Die Bibliothek für molekulare Sequenzdaten**

Die Genbank des NCBI (National Center for Biotechnology Information) gehört mit der DDBJ (DNA Data Bank of Japan) und dem EMBL (European Molecular Biology Laboratory) zu den größten international kooperierenden und öffentlichen molekularen Datenbanken (vgl. Knoop & Müller 2006, S. 74). Vergleichbar mit einer Bibliothek werden in diesen Datenbanken anstatt Bücher Nukleotidsequenzen verschiedener Organismen gesammelt, verwaltet und der Öffentlichkeit zugänglich gemacht. Werden von einem Molekularbiologen ein oder mehrere Gene einer Art sequenziert, so werden diese Sequenzdaten mit einer spezifischen Akzessionsnummer in den öffentlichen Datenbanken hochgeladen, so dass Anwender dieser Datenbanken gezielt nach diesen Sequenzen suchen und diese nutzen können. Die Tabellen 1 und 2 enthalten die Akzessionsnummern von einerseits mitochondrialen und anderseits nuclearen DNA-Sequenzen der Chamäleonarten sowie der als Außengruppe dienende Madagaskar-Legune für die es die Verwandtschaftsverhältnisse zu überprüfen gilt. Es hat sich erwiesen, dass das mitochondriale ND2-Gen (NADH dehydrogenase subunit 2), dessen codierendes Protein als Untereinheit des NADH-Dehydrogenase-Komplexes in der Atmungskette beteiligt ist, und das nukleare RAG-1-Gen (Rekombinations aktivierendes Gen 1), dessen Genprodukt eine zentrale Rolle bei der Immunantwort spielt, leistungsstarke Marker für phylogenetische Untersuchungen bei Wirbeltieren sind. Die Außengruppe (engl. *outgroup*) dient bei der Analyse als Vergleichsmaßstab. Dabei handelt es sich um eine Art, die nahe zur untersuchenden Artengruppe (Innengruppe, engl. *Ingroup*) verwandt ist, sich dennoch stammesgeschichtlich früher abgespalten hat als die betrachtete Innengruppe (vgl. Knoop & Müller 2006, S. 56)

Tabelle 1: Akzessionsnummern für die in der Gendatenbank NCBI eingetragenen Sequenzen des mitochondrialen NADH-Dehydrogenase Gens (ND 2) der zu untersuchenden Chamäleonarten sowie der Außengruppe.

|  |  |  |
| --- | --- | --- |
| **Art** | **Vorkommen** | **Akzessionsnummer** |
| *Brookesia nasus* | Madagaskar | FJ975158 |
| *Brookesia superciliaris* | Madagaskar | FJ975188 |
| *Calumma parsonii* | Madagaskar | AF448742 |
| *Calumma nasutum* | Madagaskar | AF448740 |
| *Furcifer lateralis* | Madagaskar | AF448768 |
| *Furcifer oustaleti* | Madagaskar | AF448769 |
| *Chamaeleo chamaeleon* | Afrika | AF448745 |
| *Chamaeleo calyptratus* | Afrika | AF448744 |
| *Trioceros quadricornis*  *(alt: Chamaeleo quadricornis)* | Afrika | AF448759 |
| **Außengruppe:** | | |
| *Oplurus cuvieri* | Madagaskar | OCU82685 |

Tabelle 2: Akzessionsnummern für die in der Gendatenbank NCBI eingetragenen Sequenzen des nuclearen Rekombination aktivierenden Gens (RAG-1) der zu untersuchenden Chamäleonarten sowie der Außengruppe.

|  |  |  |
| --- | --- | --- |
| **Art** | **Vorkommen** | **Akzessionsnummer** |
| *Brookesia nasus* | Madagaskar | FJ984204 |
| *Brookesia superciliaris* | Madagaskar | FJ984232 |
| *Calumma parsonii* | Madagaskar | FJ984190 |
| *Calumma nasutum* | Madagaskar | HQ130637 |
| *Furcifer lateralis* | Madagaskar | FJ984201 |
| *Furcifer oustaleti* | Madagaskar | FJ984200 |
| *Chamaeleo chamaeleon* | Afrika | FJ984184 |
| *Chamaeleo calyptratus* | Afrika | nicht vorhanden |
| *Trioceros quadricornis*  *(*alt: *Chamaeleo quadricornis)* | Afrika | FJ984188 |
| **Außengruppe:** | | |
| *Oplurus cuvieri* | Madagaskar | AY662601 |

**Anleitung: Wie findet man die zu untersuchende Nukleotidsequenz in NCBI?**

**Schritt 1:** Öffnen Sie die Gendatenbank über den folgenden Link: <http://www.ncbi.nlm.nih.gov/>

**Schritt 2:** Um die Suche auf die Nukleotidsequenz-Datenbank einzuschränken, wählen Sie auf der Startseite im Ausklappmenü des Suchfensters „Nucleotide“ an. Geben Sie im Suchfenster die Akzessionsnummer (Tabelle 1 bei der Analyse der mitochondrialen DNA bzw. Tabelle 2 bei der Analyse der nuclearen DNA) der zu untersuchenden Art ein, z.B. „FJ975158“ für *Brookesia nasus* (Abbildung 1).

**Schritt 3:** Es öffnet sich der Datenbankeintrag mit der Nukleotidsequenz zu dem gesuchten Gen. Zusätzlich lassen sich Datenbankeinträge zur Taxonomie, Publikation und Proteinübersetzung über Hyperlinks aufrufen (Abbildung 2.1 & 2.2.).

Abbildung 1: Startseite des NCBI. Gezeigt ist erstens die Anwahl der Nukleotidsequenz-Datenbank über die Auswahl „Nucleotide“ und zweites die Eingabe der Akzessionsnummer von *Brookesia nasus* in das Suchfenster.

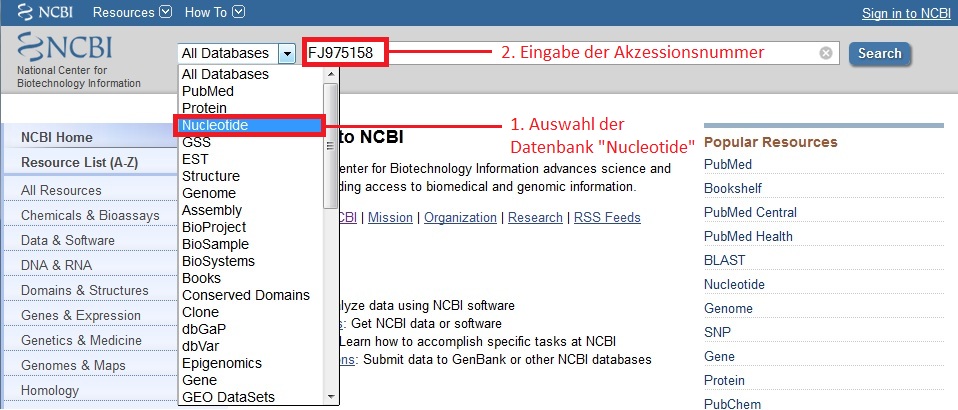


Abbildung 2.1: Exemplarischer Datenbankeintrag in NCBI des mitochondrialen NADH-Dehydrogenase Gens von *Brookesia nasus*. Gekennzeichnet sind die Angaben zur Taxonomie und Literaturreferenz.

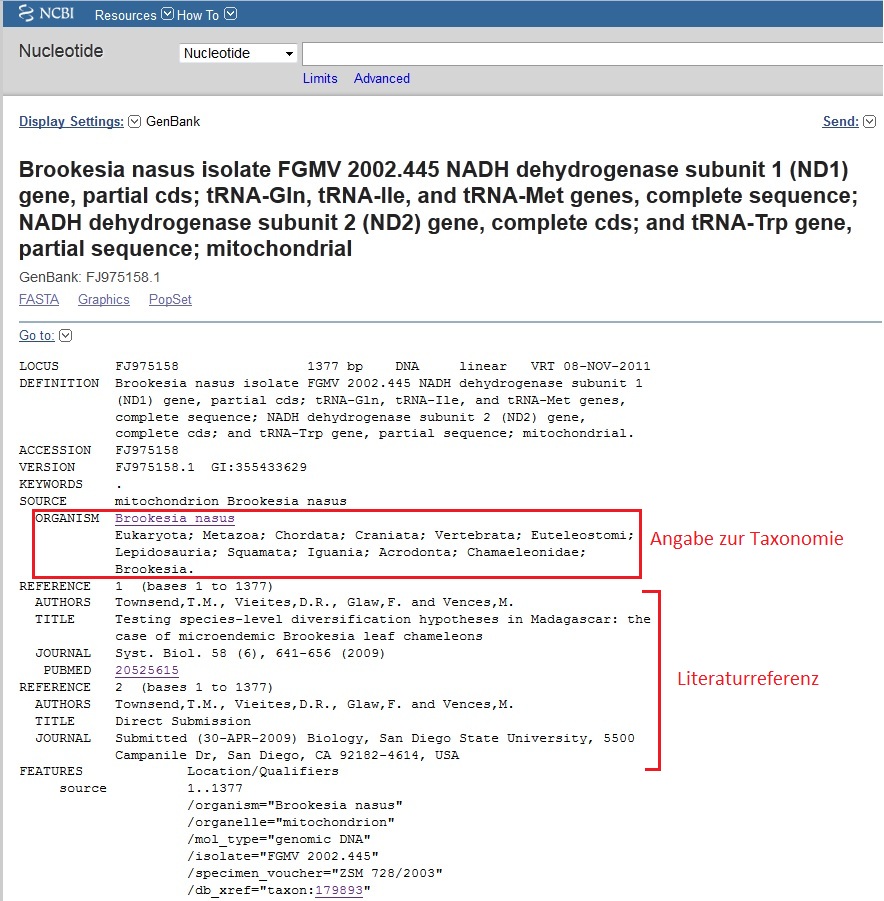
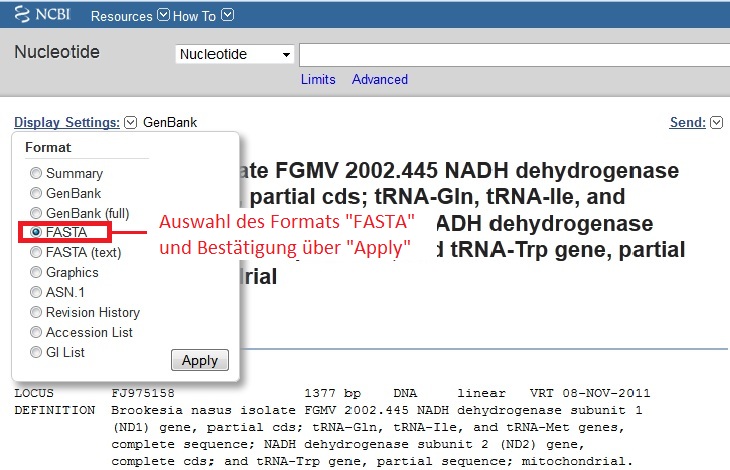


Abbildung 2.2: Exemplarischer Datenbankeintrag in NCBI des mitochondrialen NADH-Dehydrogenase Gens von *Brookesia nasus*. Gekennzeichnet sind die Angaben zum Gen und dessen Produkt sowie zur Aminosäure- und Nukleotidsequenz.



**Schritt 4:** Für die weitere Bearbeitung der Nucleotidsequenz mit dem Analyseprogramm MEGA ist eine Umwandlung des GenBank-Dateiformats in ein FASTA-Format notwendig. Dazu wird über das Ausklappmenü von „Display Settings“ das Format „FASTA“ angewählt (Abbildung 3).

Abbildung 3: Umwandlung der GenBank-Datei in eine FASTA-Datei.



**Schritt 5:** Die Einzelsequenz im FASTA-Dateiformat ließe sich nun über das Auswahlmenü von „Send“ durch die Anwahl „File“ als FASTA-Format herunterladen und auf dem eigenen Rechner abspeichern (Abbildung 4.1). Wenn man mehrere Einzelsequenzen miteinander vergleichen möchte, wie es bei der Analyse der Chamäleon-Gensequenzen der Fall ist, dann bietet es sich an die Einzelsequenzen zunächst in einer temporären Zwischenablage über das Auswahlmenü von „Send“ durch die Anwahl „Clipboard“ zu speichern (Abbildung 4.2) und anschließend als zusammengefasste Datei mit allen Sequenzen herunterzuladen (siehe nachfolgende Schritte).

Abbildung 4.1: Download einer Einzelsequenz als FASTA-Datei.

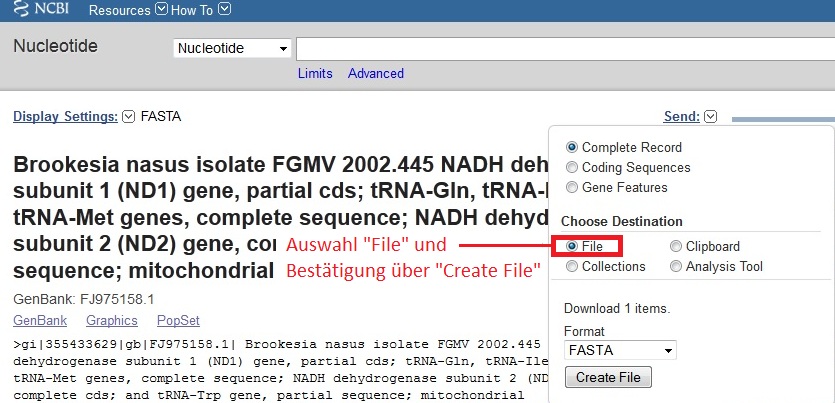
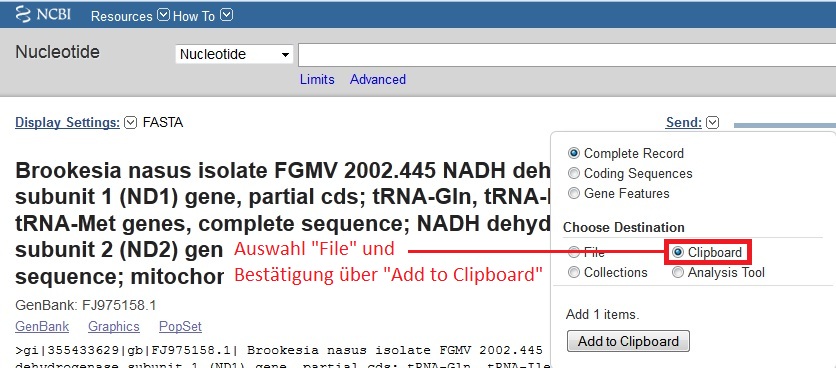


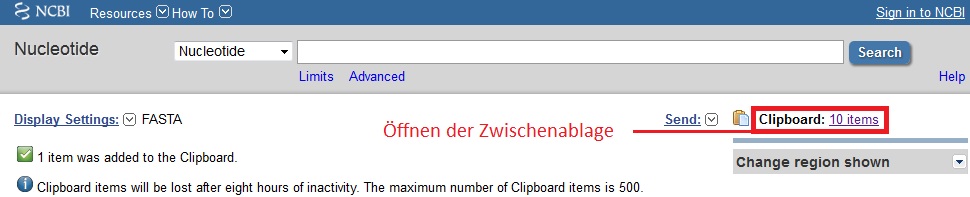
Abbildung 4.2: Hinzufügen der Einzelsequenz in die Zwischenablage (Clipboard).



**Schritt 6:** Um die Nukleotidsequenz für die nächste Art, *Brookesia superciliaris*, zu erhalten (Tabelle 1), werden die Schritte 2 bis 5 wiederholt und die erhaltene Einzelsequenz ebenfalls in die Zwischenablage kopiert. Für die weiteren in der Tabelle 1 bzw. 2 stehenden Arten wird genauso verfahren.

**Schritt 7:** Wurden alle Einzelsequenzen der zu untersuchenden Arten in die Zwischenablage kopiert, lassen sich diese durch Öffnen der Zwischenablage über den Hyperlink „Clipboard: X Items“ (X = Anzahl der in der Zwischenablage befindlichen Sequenzen) anzeigen (Abbildung 5).

Abbildung 5: Öffnen der Zwischenablage „Clipboard“.



**Schritt 8:** Die Einzelsequenzen, die in einer zusammengefassten Datei heruntergeladen werden sollen, lassen sich über Häkchen-Setzung markieren. Die markierten Sequenzen können dann über das Auswahlmenü von „Send“ durch die Anwahl „File“ als zusammengefasste FASTA-Datei heruntergeladen (Abbildung 6) und für die weitere Analyse mit dem Programm MEGA abgespeichert werden.

Abbildung 6: Herunterladen der zusammengefassten FASTA-Datei mit den Einzelsequenzen aus der Zwischenablage.

