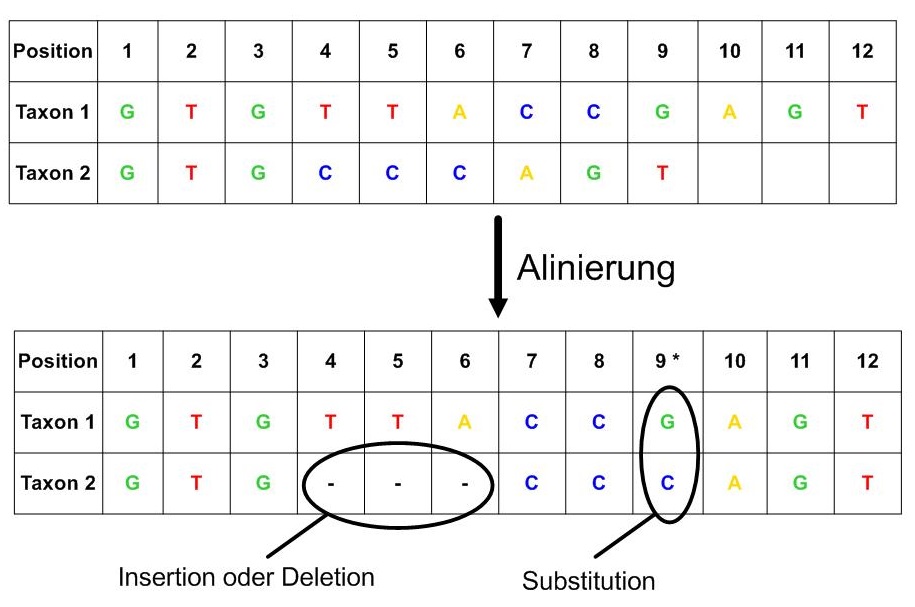
**M3: Multiple Sequenz-Alinierung mit MEGA – Der Vergleich von mehreren Nukleotidsequenzen**

Das Programm MEGA (Molecular Evolutionary Genetics Analysis) ist ein in der Evolutionsforschung weit verbreitetes Werkzeug, um nach der Bestimmung der Nukleotid-Abfolge (DNA-Sequenzierung), ein sogenanntes Sequenzalignment zu erstellen. Bei dieser Methode der Alinierung werden vergleichbare Positionen in den Nukleotid- oder Aminosäuresequenzen der untersuchten Arten untereinander angeordnet und damit eine Matrix erstellt. Die besondere Anordnung ermöglicht es Substitutionen (Nukleotidaustausche), Deletionen (Nukleotidverluste) oder Insertionen (Nukleotideinschübe), die an den Sequenzen stattfanden, ausfindig zu machen. Deletionen und Insertionen werden durch das Einfügen von Lücken (*gaps*) an der betreffenden Sequenzposition kenntlich gemacht (vgl. Knoop & Müller 2006, S. 67).

Beispielsweise hat die Alinierung des Sequenzpaares GTGTTACCGAGT/GTGCCCAGT zur Folge, dass durch das Einfügen von Lücken (-) in den Positionen 4 bis 6 der zweiten Sequenz homologe Sequenzabschnitte in Übereinstimmung gebracht werden. Diese Verschiebung der Nukleotidabfolge könnte durch Insertion von „TTA“ in der ersten Sequenz oder durch Deletion in der zweien Sequenz hervorgerufen worden sein. An der Position 9 erfolgte eine Substitution, gekennzeichnet durch ein Sternchen-Symbol (\*).

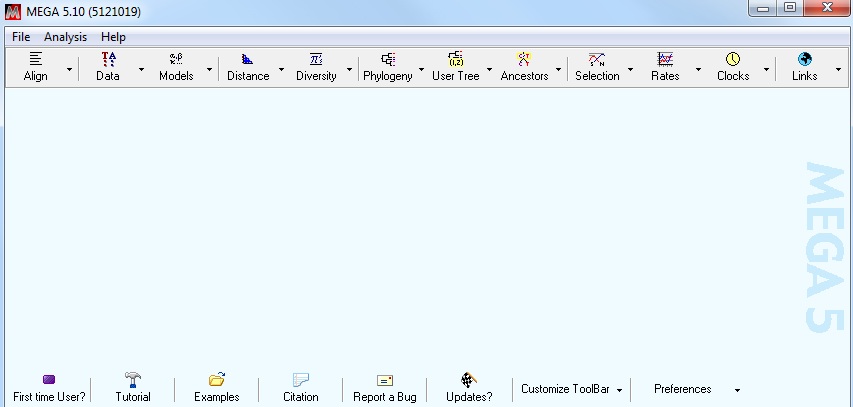


Ergebnis der Alinierung ist das Alignment, welches den Sequenzabgleich ermöglicht, um Aussagen zu den Verwandtschaftsverhältnissen treffen zu können. Je näher zwei Arten miteinander verwandt sind, desto weniger Unterschiede sind in den Positionen der betrachteten Sequenz anzutreffen. Im Gegensatz dazu sind bei weiter entfernt verwandten Arten vermehrt Stellen mit verschiedenen Nukleotiden oder Unterschiede in der Sequenzlänge zu lokalisieren (vgl. Knoop & Müller 2006, S. 67).

**Anleitung: Wie werden die zu untersuchenden Nukleotidsequenzen miteinander verglichen?**

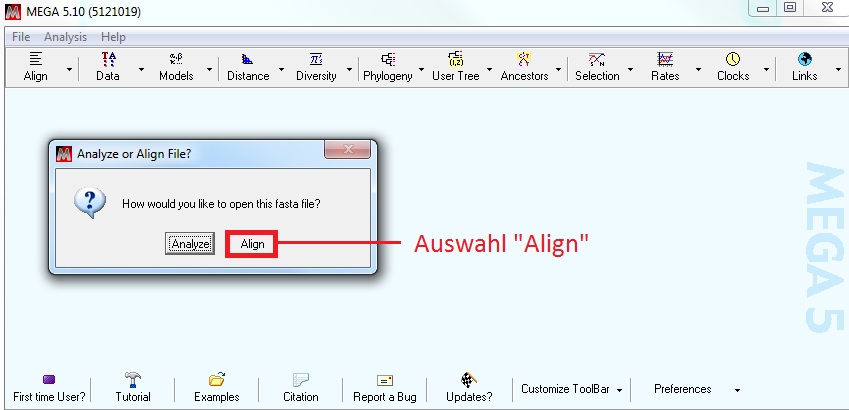
**Schritt 1:** Starten Sie das Programm MEGA, wobei sich das in Abbildung 1 gezeigte Hauptfenster von MEGA öffnet.

Abbildung 1: Hauptseite des Programms MEGA.



**Schritt 2:** Über das Menü „File“ kann unter „Open A File/Session…“ die bereits gespeicherte FASTA-Datei in das Programm geladen werden. Bei der Nachfrage „How would you like to open this fasta file?” wird die Alinierungsfunktion über „Align” angewählt (Abbildung 2).

Abbildung 2: Auswahl der Alinierungsfunktion beim Öffnen der gespeicherten FASTA-Datei.



**Schritt 3:** Die FASTA-Datei wird in einem separaten Alignment-Explorer geöffnet, wobei die Sequenzen der untersuchten Arten untereinander in Form einer Matrix dargestellt werden (Abbildung 3). Angezeigt werden die Sequenznamen, die DNA-Sequenzen in Form der Basenabfolge mit unterschiedlicher Farbmarkierung (A = Adenin, G = Guanin, T = Thymin und C = Cytosin) und die aktuelle Basenposition. Der Sequenzname ist durch ein Doppelklick über die linke Maustaste veränderbar und die Reihenfolge der einzelnen Sequenzen durch Anklicken des Nukleotidnamens beliebig verschiebbar. Des Weiteren erscheint die geöffnete FASTA-Datei auch im Hauptfenster von MEGA über das in Abbildung 4 markierte Symbol.

Abbildung 3: Anzeige des Alignment-Explorers nach der Transformation der Nukleotidsequenzen.

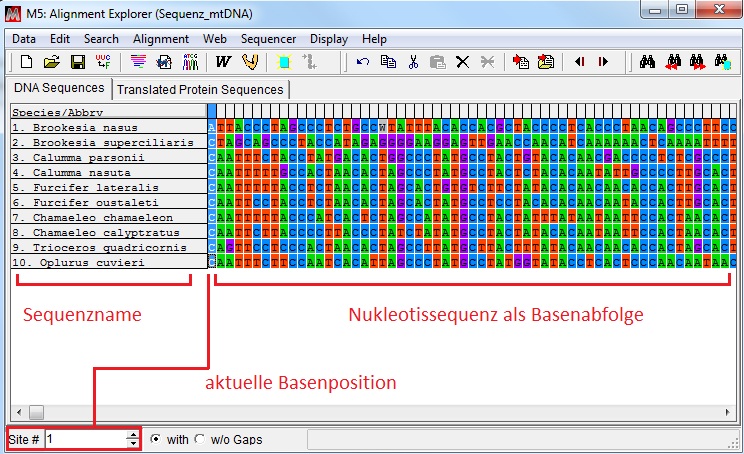
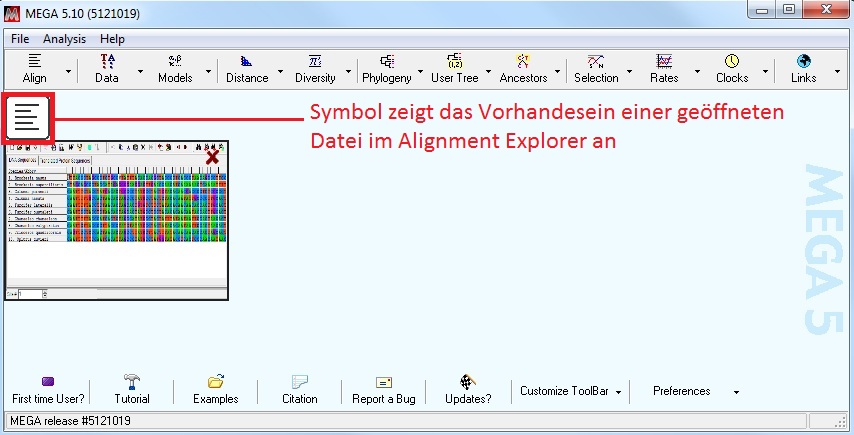


Abbildung 4: Anzeige einer geöffneten Datei des Alignment-Explorers im Hauptfenster von MEGA.



**Schritt 4:** Die Nukleotidsequenz lässt sich über die Anwahl der Registerkarte „Translated Protein Sequences“ in den Proteincode übersetzen (Abbildung 5). Dabei öffnet sich ein Fenster mit der Nachfrage „The current genetic code is: standard. Is it correct?“. Bei Verwendung von Nukleotidsequenzen aus nuklearer DNA, wird die Frage durch „Yes“ bestätigt, da die nukleare DNA in dem Programm MEGA als Standard gesetzt wird. Werden jedoch Nukleotidsequenzen aus mitochondrialer DNA verwendet, wird die Frage durch „No“ verneint und die Einstellung durch die Häkchen-Setzung bei „Vertebrate Mitochondrial“ und das Anklicken von „Ok“ geändert (Abbildung 6).

Abbildung 5: Auswahl bei der Proteinübersetzung der DNA-Sequenz.

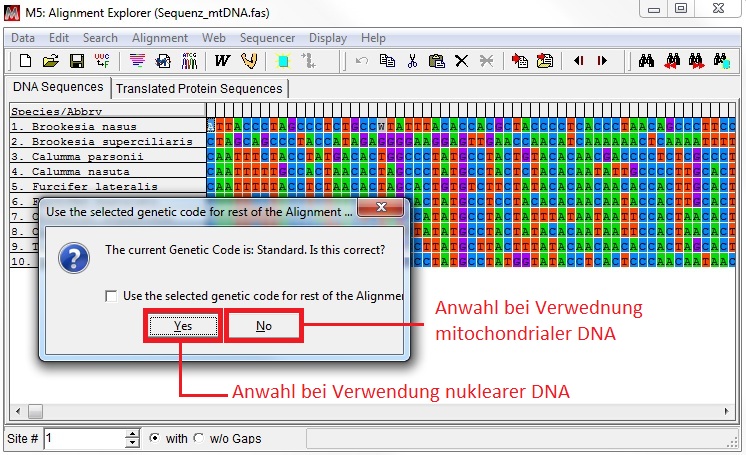
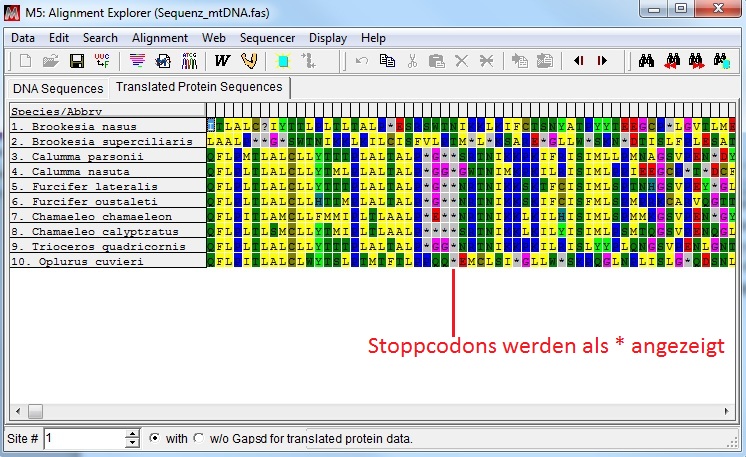


Abbildung 6: Änderung der „Standard“-Voreinstellung für die Proteinübersetzung bei Verwendung mitochondrialer DNA.



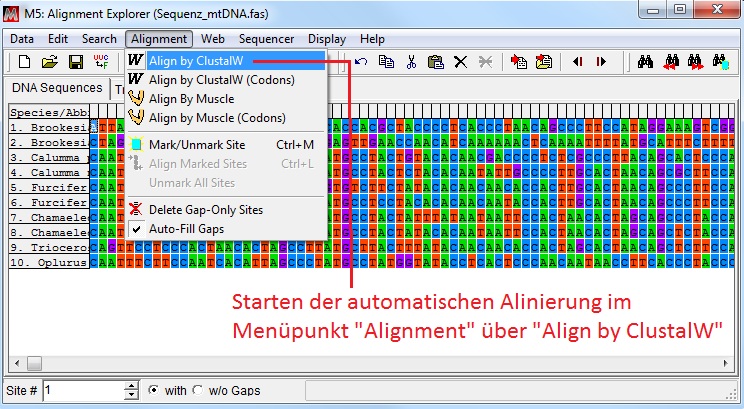
**Schritt 6:** Die Proteinübersetzung wird als Ein-Buchstaben-Code, z.B. A = Alanin, mit unterschiedlicher Farbmarkierung und Stoppcodons, die durch ein Sternchen ‚\*‘ gekennzeichnet sind, angezeigt (Abbildung 7). Unbekannte Positionen werden durch ein Fragezeichen ‚?‘ markiert. Der Wechsel in die Nukleotidsequenz-Ansicht erfolgt durch die Anwahl der Registerkarte „DNA Sequences“.

Abbildung 7: Anzeige der Proteinübersetzung im Alignment-Explorer.



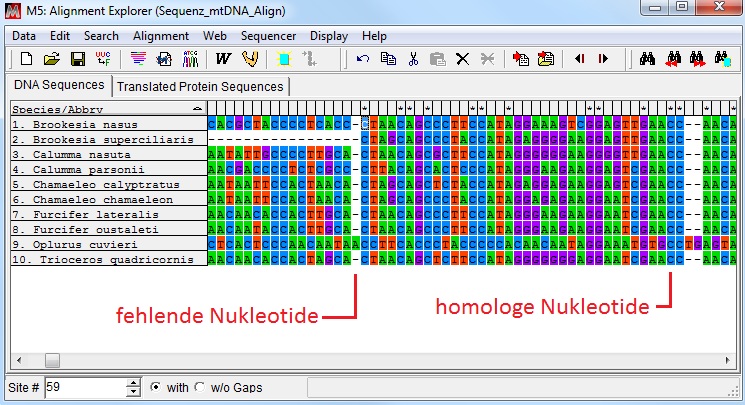
**Schritt 7:** Die automatische Alinierung der Nukleotid-Sequenzen wird im Menüpunkt „Alignment“ über die Auswahl „Align by Clustal W“ gestartet (Abbildung 8). Der Frage „Nothing selected for alignment. Select all?“ wird mit „Ok“ zugestimmt, da alle Nukleotidsequenzen, die in den Alignment-Explorer geladen wurden, aliniert werden sollen. Die Voreinstellungen des sich darauf öffnenden Fensters „ClustalW Parameters“ werden so beibehalten und mit „Ok“ bestätigt, wodurch die Berechnung des Alignments gestartet wird.

Abbildung 8: Starten der automatischen Alinierung.



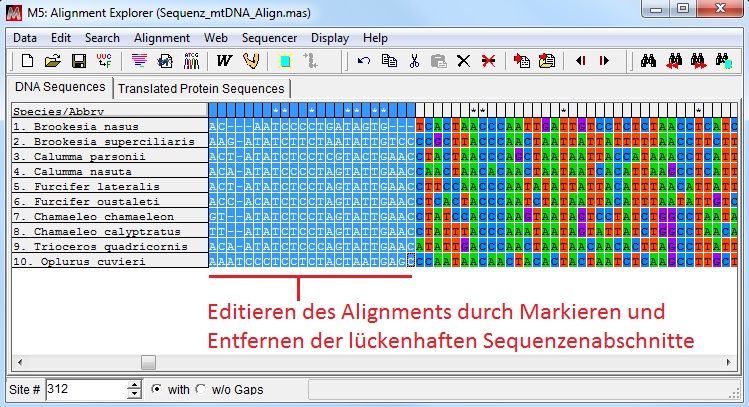
**Schritt 8:** Nach einiger Rechenzeit wird das Alignment so dargestellt, dass homologe Nukleotide übereinander stehen und fehlende Nukleotide im Alignment mit einem Bindestrich als Lücken gekennzeichnet sind (Abbildung 9).

Abbildung 9: Darstellung des Alignments nach der automatischen Alinierung.



**Schritt 9:** Als nächster Schritt erfolgt das manuelle Editieren. Für die nachfolgende Analyse und Stammbaumrekonstruktion sind Lücken auszuschließen, da nur lückenlose DNA-Abschnitte miteinander vergleichbar sind. Um dies zu gewährleisten editieren Sie das Alignment so, dass Sie die lückenhaften Abschnitte an denen beiden Enden des Alignments markieren und mit der Entf-Taste der Tastatur entfernen (Abbildung 10). Das Ergebnis ist ein lückenloses Alignment.

Abbildung 10: Manuelles Editieren des lückenhaften Alignments.



**Schritt 10:** Zur Überprüfung, ob das Alignment auch sinnvoll editiert wurde, wird wieder in die Ansicht der Proteinübersetzung über die Registerkarte „Translated Protein Sequences“ gewechselt. Sind in dem Alignment in der Proteinübersetzung viele Stoppcodons verstreut (Abbildung 11), muss in der Nukleotidansicht des Alignments die erste oder zweite Nukleotidspalte durch Markieren der Spalte/n und betätigen der Entf-Taste entfernt werden, so dass sich dann das Lesemuster um ein bzw. zwei Positionen verschiebt und ein stoppcodonfreies Alignment in der Proteinansicht entsteht (Abbildung 12). Speichern Sie anschließend das editierte Alignment als mas-Dateiformat ab.

Abbildung 11: Alignment mit Stoppcodons in der Proteinübersetzung.

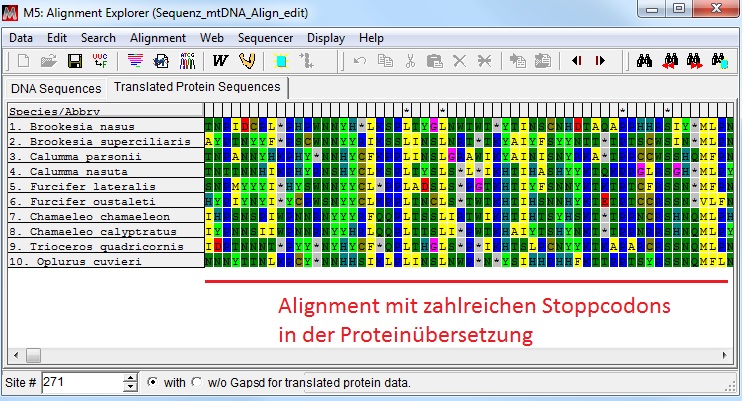


Abbildung 12: Alignment ohne Stoppcodons in der Proteinübersetzung nach Kürzung um eine Nukleotidposition.

