

Praktikumsskript

DNA-Sequenzierung

015.06.2015 – 26.06.2015

Praktikumsraum 244

SICHERHEITSHINWEISE

1. Allgemeines

Essen, Trinken, Schminken und Rauchen ist in den Kursräumen und Labors untersagt.

2. Schutzkleidung

Während des Praktikums sind Schutzkittel zu tragen. Beim Umgang mit gefährlichen Substanzen müssen Schutzhandschuhe getragen werden. Im Labor dürfen keine Kleidung, Rucksäcke, Einkaufstaschen usw. herumliegen (Spinde in den Umkleieräumen benutzen).

3. Chemikalien

Alle Chemikalienflaschen müssen mit einem Klebeetikett oder lösungsmittelfesten Stift eindeutig beschriftet sein; die Beschriftung sollte dauerhaft erkenntlich sein. Die Angaben auf den Etiketten müssen der Gefahrstoffverordnung entsprechen (Name und Gefahrensymbol). Niemals mit dem Mund pipettieren !

4. Abfälle

Alle Chemikalienabfälle werden in entsprechend gekennzeichneten Gefäßen gesammelt. Alle aus den Versuchen entstehenden Bakterienkulturen sind nach den Versuchen in den dafür vorgesehenen Abfallbehältern zu sammeln und zu autoklavieren. Das gilt auch für alle Einweg-Laborartikel wie Pipettenspitzen und Reaktionsgefäße. Glasgeräte sind nach Gebrauch zu autoklavieren.

5. Fluchtwege

Bei Feuer oder anderen Gefahren muß der Raum sofort verlassen werden. Der "NOTAUS"-Schalter ist zu drücken. Das Gebäude wird über die Flure und das nächstgelegene Treppenhaus verlassen. Fahrstühle dürfen in keinem Fall benutzt werden.

6. Gentechnik

Die im Kursraum aushängende Betriebsanweisung für gentechnische Laboratorien der Sicherheitsstufe **S1** ist zu beachten.

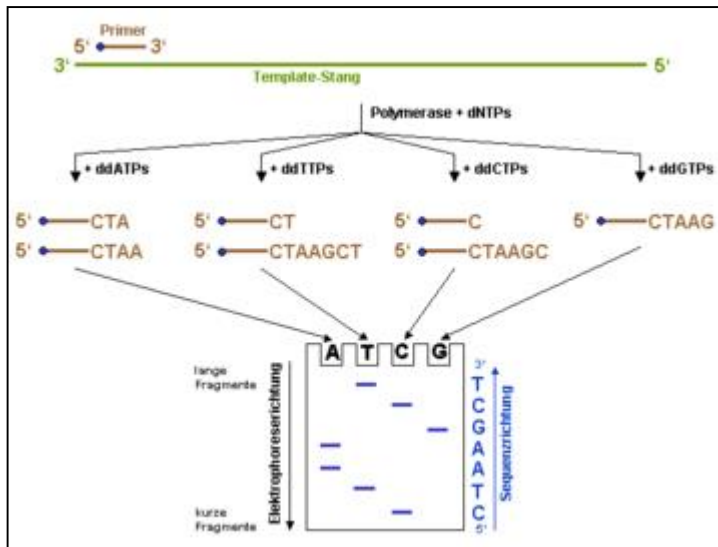
ZIEL

Im Praktikum soll in einem aufeinander aufbauenden Versuchsprogramm die DNA-Sequenzierung unter Verwendung unterschiedlicher Ausgangs-DNA erlernt werden. Dazu gehört nach Isolation von Plasmid-DNA aus rekombinanten Bakterienklonen die Sequenzierung der Plasmide mit universellen Vektorprimern (M13). Zum anderen sollen durch den Einsatz spezifischer sequenzinterner Primer fehlende Sequenzbereiche durch PCR amplifiziert und sowohl direkt als auch unter Verwendung der Plasmid-DNA (Primerwalking) sequenziert werden.

EINLEITUNG

Zur Bestimmung der Nukleotidabfolge wird die als Kettenabbruch-Synthese bekannte nichtradioaktive Nachweismethode nach Sanger *et al.* (1977) verwendet.

Dabei wird eine Matrizen-DNA in einer *cycle-sequencing*-PCR unter Einbau von 2',3'-Didesoxynukleotiden (ddNTP) amplifiziert. In vier sonst gleichen Ansätzen wird je eine der vier Basen zum Teil als Didesoxynukleotid (ddNTP) zugegeben. Diese „Kettenabbruch-Nukleotide“ besitzen keine 3'-Hydroxygruppe. Werden sie in den neusynthetisierten Strang eingebaut, ist eine DNA-Verlängerung durch eine DNA-Polymerase nicht mehr möglich. In der Folge entstehen DNA-Fragmente unterschiedlicher Länge, die in jedem Ansatz stets mit dem gleichen ddNTP enden. Da die ddNTPs zufällig eingebaut werden, bricht die DNA-Synthese an verschiedenen Stellen ab.



Prinzip der DNA-Sequenzierung nach der Dideoxy-Methode.

Der Einbau eines ddNTPs in den neusynthetisierten DNA-Strang führt zum Abbruch der Polymerisationsreaktion. Die blauen Punkte am 5'-Ende des Primers stellen eine Markierung dar (z.B. eine fluoreszierende Gruppe), mittels der die Syntheseprodukte später im Gel sichtbar gemacht werden können. Bei der Sequenzierung über Kapillar-Gelelektrophorese werden mit Fluoreszenz-Farbstoff markierte Nukleotide zur Polymerisationsreaktion eingesetzt.

Heute werden vor allem mit Fluoreszenz-Farbstoffen markierte Didesoxynukleotide eingesetzt. Jedes der vier ddNTPs wird mit einem unterschiedlichen Farbstoff gekoppelt. Diese Modifikation erlaubt es, alle vier Didesoxynukleotide in einem Reaktionsgefäß zuzugeben. Die entstehenden Abbruchprodukte werden durch Kapillar-Gelelektrophorese aufgetrennt und mit Hilfe eines Lasers zur Fluoreszenz angeregt. Die ddNTPs am Ende jedes DNA-Fragmentes senden Fluoreszenzsignale unterschiedlicher Wellenlänge und können so von einem Detektor erkannt werden. Die Abfolge der Farbsignale, die am Detektor erscheinen, gibt direkt die Nukleotidfolge des sequenzierten DNA-Stranges wieder.

AUFGABE

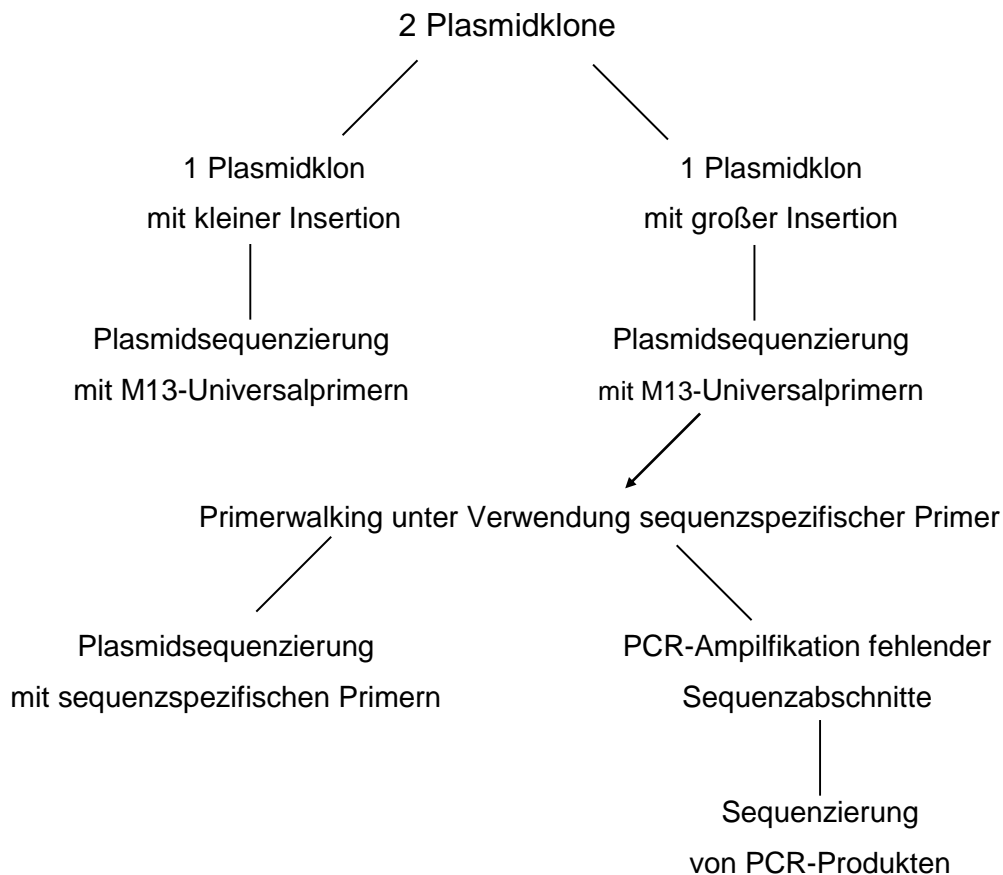
Die Aufgabe im Praktikum besteht in der Sequenzanalyse von Plasmidklonen, die im Rahmen von Forschungsarbeiten am Lehrstuhl erzeugt wurden.

Jede Gruppe erhält zwei rekombinante Bakterienklone, die nach der Plasmid-Präparation eingehend durch Agarosegelelektrophorese und Insert-PCR charakterisiert werden sollen. Homologievergleiche sollen mit Sequenzen aus der EMBL-Datenbank erfolgen.

Diese Aufgabe schließt folgende Arbeitsschritte ein:

1. Kultivierung rekombinanter Bakterienklone
2. Isolation der Plasmid-DNA
3. Charakterisierung der rekombinanten Plasmide
4. Sequenzierung von Plasmid-DNA
5. DNA-Sequenzanalyse
6. Identifizierung der Sequenzen durch Datenbankrecherchen
7. Direktes Sequenzieren der Plasmid-DNA (Primerwalking) mit spezifischen Primern

8. Amplifikation fehlender Sequenzbereiche durch PCR
9. Sequenzierung von PCR-Produkten
10. Contig-Erstellung mit einzelnen Teilsequenzen zu einer Gesamtsequenz
11. Homologievergleich und Divergenzanalyse
12. Präsentation der Ergebnisse



Ablaufschema der experimentellen Arbeiten

Das Praktikum ist so angelegt, dass Sie nach eingehender Unterweisung eine möglichst große Selbständigkeit bei der Durchführung der einzelnen Arbeitsschritte erlangen sollten. Anfallende Inkubations- und Wartezeiten sollten für Versuchsvorbereitung und die abschließende Präsentation der Ergebnisse am letzten Praktikumstag genutzt werden.

Die Methoden für die einzelnen Arbeitsschritte sind dem Handbuch „Molecular Cloning“ (Sambrook *et al.* 1989), Skripten vorhergehender Praktika am Lehrstuhl und den Produktanleitungen der verwendeten Kits zu entnehmen und werden vor jedem Versuch gemeinsam besprochen.

DNA-SEQUENZIERUNG

Das CEQ 8000 Genetic Analysis System von Beckman Coulter ermöglicht die voll automatisierte Sequenzierung von bis zu 96 markierten Proben im Mikrotiterplatten-Format. Dabei kann die Probe in einem Ansatz mit allen vier farbstoff-markierten Didesoxynukleotiden markiert werden. Jede Reihe mit je acht Proben wird automatisch denaturiert und durch Kapillarelektrophorese in acht 33 cm langen Kapillaren separiert.

Die optischen Signale werden durch laserinduzierte Fluoreszenz im jeweiligen Spektralbereich der einzelnen Farbstoffe erfasst und als Chromatogramm dargestellt.

DNA-SEQUENZANALYSE

Die Auswertung der Sequenzen erfolgt mit dem kostenfreien Programm „Staden Package“. Zuerst wird die Qualität der Sequenz ermittelt und der auswertbare Bereich definiert, um in der detaillierten Sequenzanalyse einen direkten Vergleich zwischen den Sequenzen zu ermöglichen. Weiterhin wird eine Bereinigung der Sequenzdaten von Vektorsequenzen (Vector Trimming) im Pregap-Menü und das Editieren der Sequenz unter Berücksichtigung des Chromatogramms im Gap-Menü realisiert. Hiermit ist zusätzlich das Zusammenfügen überlappender Sequenzen zu einer Gesamtsequenz (Contig) möglich.

Die überarbeiteten Nukleotidsequenzen werden über das Internet mit den Suchalgorithmen FASTA (<http://www.ebi.ac.uk/fasta33>) und BLAST (<http://www.ncbi.nlm.nih.gov/BLAST>) mit bereits bekannten Sequenzen aus der EMBL-Datenbank verglichen.

Homologie und Divergenz zwischen Sequenzen auf Nukleotid- und Proteinebene werden mit dem Programm „CLUSTAL W“ (<http://www.ebi.ac.uk/clustalw>) berechnet. Phylogenetische Analysen können mit dem frei zugänglichen Programmen MEGA (<http://www.megasoftware.net/>) und TreeView (<http://taxonomy.zoology.gla.ac.uk/rod/treeview.html>) durchgeführt werden.

AUSWERTUNG

Die aus dem Praktikum resultierenden Daten sollen nicht in Protokollform, sondern von jedem Kursteilnehmer/ jeder Gruppe in einer abschließenden Besprechung dargestellt werden.

Weiterhin ist die Präsentation verschiedener Sequenziermethoden in einem kurzen Vortrag geplant.

PURIFICATION PROTOCOL

Note

- Read **IMPORTANT NOTES** on p.3 before starting.
- All purification steps should be carried out at **room temperature**.
- All centrifugations should be carried out in a table-top microcentrifuge at **>12000 x g** (10 000-14 000 rpm, depending on the rotor type).

Use 1-5 ml of *E. coli* culture in LB media for purification of **high-copy** plasmids.

For **low-copy** plasmids use up to 10 ml of culture.

Step	Procedure
1	Resuspend the pelleted cells in 250 µl of the Resuspension Solution . Transfer the cell suspension to a microcentrifuge tube. The bacteria should be resuspended completely by vortexing or pipetting up and down until no cell clumps remain. Note: Ensure RNase A has been added to the Resuspension Solution (as described on p.3)
2	Add 250 µl of the Lysis Solution and mix thoroughly by inverting the tube 4-6 times until the solution becomes viscous and slightly clear. Note: Do not vortex to avoid shearing of chromosomal DNA. Do not incubate for more than 5 min to avoid denaturation of supercoiled plasmid DNA.
3	Add 350 µl of the Neutralization Solution and mix immediately and thoroughly by inverting the tube 4-6 times. Note: It is important to mix thoroughly and gently after the addition of the Neutralization Solution to avoid localized precipitation of bacterial cell debris. The neutralized bacterial lysate should become cloudy.
4	Centrifuge for 5 min to pellet cell debris and chromosomal DNA.
5	Transfer the supernatant to the supplied GeneJET spin column by decanting or pipetting. Avoid disturbing or transferring the white precipitate.

Step	Procedure
6	Centrifuge for 1 min. Discard the flow-through and place the column back into the same collection tube. Note: Do not add bleach to the flow-through, see p.7 for Safety Information.
7	Add 500 µl of the Wash Solution (diluted with ethanol prior to first use as described on p.3) to the GeneJET spin column. Centrifuge for 30-60 seconds and discard the flow-through. Place the column back into the same collection tube.
8	Repeat the wash procedure (step 7) using 500 µl of the Wash Solution .
9	Discard the flow-through and centrifuge for an additional 1 min to remove residual Wash Solution. This step is essential to avoid residual ethanol in plasmid preps.
10	Transfer the GeneJET spin column into a fresh 1.5 ml microcentrifuge tube (not included). Add 50 µl of the Elution Buffer to the center of GeneJET spin column membrane to elute the plasmid DNA. Take care not to contact the membrane with the pipette tip. Incubate for 2 min at room temperature and centrifuge for 2 min. Note: An additional elution step (optional) with Elution Buffer or water will recover residual DNA from the membrane and increase the overall yield by 10-20%. For elution of plasmids or cosmids >20 kb, prewarm Elution Buffer to 70° C before applying to silica membrane.
11	Discard the column and store the purified plasmid DNA at -20° C.

Material Required

Materials provided by Beckman Coulter:

- Dye Terminator Cycle Sequencing (DTCS) Quick Start Kit (P/N 608120):
 - DTCS Quick Start Master Mix
 - pUC18 Control Template (0.25µg/µL)
 - M13 -47 Sequencing Primer (1.6pmol/µL or 1.6µM)
 - Glycogen (20mg/mL)
 - Mineral Oil
 - Sample Loading Solution (SLS; see Appendix C for storage conditions)

Required materials not provided by Beckman Coulter:

- Molecular Biology Grade:
 - sterile dH₂O, 95% (v/v) ethanol/dH₂O, 70% (v/v) ethanol/dH₂O
- 3M Sodium Acetate pH5.2 - Sigma Cat # S 7899
- 100mM Na₂-EDTA pH8.0, prepared from 0.5 M Na₂-EDTA - Sigma Cat # E 7889)
- Sterile tubes - 0.5mL microfuge, 0.2mL thin wall thermal cycling tubes or plates
- Thermal cycler with heated lid

NOTICE TO PURCHASER: LIMITED LICENSE

The purchase price of this product includes a limited, non-transferable license under U.S. Patent 5,332,666; and claims in its foreign counterparts that correspond to processes for DNA sequence and fragment analysis, to use this product in DNA sequence and fragment analysis and related processes described in said patents for the internal research and development activities of the purchaser when this product is used in conjunction with an authorized DNA sequence analysis instrument for detection sequence fragments. No right to perform or offer commercial services of any kind, including without limitation reporting the results of purchaser's activities for a fee or other commercial consideration, is hereby granted, either by implication or estoppel. No other patents are licensed by purchase of this product, either by implication or estoppel. Further information relating to the purchase of licenses for DNA sequence and fragment analysis and other applications may be obtained by contacting the Director of Licensing at The Perkin-Elmer Corporation, Applied Biosystems Division, 850 Lincoln Centre Drive, Foster City, CA 94404.



GenomeLab™ Dye Terminator Cycle Sequencing with Quick Start Kit

Storage of the Cycle Sequencing kit must be in a -20°C non-frost free freezer.

1. Preparation of the DNA sequencing reaction*:

Prepare sequencing reaction in a 0.2mL thin-wall tube or microplate well. All reagents should be kept on ice while preparing the sequencing reactions and should be added in the order listed below.

dH ₂ O (to adjust total volume to 20µL)	0 - 9.5µL
DNA Template† (See Template Preparation)	0.5 - 10.0µL
Customer supplied or -47 Sequencing Primer (1.6pmol/µL or 1.6µM)	2.0µL
DTCS Quick Start Master Mix	8.0µL
TOTAL	20.0µL

†Use 0.5µL for pUC18 control template.

***Note: Mix reaction components thoroughly. Consolidate the liquid in the bottom of the tube or well by briefly centrifuging before thermal cycling.**

2. Thermal cycling program:

96°C	20 sec.
50°C	20 sec.
60°C	4 min.

for 30 cycles followed by holding at 4°C

3. Ethanol precipitation:

Precipitation in Individual Tubes

- Prepare a labeled, sterile 0.5mL microfuge tube for each sample.
- Prepare fresh Stop Solution/Glycogen mixture as follows (per sequencing reaction): 2µL of 3M Sodium Acetate (pH 5.2), 2µL of 100mM Na₂-EDTA (pH 8.0) and 1µL of 20mg/mL of glycogen (supplied with the kit). To each of the labeled tubes, add 5µL of the Stop Solution/Glycogen mixture.
- Transfer the sequencing reaction to the appropriately labeled 0.5mL microfuge tube and mix thoroughly.
- Add 60µL cold 95% (v/v) ethanol/dH₂O from -20°C freezer and mix thoroughly. Immediately centrifuge at 14,000 rpm at 4°C for 15 minutes. Carefully remove the supernatant with a micropipette (the pellet should be visible). Note: For multiple samples, always add the cold ethanol/dH₂O immediately before centrifugation.
- Rinse the pellet 2 times with 200µL 70% (v/v) ethanol/ dH₂O from -20°C freezer. For each rinse, centrifuge immediately at 14,000 rpm at 4°C for a minimum of 2 minutes. After centrifugation carefully remove all of the supernatant with a micropipette.
- Vacuum dry for 10 minutes (or until dry).
- Resuspend the sample in 40µL of the Sample Loading Solution (provided in the kit).

Precipitation in the Samples Plates

The Thermal Cycling and Ethanol precipitation can be performed in the sample plate. For instructions, see pages 19-20 in the Dye Terminator Cycle Sequencing Chemistry Protocol (P/N 718119).

4. Sample preparation for loading into the instrument:

- Transfer the resuspended samples to the appropriate wells of the sample plate (P/N 609801).
- Overlay each of the resuspended samples with one drop of light mineral oil (provided in the kit or Sigma Cat # M 5904).
- Load the sample plate into the instrument and start the desired method.

Template Preparation

1. DNA Template preparation:

Prepare sufficient template to allow for its accurate quantitation and purity verification. The quality of the DNA template will depend upon the procedure and the source of the DNA used. The following are the recommended protocols:

- QIAGEN QIAwell™ and QIAprep™ DNA isolation protocols (dsDNA and ssDNA)
- QIAGEN QIAquick™ PCR purification protocol (PCR products) *

Note: Determine the quality and quantity of template DNA by agarose gel electrophoresis.

2. DNA Template amount:

The amount of template DNA to use in the sequencing reaction depends on the form of the DNA (dsDNA plasmid, ssDNA M13, PCR product, etc.). It is important to accurately quantitate the amount (moles) of DNA when performing the DNA sequencing reaction (see formula and table below for details). The molar ratio of primer to template must be $\geq 40:1$. Listed below are the recommended amounts of DNA:

dsDNA	50-100fmol
ssDNA	25-50fmol
Purified PCR products	25-100fmol

The following table can be used to estimate DNA concentrations.

Table for estimating the dsDNA concentration.

Size (kilobase pairs)	ng for 25fmol	ng for 50fmol	ng for 100fmol
0.2	3.3	6.5	13
0.3	4.9	9.8	20
0.4	6.5	13	26
0.5	8.1	16	33
1.0	16	33	65
2.0	33	65	130
3.0	50	100	195
4.0	65	130	260
5.0	80	165	325
6.0	100	195	390
8.0	130	260	520
10.0	165	325	650
12.0	195	390	780
14.0	230	455	910
16.0	260	520	1040
18.0	295	585	1170
20.0	325	650	1300
48.5	790	1500*	1500*

For ssDNA, the values (ng) should be divided by 2.

*Use no more than 1.5µg of template DNA.

3. Template Pre-Heat Treatment

For certain plasmid DNA templates (not the included pUC18 control DNA), the following pre-heat treatment improves both signal strength and current stability.

- Dilute the template with water to the appropriate concentration.
- Heat the template at 96°C for 1 minute in a thermal cycler and then cool to room temperature before adding the remainder of the sequencing-reaction components.
- Do not add any other sequencing-reaction components to the plasmid template before performing this pre-heat treatment.
- If the raw data signal declines steeply when using this treatment, change the heating conditions to 86°C for 5 min. If the current is low or unstable following this treatment, increase the treatment to 96°C for 3 min.

*See the Detailed Dye Terminator Cycle Sequencing Chemistry Protocol (P/N 718119) for more information.

Appendix

Appendix A

Sequencing of PCR products

- All PCR products must be homogeneous in size as judged by gel electrophoresis.
- Purified PCR products
 - a) Remove unincorporated primers and dNTPs using QIAGEN QIAquick™ PCR purification system.
 - b) Use 25-100fmol of PCR product and 3.2pmoles of primer.

Unpurified PCR products

- a) For the original PCR amplification, the primer concentration should be 0.2µM or less, while the dNTP concentration should be 50µM or less.
- b) The amplification should be sufficient to produce a concentration of amplified fragment that is ≥ 10 fmol/µL
- c) Dilute this amplified fragment approximately 10 fold to result in a concentration of ≥ 1 fmol/µL.
- d) Use 5-15fmol of this diluted, unpurified PCR product and 3.2 pmoles of primer.

Appendix B

Sequencing of Large Templates

Adding 50-100fmol for large templates such as BACs, cosmids and PACs is impractical. The following procedure should be used when sequencing large templates.

1. Use 1.5µg of the template in 6µL of sterile, deionized water.
2. Pre-heat the template at 96°C for 1 minute.
3. Add the sequencing-reaction components as described in the standard protocol.
4. Cycle for 50 cycles using the appropriate cycling conditions for the primer being used.
5. Precipitate with ethanol, as previously described.

Appendix C

- Store the Sample Loading Solution in 350µL aliquots at -20°C in a non frost free freezer.
- Use each aliquot only once - do not freeze/thaw the Sample Loading Solution.

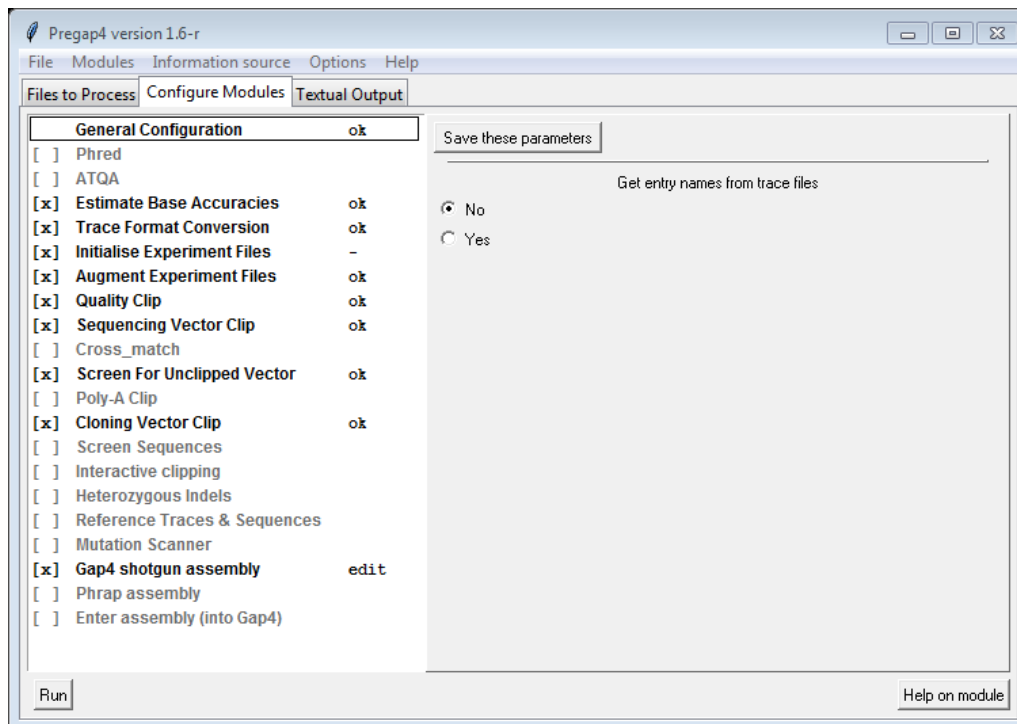
Staden Package

Vorbereitung

- Download: <http://staden.sourceforge.net/> (Gap4, Pregap4 – z.B. Version 1.6)
- Originaldaten aus Sequenzierung (*.scf) als Kopie abspeichern
- Zusammengehörige Klone in einem Ordner abspeichern
- Dateien werden durch Pregap modifiziert, neue Dateien im Ordner erstellt
- **Wichtig! Alle zusammen analysierte Dateien müssen in einem Ordner liegen.**

Sequenzen vorprozessieren in Pregap4

- Daten hochladen im Files to Process- Tab, Dateityp SCF(*.scf) beachten
- es können zusätzlich reine Sequenz als Textdatei (z.B. Referenz, Datenbankhomologe) als Dateityp *.pln geladen werden
- Sequenzen werden editiert, d. h. die Sequenz wird generell auf Qualität geprüft und ggf. eingekürzt und Vektorsequenzen können getrimmt werden
- dazu werden folgende Parameter im Fenster definiert:

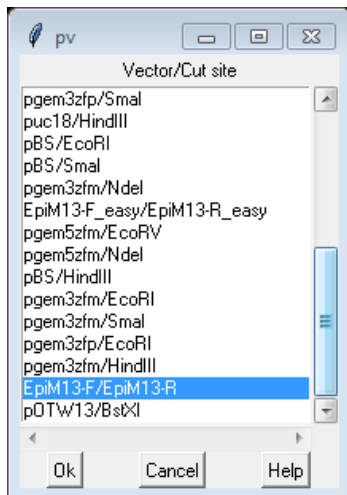
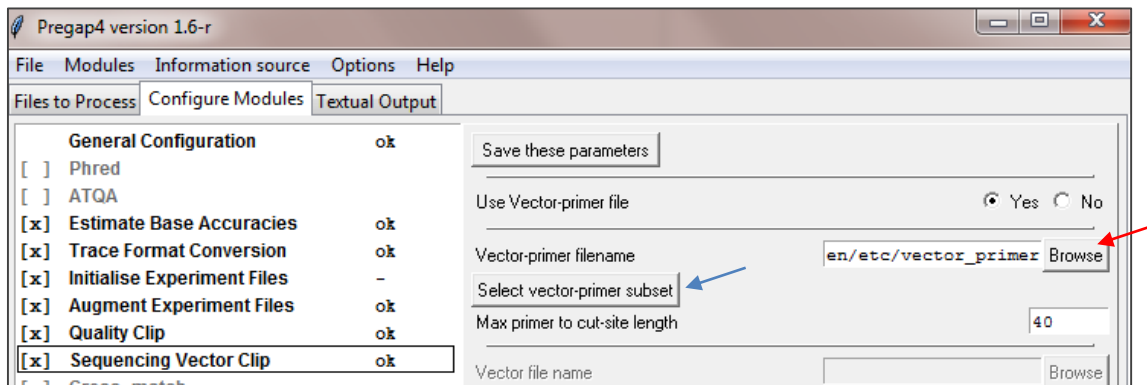


- Configure Modules-Tab:

- [x] Estimate Base Accuracies *bestimmt Zuverlässigkeit der Basen*
- [x] Trace Format Conversion
- [x] Initialise Experiment Files
- [x] Augment Experiment Files
- [x] Quality Clip *Abschneiden qualitativ schlechter Bereiche*
- [x] Cloning Vector Clip *Abschneiden der Vektorsequenzen*
 - Vektor/Primer Auswahl: Select vector-primer subset
 - dafür diesem Verlauf folgen (roter Pfeil):

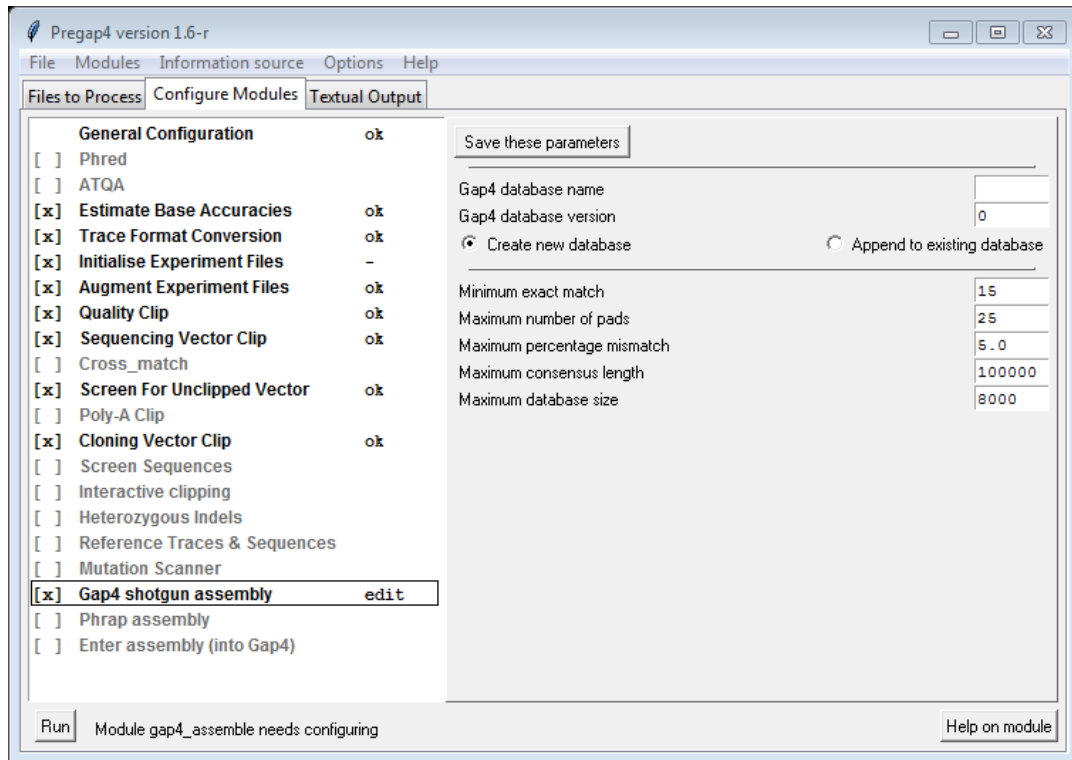
Staden Package

- C:\Program Files (x86)\Staden Package\share\staden\etc\vector_primer
- im Kurs werden zwei verschiedene Vektoren verwendet
 - 1.) pGEM T-Vektor für die Klonierung von PCR-Fragmenten
dafür Vector-primer subset: EpiM13-F/EpiM13-R auswählen (blauer Pfeil)
 - 2.) pUC18-Vektor für die Klonierung von Restriktionsfragmenten
dafür Vector-primer subset: puc18/BamHI auswählen (blauer Pfeil)



- [x] Screen For Undipped Vector *Abschneiden des Vektors überprüfen*
 - [x] Cloning Vector Clip *Abschneiden des Vektors*
 - dafür diesem Verlauf folgen:
C:\Program Files (x86)\Staden Package\share\staden\etc\vectors
und entsprechenden Vektor auswählen (pGEM T-Vektor/pUC18)
 - [x] Interactive Clipping *erlaubt manuelle Korrektur der Sequenzbewertung (Qualität/Vektor) im Trev*
 - [x] Gap4 shotgun assembly *automatische Erstellung von Contigs aus gut überlappenden Klonen*
 - Gap4 database name *Datenbank benennen*
 - Create new DB (z.B. neuer Klon) oder Append to existing DB (z.B. zus. Contig)
- Wichtig! Datenbankname so kurz und einfach wie möglich (ohne Umlaute, Zeichen)**

Staden Package



- File > Save All Parameters (in all Modules) to: Konfigurations-Datei (z.B. config.pg4)

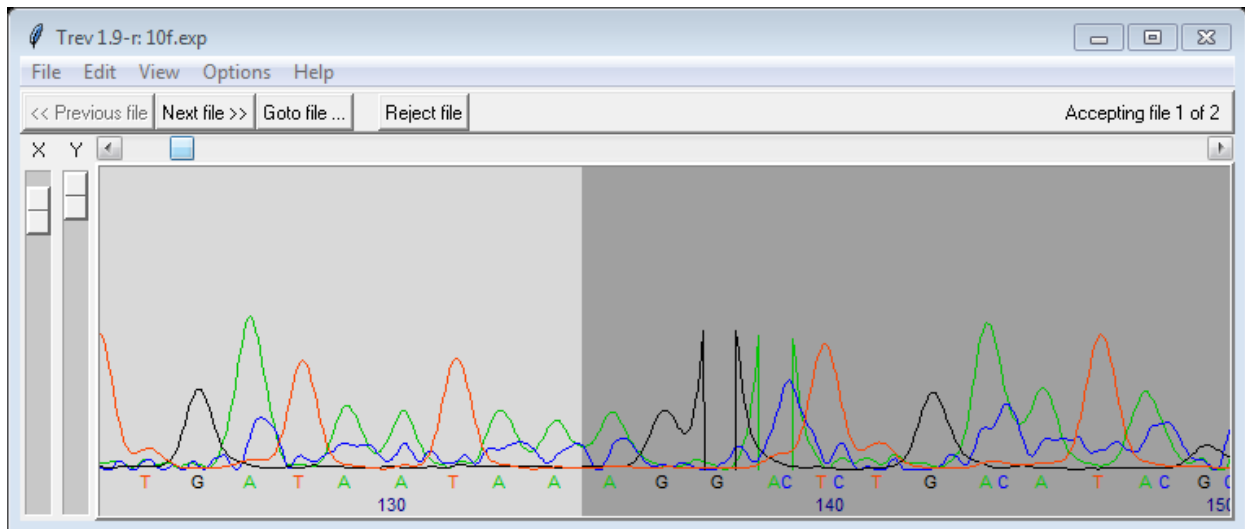
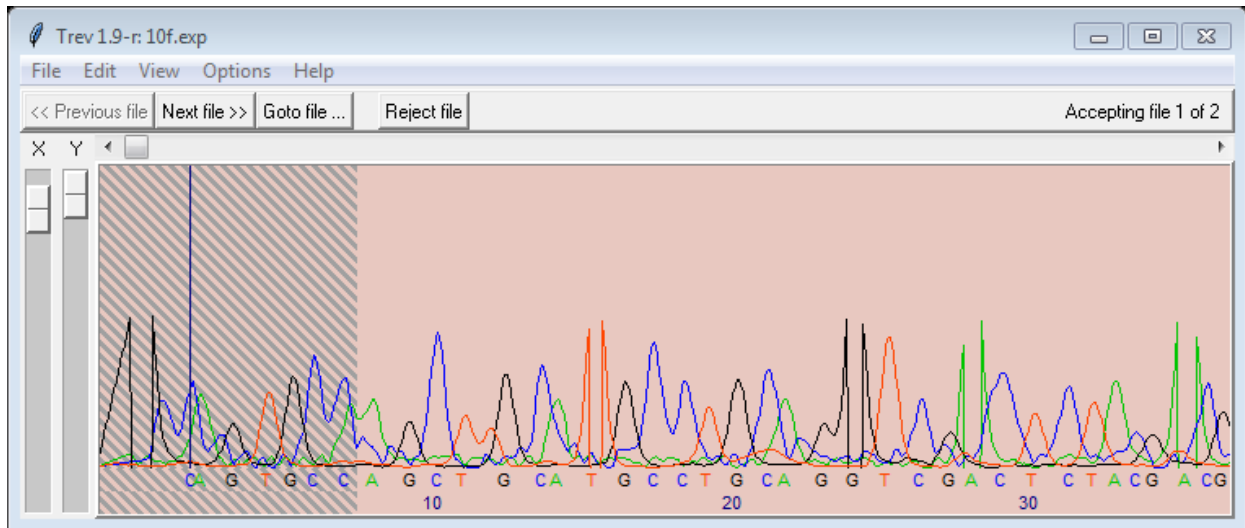
! Hier werden die konfigurierten Parameter gespeichert!

- Konfigurations-Datei laden über File > Load New Config File
- *Wichtig: Immer bei Gap4 shotgun assembly Datenbankname eintragen und wählen ob vorhanden oder neu auszuwählen (sonst Fehler oder Überschreiben der Datei)*
- Auswertung starten mittels: Run (unten links)

Interactive Clipping im Trev

- Manuelle Korrektur von annotiertem Vektor und Sequenzqualität aller Sequenzen
 - Roter Hintergrund = Vektor → wird später abgeschnitten
 - Hellgrau = gute Sequenz
 - Dunkelgrau = schlechte Sequenz → wird später ausgeblendet

Staden Package



- Für Editierung vorher auswählen: Edit > Left/Right Vector bzw. Left/Right Quality
→ Änderung im Sequenzfenster per Klick
- Rückgängig: Edit > Undo Clipping (NICHT CTRL+Z)
- Next/Previous file → Änderungen speichern
- Reject file → qualitativ schlechte Sequenzen komplett verwerfen
- (ggf. Editierung der Sequenz: View > Display Edits > Edit > Sequence
→ DEL löscht Buchstabe davor, Tippen fügt Buchstabe mittig davor ein)

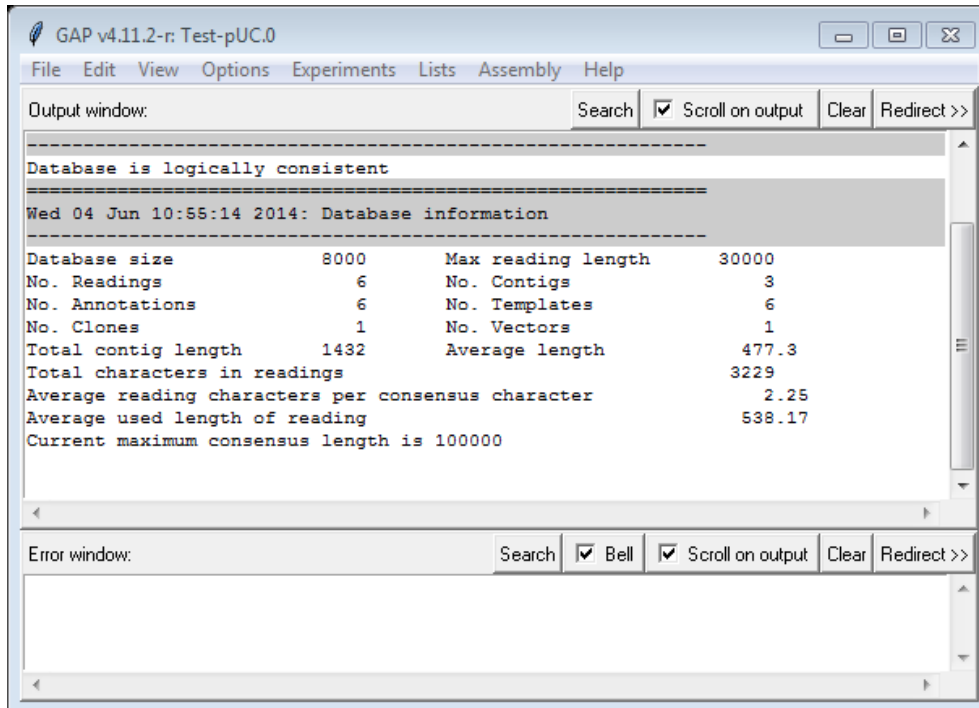
Abschluss der Vorarbeiten

- Trev-Fenster schließen → Bearbeitung in Pregap wird abgeschlossen
- Überprüfung der Prozessierung in Textual Output-Tab möglich
- *.0.aux-Datei im Sequenzordner wird erstellt
- Pregap schließen (sonst lässt sich erstellte *.aux-Datei nicht öffnen)

Staden Package

Assemblierung in Gap4

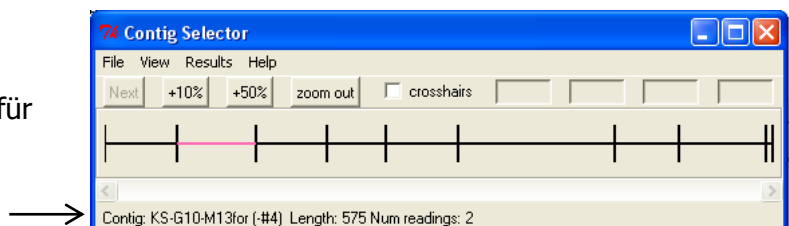
- Öffnen der *.aux-Datei → im Ordner entsteht *.o.BUSY-Datei
- Falls Datei sich nicht öffnen lässt, ggf. vorhandene *.o.BUSY-Datei im Ordner löschen



- Einmalig: Einstellungen ändern für zusätzliche Bearbeitungsoptionen:
Options > Configure Menus > User level: Expert > OK Permanent
- *Tip: Gap4 merkt sich letzten Ordner nicht → Verknüpfung zu Daten-Ordner im geöffneten Fenster erstellen*

Graphische Repräsentation der Klone im Contig Selector

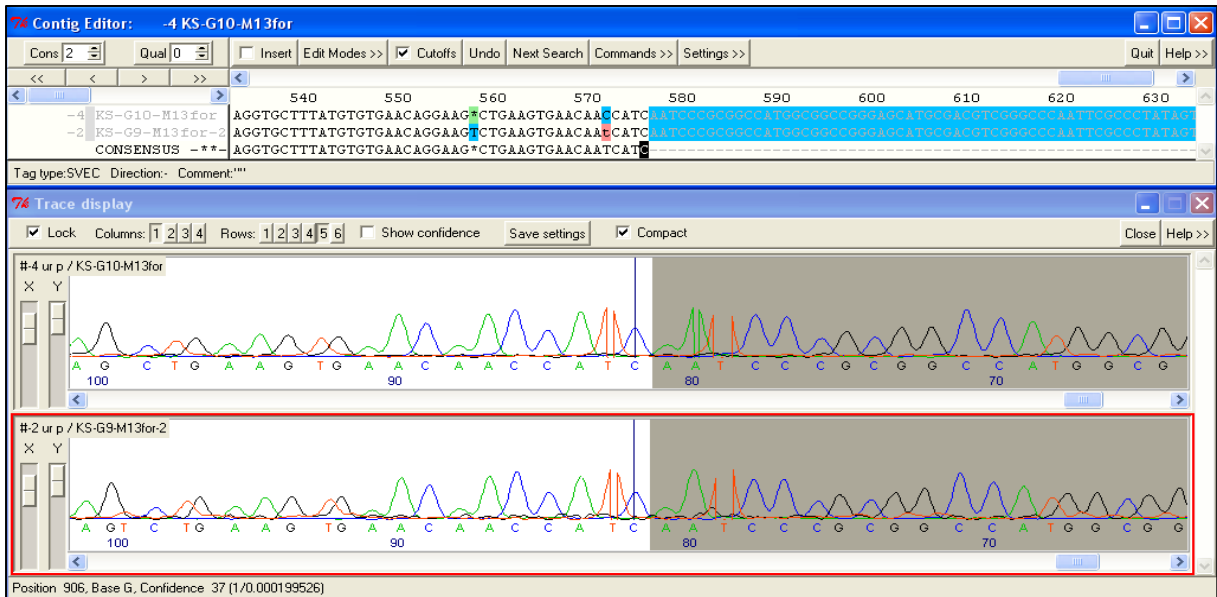
- Öffnen unter Menüpunkt **View**
- Maus über einzelne Contigs halten für Informationen (in Statusleiste)
 - Orientierung (+/-#...)
 - Länge (Length)
 - Anzahl der kombinierten Einzelsequenzen (Num reading)



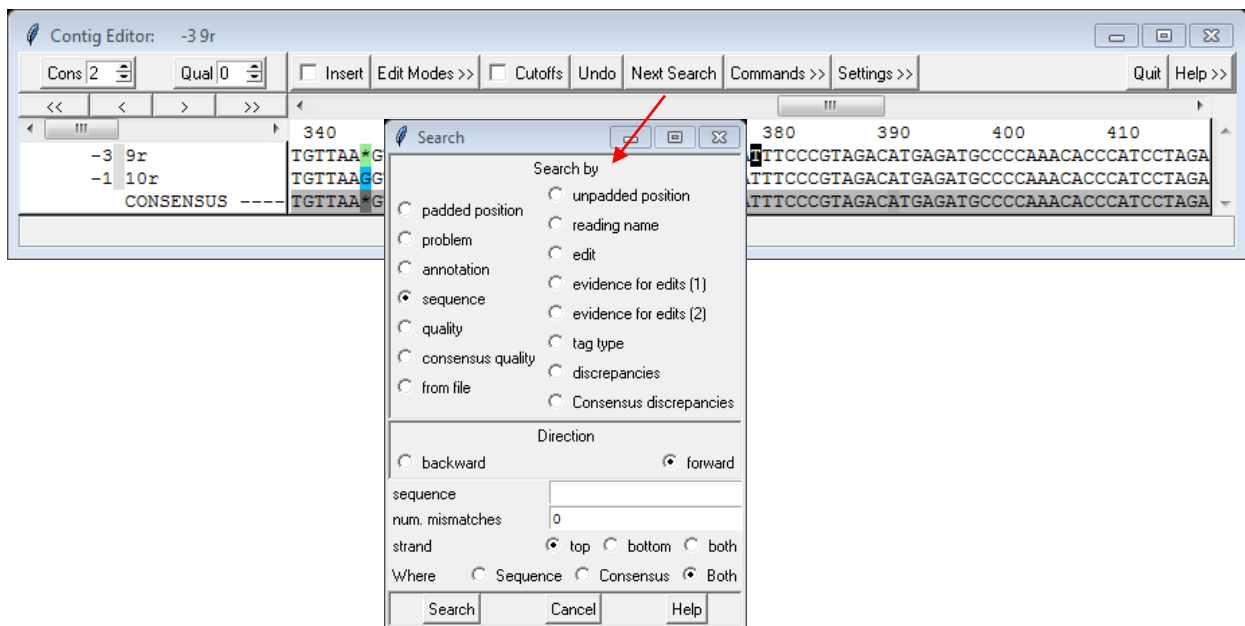
Sequenzen bearbeiten

- Rechtsklick auf Contig im Contig Selector > Edit Contig

Staden Package

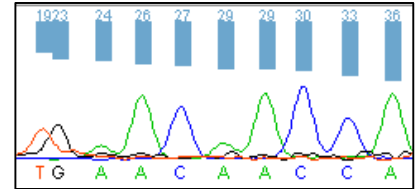


- Empfohlene Einstellungen: Settings >
 - Highlight Disagreements by background colour, case sensitive
 - blau = mismatch (Sequenzierfehler/SNPs); grün = * (potentielle Deletion)
 - Show edits → rot = editiert → dunkel = gelöscht; hell = eingefügt
 - Show Consensus Quality → hell = gut, dunkel = schlecht
 - Store undo
- Edit Modes > Mode set 2 (erlaubt erweiterte Bearbeitung, z.B. Löschen/Korrektur)
- Einmalig: Settings → Save settings (Settings & Edit mode)
- Anzeige der Rohdaten durch Doppelklick auf Einzelsequenz bzw. auf Consensus
 - zeigt alle Sequenz-Rohdaten in entspr. Reihenfolge
- Konflikte zwischen einzelnen Sequenzen oder Sequenzmotive (Primer, Schnittstelle, etc.) können gesucht werden unter Menüpunkt: Next search



Staden Package

- Trace Display-Fenster: Anzeige der Basenqualität: Show confidence = blaue Balken – zuverlässig = hoher Wert, gleichmäßiger Abstand
- Anzeige schlecht-qualitativer oder abgeschnittener Bereich: Cutoffs (blaue Sequenz)
- +/- vor Klonname = Orientierung
- Rückgängig: Undo-Feld
- Consensus: "*" taucht nicht in Fasta-File auf, "-" erscheint!
- Standard: Überschreiben der Daten/Löschen mit Backspace/Del → Einfügen: Insert wählen



Sequenz revers-komplementieren

- Hauptfenster: View > Contig selector → Rechtsklick auf Contig > Complement Contig

Sequenzen zu Contig zusammenführen > Funktion notwendig für Primerwalking

- Hauptfenster: Edit > Join Contig → Sequenzen im Contig selector wählen
- Originaldaten über Doppelklick anzeigbar
- Mismatches erscheinen in grauer Mittelleiste unter Differences als "!"
-

Automatisch

- Automatisches Alignment: Alignbefehl → automatische Lücken "*" erscheint noch als "!"
- Rückgängig: Undo im oberen UND unteren Fenster
- Änderungen annehmen über Join/Quit → speichert Sequenzen in gemeinsamem Contig mit Lücken, Nachfrage bei vielen Mismatches
- Änderungen speichern, aber ohne Join – vor Verlassen: Commands > Save Contig
-

Manuell

- Sequenzen gegeneinander verschiebbar nach Lösen von Lock
- Editierung von Unterschieden: Löschen nur wenn klar begründbar – z.B.
 - bei häufig eingefügter Base – z.B. A nach C oder C nach G
 - wenn verglichen mit qualitativ besserer Sequenz desselben Klon
 - wenn Fehlen/Überschuss klar aus Basenabständen hervorgeht
 - bei Überlagerung durch unspezifischen Peak
 - bei Fehlinterpretation des Hintergrundrauschens
 - G oft farbstoffbedingt schwach
- Keine Lücken einfügen – stattdessen Sequenzen verschieben mittels Unlock, ◀/▶, Lock
- Lücken erst nachträglich automatisch mittels Align-Befehl einfügen lassen!
- Bei Contigs mit mehreren Klonen, die Gesamtsequenz nicht überspannen: erst Lücke(*/-) einfügen in allen Einzelsequenzen und dann * im Consensus löschen (einzelnes Löschen kann Verschiebung bewirken)
- Änderungen speichern, aber ohne Join – vor Verlassen: Commands > Save Contig

Contigs bearbeiten

- Bei Contigs mit mehreren Klonen, die Gesamtsequenz nicht überspannen: Lücke einfügen in allen und * im Consensus löschen (einzelnes Löschen kann Verschiebung bewirken)
- Bei Verschiebung: falsche Sequenz → Rechtsklick → Break Contig → erneut über Join Contig vereinen mit anderen Sequenzen (Lücken bleiben erhalten)

Staden Package

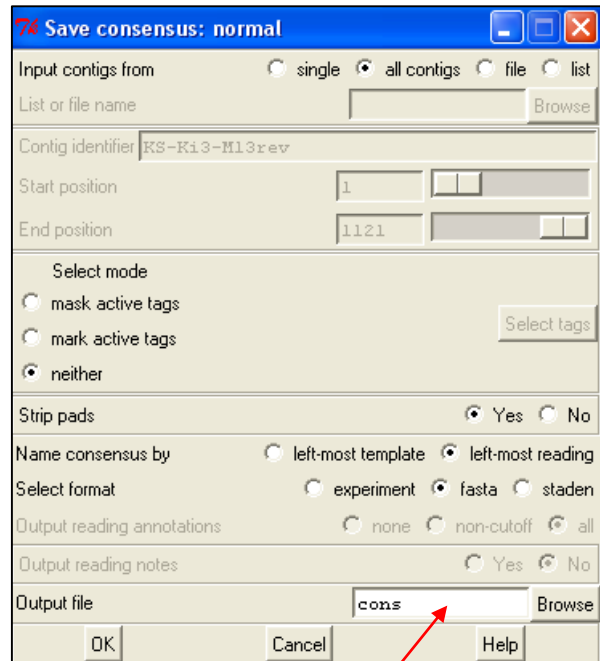
- Sequenzen nachträglich ausblenden: Cutoffs anzeigen lassen, letzte abgeschnittene Base anklicken, CTRL+◀/▶

Contigs annotieren

- Mit Maus markieren (z.B. im Consensus) → Rechtsklick → Create Tag → Type auswählen und Kommentar und Orientierung (Pfeil)
- Anwendungsmöglichkeiten: Primer, SINE, poly(A), Motive,...
- Löschen: Rechtsklick → Delete Tag

Consensus speichern

- File → Save consensus → normal
- Für einzelne Contigs oder alle im Fasta-Format
- Outputfile: *.fas eingeben
→ im Sequenz-Ordner gespeichert



cons.fas

View	Options	Experiments	List

Contig selector			
Contig list			

Results manager			

Find internal joins			
Find read pairs			
Find repeats			
Check assembly			
Sequence search			

Database information			
Template Display			
Show relationships			
Restriction enzyme map			
Stop codon map			
Quality plot			
Confidence values graph			
2nd-highest Confidence graph			
Diploid graph			
Reading coverage histogram			
Readpair coverage histogram			
Strand coverage			
List Consensus Confidence			
List Base Confidence			
Contig Navigation			
SNP candidates			

Weitere Möglichkeiten

- *Es gibt im Gap zahllose weitere Bearbeitungs- und Anzeigemöglichkeiten (z.B. Aminosäureübersetzung, Restriktionsschnittstellen-Suche,...)*
- *Also ausprobieren und freuen*

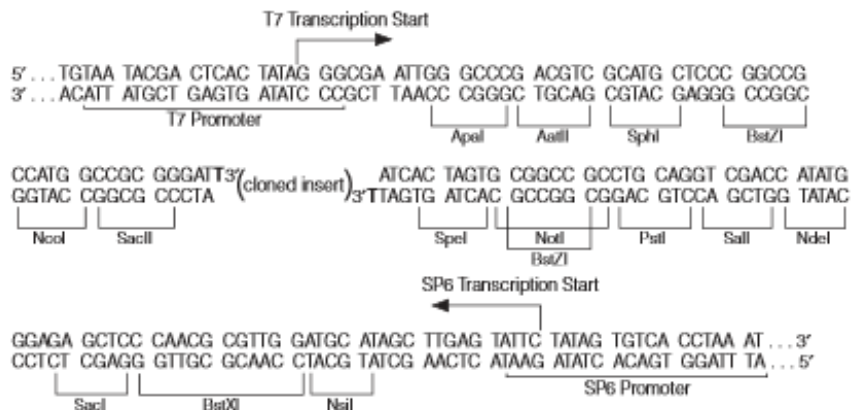
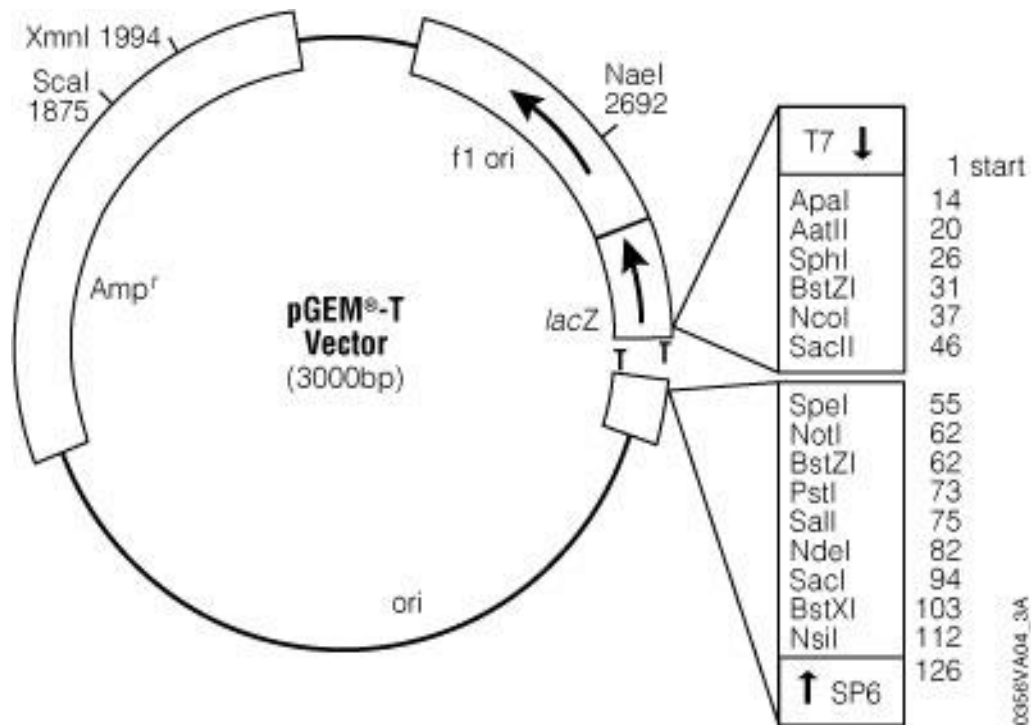
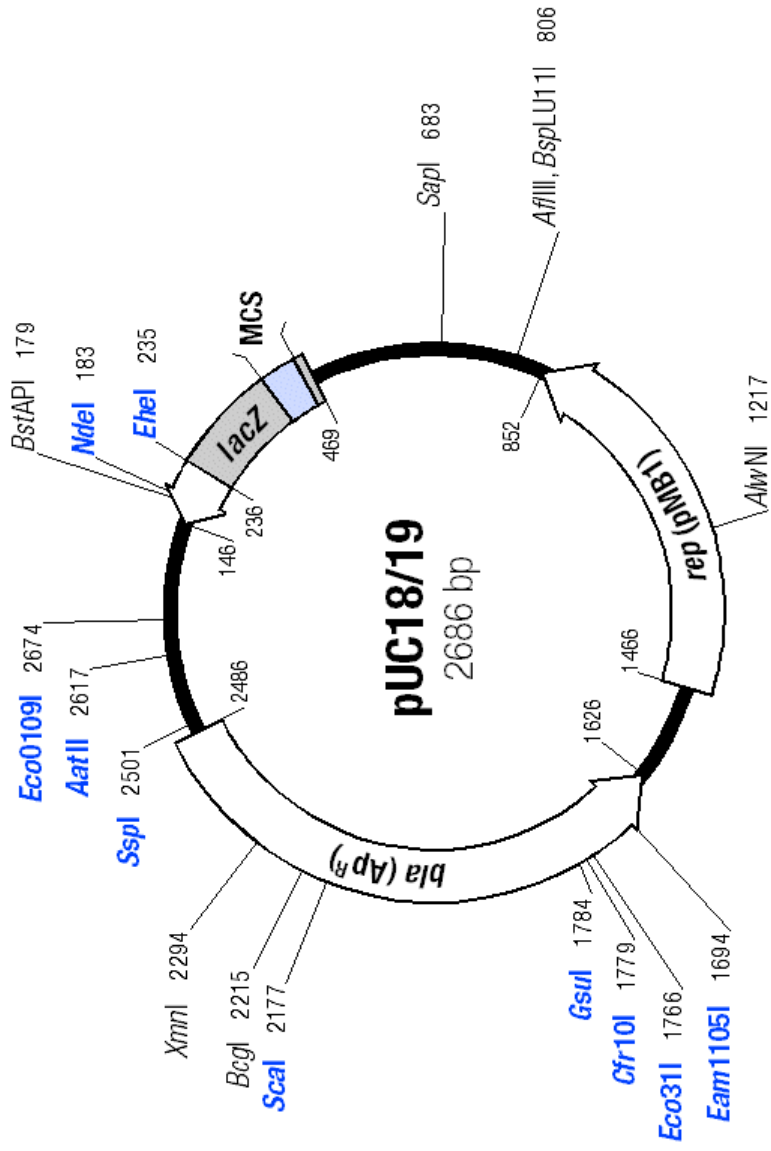


Figure 1. The promoter and multiple cloning sequence of the pGEM[®]-T Vector. The top strand corresponds to the RNA synthesized by T7 RNA polymerase. The bottom strand corresponds to the RNA synthesized by SP6 RNA polymerase.



pUC18

M13/pUC sequencing
primer (-20), 17-mer

5' GTAA AAC GAC GGCCAG TGC CAA GCT TGC ATG CCT GCA GGT CGA CTC TAG AGG ATC CCC GGG TAC CGA GCT CGA ATT CGT AAT CAT GGT CAT AGC TGT TTC CTG 3'
 3' CATT TTG CTG CCGGTC ACGGTT CGA ACG TAC GGA CGT CCA GCT GAG ATC TCC TAG GGG CCC ATG GCT CGA GCT TAA GCATTA GTA CCA GTA TCG ACA AAGGAC 5'

LacZ ← Val Val Ala Leu Ser Ala His Arg Cys Thr Ser Glu Leu Pro Asp Gly Pro Val Ser Ser Asn Thr Ile Met Thr Met

← M13/pUC reverse sequencing
primer (-26), 17-mer