

DNA-Klonierung und *in silico* Analysen

Diese Übung besteht aus zwei Teilen.

Im **Teil A** wird ein Gen über eine klassische und einfache Methode in *Escherichia coli* umklontiert. Dabei wird das Gen für die Dihydroflavonol-4-reduktase (DFR) aus einem high-copy Plasmid (pJET1.2), in die Erstklonierungen üblicherweise erfolgen, isoliert und in ein low-copy Plasmid (pGEX-6P-1) zur späteren heterologen Expression ligiert. Das umzuklonierende Gen liegt in 2 Varianten vor, die sich auf DNA und Aminosäureebene unterscheiden. Durch Sequenzierung bestimmen Sie, um welche der Variante es sich bei ihrer Probe handelt. Das Beispiel umfasst alle klassischen Klonierungsschritte: präparativer Restriktionsverdau, Agarosegelelektrophorese, Geextraktion, Ligation, Transformation von *E. coli*, Selektion, Plasmidpräparation, analytischer Restriktionsverdau und DNA-Sequenzierung.

Teil B beschäftigt sich mit der *in silico* Konstruktion eines Expressionsvektors für ein bestimmtes Gen. Beginnend mit der *in silico* Analyse eines genomischen DNA Fragmentes, werden Sie über die Identifizierung von Regionen mit Aminosäuresequenzähnlichkeiten zwischen ihrer Sequenz und öffentlich zugänglichen Sequenzen, die kodierende Regionen in ihrer DNA Sequenz lokalisieren. Sie werden dann ihre DNA Sequenz in eine cDNA Sequenz umwandeln und Oligonukleotide zur Amplifikation dieser cDNA designen.

Hinweise:

* wird von dem/r Betreuer/in ausgegeben

(A) befindet sich im Abzug für die Agarosegelelektrophorese

(B) befindet sich am Betreuertisch

(E) befindet sich im allgemeinen Entnahmebereich auf RT

(G) befindet sich in der Gefrierlade

(K) befindet sich im Kühlschrank

alle weiteren Materialien/Lösungen sind selbst herzustellen oder am Hörer/innenplatz vorhanden

Die Plasmidkarten sind am Ende dieser Vorschrift abgebildet.

Die Sequenz der 2 Varianten der dfr Gene sind im TISS verfügbar.

Gerahmte Teile: es ist unter sterilen Bedingungen zu arbeiten!

Doppelt gerahmte Teile: es ist im Abzug zu arbeiten!

Abkürzungen:

amp, Ampicillin

BLAST, Basic Local Alignment Search Tool

MCS, Multiple cloning site

min, Minuten

RG, Reaktionsgefäß

RT, Raumtemperatur

sb Wasser, steriles (bi-)destilliertes Wasser (selbst herzustellen)

TEIL A

1. Herstellung der Medien

Lösungen: Ampicillin (1000-fach) (K)

Materialien: Erlenmeyerkolben mit Deckel oder Schottflasche
Epruvetten mit Metallkappen (E)
Wägeschälchen (E)
Pepton, Hefeextrakt, NaCl (E)

Geräte: Wasserbad

Durchführung:

LB-Lennox Medium: 1 % (w/v) Pepton
0,5 % (w/v) NaCl
0,5 % (w/v) Hefeextrakt

- Substanzen in etwas dest. H₂O lösen
- mit dest. H₂O auf gewünschtes Volumen auffüllen
- je 5 mL in Epruvetten aliquotieren (Dispenser bei Betreuer/innen erhältlich)
- Epruvetten mit Metallkappen verschließen
- in Korb oder Becherglas stellen
- mit Alufolie (E) abdecken
- autoklavieren

- auf RT oder im Kühlschrank lagern

LB/amp-Platten:

- Herstellung von LB-Medium in Schottflasche oder Erlenmeyerkolben wie oben
- 1,5 % (w/v) Agar-Agar zugeben
- schwenken
- Schottflasche mit Schraubkappe verschließen (zuletzt eine Vierteldrehung retour)
- autoklavieren
- auf ca. 50 °C abkühlen lassen/temperieren (Wasserbad)

- Ampicillin zugeben
- Medium in Petrischalen gießen

- Medium fest werden lassen
- Platten können einige Tage auf RT, längerfristig im Kühlschrank gelagert werden

2. Restriktionsverdau der Plasmide zur Klonierung

Lösungen: Restriktionsenzyme *Bam*HI-HF und *Eco*RI-HF (G)
 Cut Smart Puffer für Restriktionsenzyme (K)
 Plasmid pJET1.2-DFR (G)
 Plasmid pGEX-6P-1 (G)
 sb Wasser

Geräte: Inkubator oder Thermoblock

Durchführung:

- Die 2 Varianten der pJET1.2-DFR Plasmids sowie der Expressionsvektor pGEX-6P-1 befinden sich in RG in einer Box. Verwenden sie eines der nummerierten pJET1.2-DFR Plasmide als ihr Ausgangsplasmid für die Klonierung und notieren sie die Nummer.
- Reagenzien falls nötig auftauen
- alle Reagenzien schütteln, kurz abzentrifugieren und auf Eis stellen
- Reaktionsansätze je Plasmid wie folgt pipettieren:

Plasmid-DNA	2 µg
10x CutSmart Puffer	5 µL
<i>Bam</i> HI-HF	1 µL
<i>Eco</i> RI-HF	1 µL
sb Wasser	auf 50 µL

- mit der Pipette gut mischen und kurz abzentrifugieren
- bei 37 °C über Nacht inkubieren

3. Agarosegelelektrophorese

Lösungen: 6x Purple Loading Dye (K)
 DNA-Marker: Gene Ruler 1kb DNA Ladder (K)
 DNA-Färbemittel SYBR Safe (K)
 1x TAE Laufpuffer (A)

Materialien: 250-mL-Erlenmeyerkolben mit Deckel (A)
 Agarose (E)
 Wägeschälchen für Agarose (E)
 Topfhandschuh (A)
 Agarosegelelektrophoreseschlitten (A)
 Agarosegelelektrophoresekamm (A)

Kleine Wasserwaage (A)
Pipette (A)
Meßzylinder (A)

Geräte: Mikrowelle (A)
Agarosegelelektrophoreselaufkammer mit Deckel (A)
Power Supply (A)
Gel-Imager

Durchführung:

- Entscheiden Sie aufgrund der Tabelle, welche Agarosekonzentration das Gel haben soll

Agarosekonzentration (%)	Größe linearer DNA-Fragmente, für die eine gute Auflösung erreicht wird (kb)
0,5	30 to 1
0,7	12 to 0,8
1,0	10 to 0,5
1,2	7 to 0,4
1,5	3 to 0,2
2,0	<0,2

- die entsprechende Menge Agarose für ein 50 oder 100 mL Gel in das dafür vorgesehene Wägeschälchen einwiegen

- Agarose in den dafür vorgesehenen 250-mL-Kolben geben
- 50 (kleines Gel) oder 100 mL (großes Gel) TAE-Puffer zugeben, schwenken
- Deckel auf Kolben legen
- für ca. 2 min in der Mikrowelle (800 Watt) erhitzen
- vorsichtig schwenken (Achtung: Hitzeverzug!)
- falls Agarose nicht vollständig gelöst sein sollte, noch einmal kurz aufkochen
- die Agarose etwas abkühlen lassen
- Gelschlitten in Gießkammer einspannen und mit Wasserwaage eben ausrichten
- Kamm für breite Gelspuren in den Schlitten stecken
- Agarose mit 5 μ L SYBR Safe versetzen, schwenken, in den Schlitten gießen
- eventuell entstandene Luftblasen entfernen (gelben Pipettenspitzen sind hilfreich)
- Gel ca. 20 min aushärten lassen
- gesamte Plasmidverdaue mit entsprechendem Volumen 6x Purple Loading Dye mischen
- überprüfen, ob genug TAE Laufpuffer in der Laufkammer ist
- Gel samt Schlitten in die Gelkammer legen
- vorsichtig Kamm entfernen
- 5 μ l DNA-Marker in die erste Geltasche pipettieren
- die Plasmidverdaue in die weiteren Geltaschen laden
- Deckel auf Gelapparatur setzen und richtig (!) mit dem Power Supply verbinden
- Gel für ca. 50 min bei 90 V laufen lassen

- Gel mit dem Schlitten auf den Gel-Imager legen
- falls eine ausreichende Auftrennung der DNA Fragmente erfolgt ist, das Gel ohne Schlitten auf den Gel-Imager legen, fotografieren und das Bild speichern
- falls keine ausreichende Auftrennung erfolgt ist, das Gel weiterlaufen lassen
- Gießkammer, Schlitten und Kamm mit dest. H₂O reinigen und retournieren

4. Gelextraktion der DNA Fragmente

Lösungen: GeneJET Gel Extraktions Kit (B)

Materialien: Skalpell (A)

Geräte: Gel-Imager (neben A)
Gel Cutting Guard (A)
Magnetischer Schalter (A)
Schutzbrille orange (A)
Thermoblock

Durchführung: Wenn nicht anders angegeben, bei 13000 x g zentrifugieren. Überlegen Sie sich zuerst, welche Banden Sie ausschneiden müssen.

- Guard zum Schutz des Imagers auflegen
- Gel ohne Schlitten auf den Gel-Imager legen
- Schutzbrille aufsetzen
- magnetischen Schalter anbringen damit das Gerät auch ohne Kameraabdeckung gestartet werden kann
- Transillumination starten
- zügig mit dem Skalpell die richtigen Banden ausschneiden (Gen bzw. Zielplasmid)
- Gelstücke getrennt in vorher abgewogene 2-mL-RG geben
- Imager ausschalten, mit dest. H₂O reinigen und in ursprüngliche Konfiguration zurückführen
- alle verwendeten Materialien (Schlitten, Skalpell etc.) mit dest. H₂O reinigen und retournieren
- 2-mL-RG mit Gelstücken abwägen
- das gleiche Volumen (in μL) an Binding Buffer des Gelstückgewichtes (in mg) zugeben
- RG bei 50 - 60 °C inkubieren bis sich das Gel vollständig gelöst hat (max. 10 min)
- Bis zu 800 μL der Lösung auf ein Säulchen (mit RG) auftragen und 1 min zentrifugieren
- Durchfluss im RG verwerfen
- 700 μL Wash Buffer auf die Säule geben, 1 min zentrifugieren
- Durchfluss im RG verwerfen
- die leere Säule mit RG noch einmal 1 min zentrifugieren
- Säule in ein neues 1,5-mL-RG setzen, Deckel vorher entfernen
- vorsichtig 20 μL Elution Buffer auf die Säule pipettieren, kurz einwirken lassen und 1 min zentrifugieren
- Eluat aus deckellosem RG in neues 1,5-mL-RG transferieren
- DNA-Konzentration mit dem NanoDrop messen (kontaktieren Sie dazu eine/n Betreuer/in)
- direkt weiterverwenden oder auf -20 °C aufbewahren

5. Ligation des *drf* Gens in pGEX-6P-1

Lösungen: 10x T4 Reaction Buffer (G)
T4 DNA Ligase (G)
sb Wasser

Durchführung:

- Reagenzien auftauen, schütteln, kurz abzentrifugieren, auf Eis stellen
- Menge an Plasmid und Insert berechnen (siehe Formel unten)
- in ein 1,5-mL-RG folgenden 20 μL Ansatz pipettieren:

geschnittenes Zielplasmid	x μL
geschnittenes Insert (Gen)	y μL
10x T4 Reaction Buffer	2 μL
mit sb Wasser auf 19 μL Gesamtvolumen auffüllen	
- T4 DNA Ligase 1 μL
- mit der Pipette gut mischen und kurz abzentrifugieren
- 1 h bei RT inkubieren

Formel zur Berechnung der benötigten Menge an Insert:

Das molare Verhältnis Vektor zu Insert soll für diese sticky end Ligation ca 1:3 sein. Für die Ligation werden zwischen 20-50 ng Plasmid eingesetzt. Um die benötigte ng Menge an Insert für die Ligationsreaktion auszurechnen, kann folgende Formel verwendet werden.

$[\text{ng des Zielplasmids}] \times [\text{kb Insert}] / [\text{kb Vektor}] \times (\text{Verhältnis Insert:Vector}) = \text{ng benötigtes Insert}$

Größe des *dfv* Gens: 1,026 kb

Größe des pGEX-6P-1: 4,984 kb

Daraus kann nun das benötigte Volumen für Vektor bzw Insert berechnet werden. Bitte dabei beachten, dass das Gesamtvolumen des Ligationsansatzes 20 μl nicht überschreiten darf.

6. Transformation des Ligationsansatzes in *E. coli* TOP10

Lösungen: kompetente *E. coli* TOP10 Zellen*

Materialien: Parafilm (E)

Geräte: Thermoblock
Inkubator

Durchführung:

- ein Aliquot kompetenter *E. coli* TOP10 Zellen aus dem -80°C Ultratiefkühler wird auf Eis vorsichtig aufgetaut (ca. 10 min)
- 10 μL des Ligationsansatzes dazu pipettieren und vorsichtig mischen (möglichst immer auf Eis lassen und nicht erwärmen)
- 30 min auf Eis inkubieren
- Hitzeschock für exakt 2 min bei 42 °C (Thermoblock)
- zurück auf Eis stellen
- 300 μL LB-Medium dazu mischen
- Zellen bei 37 °C 30 min regenerieren (Inkubator)
- Bakteriensuspension auf 3 LB/amp-Platten mit Drigalskyspatel ausstreichen (z.B. ein Zehntel, ein Fünftel und den Rest der Suspension)
- Platten bei 37 °C über Nacht inkubieren
- Am Morgen Platten aus dem Inkubator nehmen, auf Kolonien kontrollieren, dann mit Parafilm verschließen und bei 4 °C aufbewahren

7. Anzucht von *E. coli* Einzelkolonien

Lösungen: Ampicillin (1000-fach) (K)

Materialien: Zahnstocher (E)

Geräte: Schüttelinkubator

Durchführung:

- drei 5-mL-Aliquote LB-Medium in Epruvetten mit Ampicillin versetzen
- jede Epruvette mit einer *E. coli* Einzelkolonie beimpfen, dazu können sterile gelbe Pipettenspitzen oder sterile Zahnstocher verwendet werden
- bei 37 °C und 200 rpm über Nacht im Schüttelinkubator inkubieren. Kulturen erst am späteren Nachmittag inkubieren.

8. Isolierung der Plasmid-DNA

Lösungen: GeneJET Plasmid Miniprep Kit (B)

Durchführung: Wenn nicht anders beschrieben, bei 13000 x g zentrifugieren.

- 2 mL jeder der drei *E. coli* Kulturen in ein 2-mL-RG überführen
- 2 min bei 6800 x g zentrifugieren
- Überstand verwerfen
- weitere 2 mL Kultur hinzufügen und 2 min bei 6800 x g zentrifugieren
- Falls noch Kulturflüssigkeit vorhanden, Ernte nochmals wiederholen
- nach letztem Zentrifugieren den Überstand komplett abnehmen
- Zellpellet in 250 μ L Resuspension Solution durch Schütteln resuspendieren
- 250 μ L Lysis Solution dazu pipettieren
- das RG 4-6 mal schwenken
- 350 μ L Neutralization Solution dazu pipettieren
- das RG 4-6 mal schwenken
- 5 min zentrifugieren
- Überstand vorsichtig in Säulchen pipettieren
- 1 min zentrifugieren
- 500 μ L Wash Solution in das Säulchen pipettieren
- 1 min zentrifugieren
- noch einmal 500 μ L Wash Solution in das Säulchen pipettieren
- 1 min zentrifugieren
- leeres Säulchen noch einmal 1 min zentrifugieren
- Säule in ein neues 1,5-mL-RG geben, Deckel des RGs entfernen
- 50 μ L Elution Buffer vorsichtig auf den Säulchenboden pipettieren
- 1 min zentrifugieren
- Eluat in neues RG (mit Deckel) transferieren
- benennen Sie Ihre Plasmide; üblich wäre:
 - ein p für Plasmid
 - ein Hinweis auf das Plasmid Backbone
 - ein Hinweis auf das Insert
 - und/oder Ihre Initialen
 - eine laufende Nummer
 - zB pGEX-DFR-1
- DNA-Konzentrationen mit dem NanoDrop messen (kontaktieren Sie dazu eine/n Betreuer/in)
- weiterverwenden oder auf -20 °C aufbewahren

9. Kontrollverdau (Restriktionsanalyse)

Lösungen: Restriktionsenzyme *Pst*I-HF, *Bam*HI-HF und *Eco*RI-HF (G)
CutSmart Puffer
6x Purple Loading Dye (K)
DNA-Marker: Gene Ruler 1kb DNA Ladder (K)
DNA-Färbemittel SYBR Safe (K)
1x TAE Laufpuffer (A)

Materialien: 250-mL-Erlenmeyerkolben mit Deckel (A)
Agarose (E)
Wägeschälchen für Agarose (E)
Topfhandschuh (A)
Agarosegelelektrophoreseschlitten (A)
Agarosegelelektrophoresekamm (A)
Kleine Wasserwaage (A)
Pipette (A)

Geräte: Mikrowelle (A)
Agarosegelelektrophoreselaufkammer mit Deckel (A)
Power Supply (A)

Gel-Imager

Führen Sie 2 Kontrollverdau durch: einer soll das Plasmid linearisieren (*Pst*I-HF) und der zweite daraus das Insert wieder ausschneiden (*Bam*HI-HF und *Eco*RI-HF). Stellen Sie für beide Kontrollverdau je einen Mastermix her.

Anzahl der Proben je Mastermix:

___ Plasmid Proben	___
Aufschlag	1
Gesamtanzahl	___

Mastermix pro Kontrollverdau (sb Wasser dient zum Auffüllen des Volumens):

Reagens	Konz.	Endkonz./-menge	1 Probe [μ L]	___ Proben [μ L]
CutSmart Puffer	10x	1x		
<i>Pst</i> I-HF	-	-	0,5	
Plasmid DNA-Probe		ca. 500 ng		
sb Wasser	-	-		
Gesamt	-	-	20	

Reagens	Konz.	Endkonz./-menge	1 Probe [μ L]	___ Proben [μ L]
CutSmart Buffer	10x	1x		
<i>Bam</i> HI-HF	-	-	0,5	
<i>Eco</i> RI-HF	-	-	0,5	
Plasmid DNA-Probe		ca. 500 ng		
sb Wasser	-	-		
Gesamt	-	-	20	

Durchführung:

- pro Mastermix ein 1,5-mL-RG auf Eis bereitstellen
- Reagenzien und Plasmid-DNA Proben auftauen, schütteln, kurz abzentrifugieren, auf Eis stellen
- Mastermixe herstellen: die oben berechneten Mengen der Reagenzien (außer der Proben) in das jeweilige RG pipettieren, gut schütteln und kurz abzentrifugieren
- die entsprechenden Volumina der jeweiligen Mastermixe in 1,5-mL-RG vorlegen
- die entsprechenden Volumina Plasmid-DNA Proben zu beiden vorgelegten Mastermixen mischen
- 60 min bei 37 °C inkubieren
- Danach alle Ansätze mit entsprechendem Volumen 6x-Ladepuffer versetzen
- je 15 μ L auf einem 0,8%igen Agarosegel analysieren (siehe dazu Punkt 3. Agarosegelelektrophorese)
- Gellauf: 50 min, 90 V
- Gel am Gel-Imager fotografieren und Bild speichern

10. Sequenzierung

Durchführung:

- ein Plasmid mit korrektem Restriktionsmuster auswählen
- falls erforderlich diese Plasmid-DNA auftauen, schütteln, kurz abzentrifugieren
- 700 – 1200 ng der Plasmid-DNA in ein 1,5-mL-RG pipettieren

- mit sb Wasser auf ein Endvolumen von 12 μ L auffüllen
- das RG bei einem/r Betreuer/in abgeben
- das Sequenzierergebnis wird per e-mail an Sie übermittelt
- Kontrollieren Sie die Sequenz des *dfr* Gens und identifizieren Sie um welche der 2 Varianten es sich handelt. Die Sequenzen der verschiedenen *dfr* Gene werden im TISS zur Verfügung gestellt.

TEIL B

11. Identifizierung der kodierenden Region in einer DNA Sequenz mit BLAST

Die für diesen Teil nötige DNA Sequenz finden Sie im TISS. Im ersten Schritt werden Regionen mit Ähnlichkeiten in der Aminosäuresequenz zwischen ihrer Sequenz und öffentlich zugänglichen Sequenzen identifiziert. Dazu wird ihre DNA Sequenz mit Aminosäuresequenzen der NCBI Datenbank (National Center for Biotechnology Information) verglichen. Für diesen Vergleich verwenden Sie das BLAST (Basic Local Alignment Search Tool) Programm. Unter den verschiedenen BLAST Programmen wählen sie dabei das blastx Programm aus. Blastx übersetzt ihre Sequenz in alle 6 Leserahmen und vergleicht diese zur Proteindatenbank am NCBI.

Materialien: DNA Sequenz
Abbildung Expressionsvektor
Sequenz Expressionsvektor
Gene Runner
Unterlagen Vorbesprechung

Durchführung:

- Gehen Sie auf die NCBI Webseite: <https://www.ncbi.nlm.nih.gov/>
- Suchen Sie das BLAST Programm und klicken Sie es an
- Suchen Sie das blastx Programm und klicken Sie es an
- Kopieren Sie ihre DNA Sequenz im FASTA Format in das offene Fenster: Enter Query Sequence.
Das FASTA Format sieht so aus:
>definition line (=Name ihrer Sequenz)
catgagagatctagtctagctgat (ihre Sequenz selbst).
- Klicken Sie auf den BLAST Button, um ihren Sequenz abzuschicken. Das Resultat erscheint dann nach ein paar Sekunden bis Minuten.

Im Fenster Blast Results sehen Sie dann die Verteilung der Treffer (Hits) in Bezug auf ihre übersetzte Query Sequenz. Ein Farbcode gibt dabei an wie ähnlich die Sequenzen zu ihrer Query Sequenz sind. Diese farbigen Balken entsprechen in der Regel den Exons ihres Gens und sind meist durch dünne schwarze Striche unterbrochen, die den Introns entsprechen. Aus dieser Abbildung lässt sich somit grob bestimmen aus wie vielen Exons bzw Introns ihr Gen besteht. Außerhalb dieses übersetzten Bereichs ihres Gens befinden sich weitere nicht konservierte Regionen, das sind die Promoter- und Terminatorregion. Unter diesen Hits befinden sich die Sequences producing significant alignments, die überspringen Sie und gehen direkt zu Alignments. Überprüfen sie zuerst in welcher Orientierung sich das Gen in ihrer DNA befindet. Ihre Sequenz ist das Query. Die Sequenzen der NCBI Datenbank sind das Sbjct (Subject). Bei der Überprüfung achten Sie darauf, dass die Nummerierung der Aminosäuren der Proteine der Datenbank immer in der richtigen Orientierung vom N- zum C-Terminus angegeben wird. Ihre Query Sequenz orientiert sich aber an der Proteinsequenz des Sbjct. Da ihr Gen in der DNA in 2 Orientierungen (+ oder - Strang) vorkommen kann, ist es für die weitere Analyse einfacher, das Gen in der richtigen Orientierung zu haben. Sind ihre Frames (Leserahmen) positiv, also +1, +2 oder +3, ist ihre DNA Sequenz in der richtigen Orientierung. Sind ihre Leserahmen -1, -2, -3, dann müssen sie ihre Sequenz mit Hilfe eines DNA Programmes umdrehen.

Durchführung:

- Überprüfen Sie die Orientierung ihres Gens an Hand des BLAST Ergebnisses
- Falls sich das Gen in der falschen Orientierung befindet, wiederholen den blastx Vorgang mit dem Gegenstrang. Drehen Sie dafür die DNA Sequenz um (siehe 12).

12. Änderung der Orientierung der DNA

Durchführung:

- Installieren Sie das DNA Programm Gene Runner (ein aktueller Link wird ihnen am Kursanfang zur Verfügung gestellt)
- Öffnen Sie den Gene Runner. Gehen sie zu Go to File → New → Nucleic acid sequence
- Kopieren Sie ihre DNA Sequenz und fügen sie die DNA Sequenz ein
- Gehen Sie zu Edit → Sequence Orientation → Klicken sie auf Reverse and Complement.
- Somit haben Sie die umgekehrte, komplementäre DNA Sequenz. Sie können die (obere) Sequenz im Gene Runner markieren und direkt mittels Copy und Paste in das blastx Fenster einfügen. Modifizieren Sie die Sequenz noch so, dass Sie dem FASTA Format entspricht.
- Wiederholen sie nun den BLAST wie oben angegeben. Speichern bzw drucken sie die ersten Alignments für die folgende Analyse aus.

13. Übersetzung der DNA Sequenz und Darstellung als DNA/Protein Alignment

Es sind nun alle Informationen vorhanden, um ein DNA Protein Sequenz Alignment zu machen. Prinzipiell geht das auch im Gene Runner, nur lässt sich dieses Alignment in den neueren Versionen nicht mehr exportieren und abspeichern (Gehen Sie auf View → Translation und übersetzen Sie Frame 1-3). Als Alternative stehen z.B. Programme wie EMBOSS Sixpack zur Verfügung.

Durchführung:

- Gehen sie am EMBL-EBI auf das Tool EMBOSS Sixpack (https://www.ebi.ac.uk/Tools/st/emboss_sixpack).
- Kopieren Sie ihre DNA Sequenz und fügen Sie die DNA Sequenz ein. Gehen Sie auf Step 2 Select Parameters bei REVERSE auf no (Sie bekommen dann nur 3 Leserahmen)
- Drücken Sie submit
- Sie erhalten nun ihre DNA Sequenz mit den 3 Leserahmen. Klicken Sie auf Download Sequence File. Kopieren Sie die ganze DNA Protein Sequenz und fügen Sie alles in ein Word Dokument ein. Verwenden Sie eine Schriftart bei der jedes Zeichen die gleiche Größe hat wie z.B. Courier. Im Falle des EMBOSS Sixpack Files befindet sich die Aminosäure über dem ersten Nucleotid des Codons (Triplett). Stops in ihrer Sequenz sind durch * gekennzeichnet.
- Suchen Sie nun den Anfang ihres Proteins mit Hilfe ihres blastx Ergebnisses. Schauen Sie sich dafür die erste Zahl der Aminosäuren der sbjt Line der verschiedenen Reading Frames im blastx Resultat an. Die kodierende Region in ihrem Sbjct line beginnt idealerweise mit Aminosäure 1, wenn ihr erstes Exon mit dem Start ATG/Methionin beginnt. Falls aber der Anfang des Proteins nicht konserviert ist, müssen Sie ihr DNA/Protein Sequenz Alignment upstream nach einem ATG/Methionin absuchen. Zur Vereinfachung befindet sich das Methionin im ihrem Fall immer im selben Leserahmen, es könnte aber auch ein Leserahmenwechsel möglich sein. Verwenden Sie das erste Methionin im Leserahmen als ihr Start Methionin.
- Im nächsten Schritt wird nun das Ende des ersten Leserahmens in ihrem DNA/Proteinalignment gesucht. Anhand des blastx Ergebnisses können Sie nun die Sequenz bis zum Ende des Exons/Leserahmens nachverfolgen. Markieren Sie das Ende des ersten Leserahmens im File.
- Auf das Exon folgt logischerweise ein Intron. Der Anfang eines Introns beginnt in den meisten Fällen (90-95%) mit der Sequenz GT. Es kommt aber vor, dass sich beim Übergang von Exon zu Intron mehrere GTs befinden, die in Frage kommen könnten. Sie müssen daher das richtige finden. Dabei müssen Sie jenes GT nehmen, das am besten zu ihrem blastx Resultat korreliert: Der Anfang des Introns beginnt daher meist nach dem Bereich, der für die letzten konservierten Aminosäuren kodiert. Markieren Sie das GT.
- Gehen Sie dann zum Ende des Introns. Das Intron endet nach einem AG, gefolgt vom zweiten Exon. Nehmen Sie wieder ihren blastx Vergleich, suchen das zweite Exon und suchen sie die ersten Aminosäuren des 2ten Exons in ihrem DNA/Protein Alignment. Markieren Sie mögliche passende AGs. Das richtige AG befindet sich meist kurz vor dem Bereich, der für die ersten konservierten Aminosäuren des zweiten Exons kodiert. Nach dem AG kann es zu einem Wechsel des Leserahmens kommen, es muss aber nicht. Vergleichen sie dazu die Leserahmen im blastx Resultat.
- Wenn Sie nun die Exon/Intron Grenzen gefunden haben, überprüfen Sie ob die Grenzen korrekt sind. Das machen Sie durch Translation der entstehenden Sequenz kodierenden Sequenz. Es

kommt häufig vor, dass ein Intron ein Triplet unterbricht. 1-2 Nucleotide befinden sich dann vor dem Intron (vor dem GT), während die 2-1 Nucleotid(e) nach dem AG folgen, bevor das erste vollständige Triplet anschließt. Falls kein vollständiges Triplet entsteht, müssen Sie Anfang und Ende ihres Introns noch einmal überprüfen und nach Alternativen für ihr gewähltes GT oder AG suchen. Achten Sie dabei aber immer auf die konservierten AA, die erhalten werden müssen. Außerdem soll die Anzahl der Aminosäuren in dem Bereich, den der verwandte Proteine entsprechen. Es sollen also keine Teile hinzukommen oder wegfallen.

- Weiters ist darauf zu achten, dass die Spaltung eines Triplets durch ein Intron auch Konsequenzen auf die *in silico* übersetzte Aminosäuresequenz haben kann und eventuell eine Änderung der AA bewirken kann. Überprüfen Sie daher ob das neue Triplet eine Änderung der AA zur Folge hat und korrigieren Sie es, wenn nötig, im File.

Beispiel AGTGNNNNNNNNNAGCC. Das Triplet AGT kodiert für Serin. Aber durch die Entfernung der Intronsequenz/des GTs (unterstrichen) entsteht ein neues Triplet nämlich ACC, welches für Threonin kodiert.

- In weiterer Folge werden nun die anderen Introns/Exons lokalisiert, bis zum Ende der kodierenden Sequenz. Speichern Sie das Dokument am besten in Kopie ab.
- Machen Sie nun das DNA Protein Alignment. Entfernen Sie alle nicht übersetzten Leserahmen und AA bis nur mehr die DNA und die kodierte Aminosäuresequenz übrigbleibt. Achten Sie darauf, dass Sie durch das Löschen nicht die Aminosäuren verschieben.
- Extrahieren Sie nun die Aminosäuresequenz
- Extrahieren Sie nun den kodierenden Teil der cDNA Sequenz beginnend beim ersten ATG bis zum Stop Codon.
- Führen Sie eine Kontrolle der erhaltenen Sequenzen durch
 Kodierender Teil cDNA Sequenz: Es muss ein durchgehender Leserahmen ohne unterbrechende Stops vorhanden sein. Das können Sie im Gene Runner machen, indem sie die Sequenz nehmen, hineinkopieren und übersetzen.
 Extrahierte Aminosäuresequenz: Das extrahierte Protein können sie mit dem übersetzten Protein aus dem Gene Runner vergleichen. Nehmen Sie dazu statt blastx PROTEIN BLAST (blastp). Klicken sie PROTEIN BLAST an und dann auf Align two or more sequences. Fügen Sie die Aminosäuresequenzen in die 2 Fenster ein und klicken auf BLAST. Die Sequenzen sollten dann 100% identisch sein. Zusätzlich können Sie noch die Aminosäuresequenz gegen die NCBI Datenbank testen. Machen sie das auch wieder über blastp (siehe blastx Beschreibung). Mit diesem Vergleich zu den nahe verwandten Proteinen können Sie analysieren ob ihr Protein von der Größe passt und es keine zusätzlichen/fehlenden Aminosäuren aufweist.

14. Design von Oligonukleotiden für die Amplifikation der cDNA und Klonierung in einen Expressionsvektor

Basierend auf der erhaltenen cDNA Sequenz werden nun Oligonukleotide für die *in silico* Amplifikation der cDNA designt, um die cDNA in einen Expressionsvektor über Restriktionsschnittstellen zu insertieren. Die Herstellung des Expressionsvektors erfolgt dabei über eine konventionelle Klonierung basierend auf Restriktionsverdau und Ligation sowie über eine ligationsunabhängige Klonierung.

Durchführung:

A. Konventionelle Klonierung

- Überprüfen Sie an Hand der Abbildung ihres Expressionsvektors (im TISS), welche Restriktionsenzyme in der Multiple Cloning Site (MCS) ihres Vektors schneiden.
- Überprüfen Sie welche dieser Restriktionsenzyme der MCS nicht in der cDNA schneiden. Öffnen Sie dafür den Gene Runner, kopieren ihre cDNA wie vorher beschrieben hinein und gehen Sie zu Analysis → Nucleic Acid → Restriction Sites. Bei Output format klicken Sie nur List non-cutters an. Vergleichen Sie nun welche die Restriktionsenzyme der MCS mit den erhaltenen Non-Cuttern.
- Wählen 2 Restriktionsenzyme, die im Expressionsvektor in der MCS schneiden und nicht in der cDNA schneiden aus. Am besten geeignet sind Restriktionsenzyme, die am 5' Ende und am 3'

Ende der MCS schneiden. Wichtig ist noch die Reihenfolge der Restriktionsenzyme in der MCS für das Primerdesign zu beachten, damit die cDNA richtig orientiert ist.

- Primer werden, wie für DNA üblich, immer in 5' zu 3' Richtung geschrieben. Der sogenannte Forward Primer beginnt beim Start Methionin (ATG) und dann folgen weitere 16-18 Nukleotide des Anfangs der cDNA. Um die cDNA in ein Plasmid zu klonieren, wird noch zusätzlich die Sequenz für eine Restriktionsenzymstichstelle an den Anfang des Primers gehängt: im Beispiel GAATTC für *EcoRI*.

Beispiel Forward Primer: 5'-GAATTC-ATG-gagctagatagtaataatgata-3'
(*EcoRI* Schnittstelle-StartATG-cDNA spezif. Sequenz)

Der Reverse Primer wird analog zusammengestellt. Er beginnt mit dem letzten Teil der kodierenden Sequenz einschließlich Stop Codon gefolgt von der gewählten Restriktionsschnittstelle. Der Reverse Primer würde dann natürlich in die falsche Richtung zeigen. Daher muss die Sequenz umgedreht werden und die komplementäre Sequenz gebildet werden (Reverse and complement).

Beispiel Reverse Primer: 5'-gggtatatatatgtat-TAG-GGATCC-3'.
(cDNA spezifische Sequenz-StopTAG-*BamHI* Schnittstelle)

Reverse and Complement: 5' GGATCC-CTA-atatataacc-3'

Nun hängt es noch von der Art ihres Expressionsvektors ab, ob die Oligonukleotide noch weiter verändert werden müssen. Expressionsvektoren enthalten oft Tags die N oder C-Terminus an das Protein angehängt werden. Daher ist zu beachten, dass das Protein und die Tags im selben Leserahmen sind.

Stop Codons sind meist im Expressionsvektor nach der MCS enthalten. Die Benutzung des eigenen Stop Codons der cDNA Sequenz ist aber vorzuziehen, da dann keine zusätzliche AA an C-Terminus des Proteins hängen. Falls aber ein Tag an den C-Terminus gehängt wird, muss natürlich das Stop Codon entfernt werden, da sonst der Tag nicht übersetzt wird. Ist am C-Terminus ein Tag, muss auch noch darauf geachtet werden, dass der Leserahmen des Proteins und des Tags einen gemeinsamen Leserahmen bilden. Es kann daher sein, dass nach dem letzten für eine Aminosäure kodierenden Codons noch zusätzlich Nucleotide (1 oder 2) eingefügt werden müssen, um einen durchgehenden Leserahmen zu erhalten.

In unserem Fall hat der Expressionsvektor einen N-terminalen Tag (Glutathion-S-Transferase), Auch hier ist darauf zu achten, dass ein durchgehender Leserahmen vorhanden ist. Das zu exprimierende Gen muss daher im selben Leserahmen wie der Tag sein. Daher kann es nötig sein, 1-2 zusätzliche Nucleotide in den Primer einzufügen.

Beispiel Forward Primer: 5'-GAATTC- N₁₋₂-ATG-gagctagatagtaataatgata-3'

Beispiel Reverse Primer 5' GAATTC-N₁₋₂-atatataaccctgctcgt-3'

- Überprüfen Sie daher an Hand ihres Expressionsvektors wie sich die Klonierung ihres cDNA Konstrukts in den Expressionsvektor auf den Leserahmen auswirkt. Adaptieren Sie dann die Sequenz des Primers entsprechend. Verwenden Sie dabei aber keine Ns, sondern A, G C, oder T.
- Nun ist ihr Primer fertig. Bevor Primer bestellt werden, muss aber noch Folgendes beachtet werden:
 - Das Hinzufügen zusätzliche Nucleotide an den Enden der Oligonucleotide ist oft notwendig, damit die Restriktionsenzyme verlässlich schneiden können. Dafür kann man entsprechende Listen der Hersteller konsultieren.
 - Die Annealing Temperatur der Primer sollte ähnlich sein. Die Temperatur wird dabei nur aus der Sequenz des cDNA Teils des Primers berechnet (ohne zusätzliche Nucleotide oder Restriktionsenzymstichstellen). Um halbwegs gleiche Temperaturen zu erreichen, kann daher der cDNA Sequenz Anteil der Primer noch verkürzt oder verlängert werden.
 - Analyse der Sekundärstruktur der Primer: Primer können Strukturen ausbilden, die die PCR

Reaktion stören. Für die Analyse gibt es verschiedene Online Programme bzw kann das auch im Gene Runner gemacht werden. Basierend auf den Ergebnissen kann die DNA Sequenz der Primer noch entsprechend modifiziert werden ohne dabei die Aminosäurezusammensetzung des kodierenden Bereichs des Gens zu ändern.

B. Ligationsunabhängige Klonierung

- Basierend auf der cDNA Sequenz, der Sequenz des Expressionsvektors und passenden Restriktionsenzymstellen werden Oligonukleotide für eine Ligationsunabhängige Klonierung designt. Als Beispiel verwenden Sie das In-fusion System von Takara.
- Öffnen Sie die Website <https://www.takarabio.com/learning-centers/cloning/primer-design-and-other-tools> und folgen Sie den Anweisungen auf der Website.
- Bei dieser Art der Klonierung ist im Prinzip nur eine Restriktionsschnittstelle im Expressionsvektor nötig ist und es spielt auch keine Rolle mehr ob das Restriktionsenzym im cDNA Fragment schneidet, da sie es nicht mehr schneiden müssen. Wählen sie also entweder eine oder zwei Restriktionsschnittstellen.
- Beachten sei aber wieder, dass ihre Sequenzen in frame mit ihren Tags sein müssen. Das Programm macht das nicht.

15. Abgabe

- Abgabe eines DIN-A4-Blatts mit der Nummer des Ausgangsplasmids, den ausreichend beschrifteten Gelbildern der präparativen Plasmidverdaue und Kontrollverdaue sowie des ausgewerteten Sequenzierungsergebnisses (welches *dfv* Gen haben Sie umklont?). Wählen Sie außerdem eine Kombination aus 2 Enzymen aus, die zur Überprüfung des Vektors verwendet werden könnte. Dabei soll ein Restriktionsenzym im Vektor und eines in der cDNA schneiden. Geben Sie die Enzyme sowie das zu erwartende Schnittmuster (= Größen der Fragmente in kb) an. Verwenden sie nur Enzyme die ein sinnvolles Schnittmuster ergeben.
- Abgabe des Teils B als doc Datei per e-mail an bernhard.seiboth@tuwien.ac.at mit folgenden Ergebnissen:
 1. Ihre Original Query DNA Sequenz
 2. Nukleotid/Aminosäure Alignment
 3. Extrahierte Protein Sequenz und cDNA Sequenz
 4. Name des Expressionsvektors und der 2 Restriktionsenzyme für Klonierung
 5. Sequenz Forward und Reverse Primer für Klonierung mittels konv. Klonierung
 6. Sequenz Forward und Reverse Primer für Klonierung durch In-fusion System

Plasmidkarten:



