

DNA-Klonierung und *in silico* Analysen

Diese Übung besteht aus zwei Teilen:

Im **Teil A** wird ein Gen über eine klassische und einfache Methode aus einem high-copy Plasmid (pJET1.2-DFR) in ein low-copy Plasmid (pGEX-6P-1) umkloniert. Klonierungsvektoren wie pJET1.2 werden häufig dazu verwendet, um PCR Produkte direkt nach der Amplifikation zu klonieren. In einen weiteren Schritt können sie dann aus dem Vektor wieder herausgeschnitten werden, um sie in einen Expressionsvektor wie pGEX-6P-1 zur heterologen Expression umzuklonieren. Das umzuklonierende Gen ist die Dihydroflavonal-4-reduktase (DFR), die in 2 Varianten vorliegt, die sich auf DNA und Aminosäureebene unterscheiden. Nach der Umklonierung bestimmen Sie durch Sequenzierung, um welche der Variante es sich bei ihrer Probe handelt. Das Beispiel umfasst alle klassischen Klonierungsschritte: präparativer Restriktionsverdau, Agarosegelelektrophorese, DNA Gelextraktion, Ligation, Transformation von *E. coli*, Selektion, Plasmidpräparation, analytischer Restriktionsverdau, DNA-Sequenzierung und Analyse.

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 aus ihrer DNA Sequenz die cDNA Sequenz extrahieren und Oligonukleotide zur Amplifikation dieser cDNA für die Klonierung in einen Expressionsvektor 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 der Vektoren befinden sich am Ende dieser Vorschrift.

Die Sequenz der 2 Varianten des dfr Gens 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 (1000x) (K)

Materialien: Erlenmeyerkolben mit Deckel oder Schottflasche
Eproutetten 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 geeigneten Volumen dest. H₂O lösen
- mit dest. H₂O auf gewünschtes Volumen auffüllen
- je 5 mL in Eproutetten aliquotieren (Dispenser bei Betreuer/innen erhältlich)
- Eproutetten 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)
rCutSmart™ Puffer für Restriktionsenzyme (K)
Plasmid pJET1.2-DFR (G)
Plasmid pGEX-6P-1 (G)
sb Wasser

Geräte: Inkubator oder Thermoblock

Durchführung:

- Die pJET1.2-DFR Plasmide befinden sich in RG in der Plasmidbox für Übung 6 und sind durchnummeriert. Entnehmen Sie ein RG, das eine der 2 Varianten des dfr Gens enthält, und verwenden Sie das darin befindliche Plasmid als ihr Ausgangsplasmid für die Klonierung. Notieren Sie die Nummer, damit nachvollziehbar ist, um welches Gen es sich handelt.
- Reagenzien falls nötig auftauen
- alle Reagenzien schütteln, kurz abzentrifugieren und auf Eis stellen
- Reaktionsansätze je Plasmid wie folgt pipettieren:

Plasmid-DNA	2 µg
10x rCutSmart 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 für 4 h oder 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)
Gelelektrophoreseschlitten (A)
Gelelektrophoresekamm (A)
Kleine Wasserwaage (A)
Pipette (A)
Meßzylinder (A)

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

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 µL SYBR Safe versetzen, schwenken, in den Schlitten gießen
- eventuell entstandene Luftblasen entfernen (gelben Pipettenspitzen sind dabei hilfreich)
- Gel ca. 20 min aushärten lassen
- gesamte Plasmidverdaue mit entsprechendem Volumen 6x Purple Loading Dye mischen
- Gel samt Schlitten in die Gelkammer legen
- vorsichtig Kamm entfernen
- überprüfen, ob genug 1x TAE Laufpuffer in der Laufkammer ist
- 5 µL DNA-Marker in die erste Geltasche pipettieren
- Plasmidverdaue getrennt in die weiteren Geltaschen laden, dabei immer eine Tasche freilassen
- Deckel auf Gelapparatur setzen, Laufrichtung überprüfen und mit Power Supply verbinden
- Gel für ca. 50 min bei 90 V laufen lassen

- Gel mit dem Schlitten auf den Gel-Imager legen und analysieren
- falls eine ausreichende Auftrennung der DNA Fragmente erfolgt ist, das Gel ohne Schlitten in die ChemiDoc legen, fotografieren und das Bild speichern (UV Expositionszeit möglichst kurzhalten)
- Überprüfen Sie an Hand des Bildes ob die erwarteten DNA Banden vorhanden sind und extrahieren Sie dann die richtigen Fragmente (Punkt 4)
- falls keine ausreichende Auftrennung erfolgt ist, das Gel weiterlaufen lassen
- Zum Schluss Gießkammer, Schlitten und Kamm mit dest. H₂O reinigen und retournieren

4. Gelextraktion der DNA Fragmente

Lösungen: Monarch DNA Gel Extraction Kit (B)

Materialien: Skalpell (A)

Geräte: ChemiDoc (neben A)

Thermoblock

Durchführung: Wenn nicht anders angegeben, bei 13000 x g zentrifugieren. Überlegen Sie sich vorher, welches Bandenmuster Sie erwarten und welche Banden (Gen bzw. Zielplasmid) Sie ausschneiden.

- Gel ohne Schlitten in die ChemiDoc auf den Blue Screen legen
- Schutzbrille aufsetzen
- UV Licht einschalten
- Zügig mit dem Skalpell die richtigen Banden getrennt ausschneiden, Agarosegelstücke bei Bedarf eventuell noch zuschneiden, um benötigte Menge an Gel Dissolving Buffer zu minimieren. Gelstücke getrennt in vorher abgewogene 2-mL-RG geben
- UV Licht ausschalten, Screen mit dest. H₂O reinigen
- alle verwendeten Materialien (Schlitten, Skalpell etc.) mit dest. H₂O reinigen und retournieren
- 2-mL-RG mit Gelstücken abwägen
- 4 Vol Monarch Gel Dissolving Buffer zugeben (Beispiel: 400 µl Puffer pro 100 mg Gelstück). Falls das Gelstück schwerer als 150 mg ist, 3-3,5 Vol Monarch Gel Dissolving Buffer zugeben.
- RG bei 50 °C inkubieren, invertieren sie dabei regelmäßig das RG, bis das Gelstück vollständig aufgelöst ist. Dauer ca. 5-10 min.
- Setzen Sie die Säule in das Sammel RG ein und tragen Sie die Probe auf die Säule auf.
- 1 min zentrifugieren und Durchfluss verwerfen.
- Setzen Sie die Säule wieder in das Sammel RG ein.
- 200 µL DNA Wash Buffer auf die Säule geben, wieder 1 min zentrifugieren.
- Durchfluss verwerfen und noch einmal die Säule mit 200 µL DNA Wash Buffer waschen.
- 1 min zentrifugieren und dann die Säule in ein neues 1,5-mL-RG (ohne Deckel) setzen. Dabei vermeiden, dass die Säule in Kontakt mit dem Wash Buffer kommt. Im Zweifel die Zentrifugation wiederholen.
- Vorsichtig 10 µL DNA Elution Buffer in die Mitte der Säule pipettieren, 1 min einwirken lassen und wieder 1 min zentrifugieren.
- Eluate mit DNA aus deckellosen RG in neues 1,5-mL-RG transferieren.
- DNA-Konzentration mit dem NanoDrop messen (kontaktieren Sie dazu eine/n Betreuer/in).
- Eluate 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 sollen optimalerweise zwischen 20-50 ng Plasmid eingesetzt werden. Die benötigte ng Menge an Insert für die Ligationsreaktion kann durch folgende Formel berechnet werden:

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

Größe *dfc* Gen: 1,026 kb
Größe pGEX-6P-1: 4,984 kb

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

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

Lösungen: kompetente *E. coli* TOP10 Zellen (bekommen Sie von Betreuer/in)

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 zu den *E. coli* TOP10 Zellen pipettieren und vorsichtig mischen, dabei möglichst immer auf Eis lassen und nie erwärmen
- 30 min auf Eis inkubieren
- Hitzeschock für 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. Zur Lagerung mit Parafilm verschließen und bei 4°C aufbewahren.

7. Anzucht von *E. coli* Einzelkolonien für die Plasmidextraktion

Lösungen: Ampicillin (1000-fach) (K)

Materialien: Zahnstocher (E)

Geräte: Schüttelinkubator

Durchführung:

- fünf 5-mL-Aliquote LB-Medium in Epprovetten mit Ampicillin versetzen
- jede Epprovette 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. Dafür Kulturen erst am späteren Nachmittag beimpfen.

8. Isolierung der Plasmid-DNA

Lösungen: Monarch Plasmid Miniprep Kit (B)

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

- 2 mL jeder der vier *E. coli* Kulturen in ein 2-mL-RG überführen
- 30 s zentrifugieren
- Überstand verwerfen
- weitere 2 mL Kultur hinzufügen und noch 1x 30 s zentrifugieren

- Falls noch Kulturflüssigkeit vorhanden, nochmals wiederholen
- nach letztem Zentrifugieren Überstand komplett abnehmen
- Zellpellet in 200 µl Plasmid Resuspension Buffer (B1, pink). durch Schütteln oder pipettieren vollständig resuspendieren
- Zellen durch Zugabe von 200 µl Plasmid Lysis Buffer (B2, blau/grün) lysieren. Das RG sofort umdrehen und 5-6 Mal vorsichtig schwenken, bis die Farbe in ein dunkles Rosa übergeht und die Lösung klar und zähflüssig ist. Nicht vortexen! Eine Minute lang inkubieren.
- Zugabe von 400 µl Plasmid-Neutralisierungspuffer (B3, gelb). Das RG vorsichtig invertieren, bis die Farbe gleichmäßig gelb ist und sich ein Präzipitat bildet. Nicht vortexen! 2 Minuten inkubieren.
- 5 min bei 16000 x g zentrifugieren.
- Überstand vorsichtig in Säulchen mit RG pipettieren und 1 min zentrifugieren.
- Durchfluss verwerfen.
- Säulchen wieder in RG stellen und 200 µl Plasmid Wash Buffer 1 auf das Säulchen pipettieren
- 1 min zentrifugieren.
- 400 µl Plasmid Wash Buffer 2 hinzufügen und für 1 min zentrifugieren.
- Überführen Sie die Säule in ein sauberes 1,5-ml-RG. Achten Sie darauf, dass die Spitze der Säule nicht mit dem Durchfluss in Berührung kommt. Im Zweifelsfall noch 1x zentrifugieren Deckel des RGs entfernen.
- Tragen Sie 30 µl DNA Elution Buffer in die Mitte der Matrix auf. Warten Sie 1 min und zentrifugieren Sie dann 1 min., um die DNA zu eluieren.
- Eluat in neues RG (mit Deckel) transferieren
- DNA-Konzentrationen mit dem NanoDrop messen (kontaktieren Sie dazu eine/n Betreuer/in)
- Weiterverwenden oder auf -20°C aufbewahren
- 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

9. Kontrollverdau (Restriktionsanalyse)

Lösungen: Restriktionsenzyme *PstI*-HF, *BamHI*-HF und *EcoRI*-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)
 Gelelektrophoreseschlitten (A)
 Gelelektrophoresekamm (A)
 Kleine Wasserwaage (A)
 Pipette (A)

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

Führen Sie 2 Kontrollverdau durch: einer soll das Plasmid linearisieren (*PstI*-HF) und der zweite daraus das Insert wieder herausschneiden (*BamHI*-HF und *EcoRI*-HF). Stellen Sie für beide Kontrollverdau je einen Mastermix her.

Anzahl der Proben je Mastermix:

Plasmid Proben	1
Aufschlag	1
Gesamtanzahl	2

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 den 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 zur Sequenzierung abgeben
- das Sequenzierergebnis wird per e-mail an Sie übermittelt
- Kontrollieren Sie die Sequenz des für sie sequenzierten Teils des *dfr* Gens und identifizieren Sie um welche der 2 Varianten es sich handelt. Die Sequenzen der verschiedenen *dfr* Gene werden dazu 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)
catgagatctagctagctgat... (ihre Sequenz).
- Klicken Sie auf den BLAST Button, um ihre 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. An Hand dieser Abbildung lässt sich 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 negativ (-1, -2, -3), dann müssen sie ihre DNA Sequenz mit Hilfe eines geeigneten Programmes umdrehen.

Durchführung:

- Überprüfen Sie die Orientierung ihres Gens an Hand des BLAST Ergebnisses
- Falls sich das Gen in der falschen Orientierung befindet, wiederholen Sie 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 (aktueller Link wird am Kursanfang zur Verfügung gestellt) oder ein anderes DNA Programm. Im folgenden die Anleitung für den Gene Runner, die Schritte können aber mit jedem DNA Programm durchgeführt werden.
- 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 sich die ersten Alignments für die folgende Analyse aus.

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

Es sind nun alle Voraussetzungen 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 statt 6 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). Stop Codons 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 sbjct 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 in ihrem Fall immer im selben Leserahmen. Bei anderen Genen kann 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.
- Im nächsten Schritt erfolgt nun die Identifizierung der Exons/Introns. Auf das erste Exon mit dem Start Methionin 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 passendste GT.
- Gehen Sie dann zum Ende des Introns. Das Intron endet nach einem AG, gefolgt vom nächsten Exon. Nehmen Sie wieder ihren blastx Vergleich. Suchen Sie das zweite Exon und 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, muss aber nicht. Vergleichen sie dazu einfach 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, kodierenden Sequenz. Es kommt häufig vor, dass ein Intron ein Triplett unterbricht: 1-2 Nucleotide eines Codons befinden sich dann vor dem Intron (vor dem GT), während die dazugehörigen Nucleotid(e) nach dem AG

folgen. Falls kein vollständiges Triplet durch Herausschneiden ihres Introns entsteht, müssen Sie Anfang und Ende ihres Introns noch einmal überprüfen und dabei nach Alternativen für ihre gewählten GTs oder AGs 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 verwandten Proteine entsprechen. Es sollen also möglichst keine Aminosäuren 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 AGTNNNNNNNNNAGCC: Das Triplet AGT kodiert für ein Serin. Aber durch die Entfernung der Intronsequenz, also GT bis einschließlich AG (unterstrichen), entsteht ein neues Triplet und es kommt zu A zwei C hinzu. ACC kodiert für ein Threonin.

- 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 aus diesem Alignment die Aminosäuresequenz.
- Extrahieren Sie nun aus dem Alignment den kodierenden Teil der cDNA Sequenz beginnend beim ersten ATG bis zum Stop Codon.
- Führen Sie eine Kontrolle der zwei erhaltenen Sequenzen durch:
Kodierender Teil der cDNA Sequenz: Es muss ein durchgehender Leserahmen ohne unterbrechende Stops vorhanden sein. Das können Sie z.B. im Gene Runner machen, indem sie die Sequenz nehmen, hineinkopieren und übersetzen.
Extrahierte Aminosäuresequenz: Das extrahierte Protein können Sie mit dem übersetzten Protein der cDNA aus dem Gene Runner vergleichen. Nehmen Sie dazu statt blastx (siehe oben) den 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. Mit diesem Vergleich zu den nahen verwandten Proteinen können Sie sehen, ob ihr Protein von der Größe passt, also 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 ihres Gens werden nun Oligonukleotide für eine *in silico* Amplifikation der cDNA designt, um ihre cDNA in einen passenden 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) 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 List non-cutters an. Vergleichen Sie nun die Restriktionsenzyme der MCS mit den erhaltenen Non-Cuttern.
- Wählen Sie 2 Restriktionsenzyme, die im Expressionsvektor in der MCS schneiden aber nicht in der cDNA, aus. Am besten geeignet sind Restriktionsenzyme, die am 5' Ende und am 3' Ende der MCS schneiden. Wichtig ist es, die Reihenfolge der Restriktionsenzyme in der MCS für das

Primerdesign zu beachten, damit die cDNA richtig orientiert ist. Die Restriktionsschnittstelle für den Forward Primer soll also am 5' Ende sein.

- 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 Nucleotide des Anfangs der cDNA. Um die cDNA in ein Plasmid zu klonieren, wird noch zusätzlich die Sequenz für eine Restriktionsenzymsschnittstelle an den Anfang des Primers gehängt: in diesem Beispiel GAATTC für *EcoRI*.

Beispiel Forward Primer: 5'-GAATTC-ATG-gagctagatagtaata-gata-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 so 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)

Nach Reverse and Complement: 5' GGATCC-CTA-atatatacccc-3'

Nun hängt es noch von der Art ihres Expressionsvektors ab, ob die Oligonucleotide noch weiter verändert werden müssen. Expressionsvektoren enthalten oft Tags, die N oder C-Terminus an das Protein angehängt werden. Daher ist zu beachten, dass ihr Gen für 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 werden soll, muss aber das Stop Codon ihres Gens entfernt werden, da sonst der Tag nicht übersetzt wird. Soll am C-Terminus ein Tag angehängt werden, muss auch noch darauf geachtet werden, dass der Leserahmen ihres Gens und der des Tags einen gemeinsamen Leserahmen bilden. Es kann daher sein, dass nach dem letzten für eine Aminosäure kodierenden Codons ihres Gens noch zusätzlich Nucleotide (1 oder 2) eingefügt werden müssen, um einen durchgehenden Leserahmen zu erhalten.

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

In unserem Fall hat der Expressionsvektor mit der Glutathion-S-Transferase einen N-terminalen Tag, 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 Forward Primer einzufügen.

Beispiel Forward Primer: 5'-GAATTC- N₁₋₂-ATG-gagctagatagtaata-gata-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 für den Primer keine Ns, sondern A, G, C, oder T.
- Nun sind ihre 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 konsultieren Sie entsprechende Listen der Hersteller.
 - Die Annealing Temperaturen der Primer sollte ähnlich sein. Die Temperatur wird dabei nur aus der Sequenz des cDNA Teils des Primers berechnet (ohne zusätzliche Nucleotide oder Restriktionsenzymsschnittstellen). Um halbwegs gleiche Temperaturen zu erreichen, wird daher der cDNA Sequenz Anteil der Primer noch verkürzt oder verlängert.
 - Analyse der Sekundärstruktur der Primer: Primer können Strukturen ausbilden, die die PCR 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 daß dabei die Aminosäurezusammensetzung des kodierenden Bereichs des Gens verändert werden darf.

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 eigentlich nur eine Restriktionsschnittstelle im Expressionsvektor nötig ist und es spielt auch keine Rolle mehr ob das Restriktionsenzym im cDNA Fragment schneidet, da Sie es vorher nicht mehr schneiden müssen. Wählen Sie also entweder eine oder zwei Restriktionsschnittstellen für die Klonierung.
- Beachten sei aber wieder, dass ihre Gensequenz in frame mit ihren Tag sein muss. Das Programm macht das nicht für Sie.

15. Abgabe

Erfolgt als ein pdf im TUWEL

Von Teil A:

1. Nummer des Ausgangsplasmids
2. Ausreichend beschriftete Gelbilder der präparativen Plasmidverdaue und Kontrollverdaue
3. Ausgewertetes Sequenzierungsergebnis (welches *dfr* Gen haben Sie umkloniert?).

Von Teil B:

4. Ihre Original Query DNA Sequenz
5. Nukleotid/Aminosäure Alignment
6. Extrahierte Protein Sequenz
7. Extrahierte cDNA Sequenz
8. Name des Expressionsvektors und der 2 Restriktionsenzyme für Klonierung
9. Sequenz Forward und Reverse Primer für Klonierung mittels konv. Klonierung
10. Sequenz Forward und Reverse Primer für Klonierung durch In-fusion System

Plasmidkarten:



