

7. Appendices

Appendix 1 Learning Contract

BU Bournemouth University		LEARNING CONTRACT: INDEPENDENT RESEARCH PROJECT	
Student Name:	Lydia Doherty		
Degree Programme:	BSc (Hons) Biological Sciences		
Proposed Project Title:	Investigating the location of the molecular recognition site of the glucuronide transporter (gusB) of <i>Escherichia coli</i> , using gene fusion with xylE.		
Supervisor:	Dr Wei-Jun Liang		
Research Proposal Attached	<input checked="" type="checkbox"/> YES	<input type="checkbox"/> NO	and includes:
<input checked="" type="checkbox"/> YES	<input type="checkbox"/> NO	Risk Assessment for fieldwork and evidence of COSHH assessment for all laboratory procedures (online risk assessment completed)	
<input type="checkbox"/> YES	<input checked="" type="checkbox"/> NO	Completed booking forms for all field equipment	
<input type="checkbox"/> YES	<input checked="" type="checkbox"/> NO	Letters of permission where appropriate providing evidence of access to such things as field sites and/or museum archives	
<input type="checkbox"/> YES	<input checked="" type="checkbox"/> NO	Completed Ethics Checklist	
Copies of all relevant forms may be found on myBU - SciTech tab - Projects - Project Forms			
INTERIM INTERVIEW - Progress evaluation			
<p>The nature of this review should be clearly defined and agreed. Please complete the box below with the agreed details including the agreed submission date which is normally the first week of November in Level 6/H. Submission is via a formal tutorial with the supervisor.</p> <p>Looking at the quality of laboratory work which has already commenced.</p> <p>Checking the progress of the laboratory work and/or writing up.</p>			
Assessment Due:	17/11/2017		
FINAL ASSESSMENT - RESEARCH PAPER/REPORT			
<p>This assessment is normally governed by the guidance provided in the Independent Research Project Guide. Any variance in terms of format and word limit should be agreed and specified in the box below. Submission date cannot however be changed unless evidence of mitigating circumstances are provided in accordance with the standard BU Guidelines.</p> <p>Maximum of 12,000 words</p>			

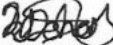
As the student undertaking the above project I agree to:

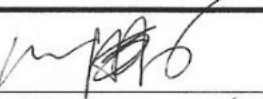
- E-mail my supervisor on a fortnightly basis with a progress report
- Meet with my supervisor at least once a month to discuss progress and I understand that it is my responsibility to organise these meetings
- Comply with the terms of this learning contract and the guidance set out in the Guide to Independent Research Projects
- I understand that this is an *independent* project and that I am solely responsible for its completion
- I agree to comply with all laboratory and fieldwork protocols established by the Faculty.

As the supervisor of this project I agree to:

- Meet with the student undertaking this project on at least a monthly basis and to respond to the progress e-mails as appropriate
- To meet formally with the student during the first week in November to undertake the interim interview
- To provide guidance and support to the student undertaking this project bearing in mind that it is an *independent* research project. This is inclusive of commenting on drafts of the final report in a timely fashion.

Both of the undersigned parties agree to be bound by this learning contract:

Student Signature:	
PRINT NAME:	Lydia Kristen Doherty
Date:	30/06/2017

Supervisor Signature:	
PRINT NAME:	Wei-Jun Liang
Date:	29/06/2017

When completed, this form should be handed in to SciTech Admin (C114) and a copy retained by the student to be included in an appendix to the final IRP document.

Appendix 2 Interim Report

Independent Research Project Interim Interview - Agreed Comments Form

Student Name: Lydia Doherty	Programme: BSc (Hons) Biological Sciences
Date: 16.11.17	IRP Title: Deciphering the location of molecular recognition on the glucuronide transporter (GusB) by gene fusion.
Supervisor Name: Wei-Jun Liang	



Agreed comments – to include progress and plans for completion:

Progress: Lab work is ~33% complete. Have successfully amplified half of my gene fragments in preparation for fusion PCR. Introduction is also ~33% complete. I have made a clear plan of where I intend to take this project. *Currently read and critically analysed ~40 scientific papers.*

Plans for completion: I plan to carry on writing my introduction alongside conducting my laboratory work. I will continue with primary PCR until I have at least 4 pairs of gene fragments successfully amplified. I currently have 2 pairs. I will then move forward with the rest of my laboratory work. I will design and order flanking primers to test whether there is a functional copy of XylE in the genome of *E. coli* mc1061. If I have time I will attempt to fuse GusB and XylE together, separated by GFP. In addition, I will attempt to separate the two domains of GusB with GFP to see if there's still functionality.

*Will read more scientific papers.
Create null hypothesis and alternative hypothesis.
Create a more detailed gantt chart.*

Two copies of this form are needed – student to retain one copy and include in the appendices of the dissertation the other is to be emailed to the supervisor.

Student Signature: 	Supervisor Signature:  17-11-2017
---	--

Appendix 3 COSHH Risk Assessment

Bournemouth University COSHH ASSESSMENT FORM DNA EXTRACTION AND PCR

1. Assessor: L. Doherty		2. Assessment Date: 24/03/2017		3. Assessment Review: Next review date: 24/03/2018	
4. Summary of process or method (or make specific reference to written protocol to be used): It will involve DNA extraction from non-pathogenic <i>E. coli</i> , DNA amplification by PCR, electrophoresis, making transformant, restriction digest, functionality tests and glucuronide transport assays using chromogenic chemicals, such as X-Gluc, p-nitrophenol, p-nitrophenyl glucuronide, TrisBase, EDTA, SDS, NaOH, Ethanol.					
5. Key Activity/Task (in relation to exposure potential e.g. mixing, filling, spraying, etc.): Mixing chemicals using pipettes, handling bacteria, weighing powders				6. People who could come to harm (number & roles e.g. students) Staff and students in lab.	
7. Duration of Exposure (minutes, hours and how often): 3 – 8 hours, up to five days a week.				8. Location and Conditions of Use (e.g. lab, room, temp etc.): Dorset house DNA lab	
9. Hazardous ingredients: (copy form/add more rows as req'd)	10. Quantities Used	11. Workplace Exposure Limit (WEL)	12. Risk/safety phrases	13. Actual Potential Route of Exposure (E.g. by inhalation)	14. Datasheet Attached? Y/N
A Multiple <i>E. coli</i> strains	100 ml	N/A	Non-pathogenic	Inhalation, skin/eye contact, ingestion	N/A
B 10 mM dNTP	>10 ml	N/A	Not hazardous	Inhalation, skin/eye contact, ingestion	N/A

I: Health & Safety/Public/Hazardous Substances/COSHH assessment form 2003

Version 1 May 1

C	1M MgCl ₂	>10 ml	N/A	Not a hazardous substance or mixture according to Regulation (EC) No. 1272/2008. Not a hazardous substance or mixture according to EC-directives 67/548/EEC or 1999/45/EC.	Inhalation, skin/eye contact, ingestion	N/A
D	PCR primers	>10 ml	N/A	Not hazardous	Inhalation, skin/eye contact, ingestion	N/A
E	GoTaq polymerase	>10 ml	N/A	Not hazardous	Inhalation, skin/eye contact, ingestion	N/A
F	Agarose	100 g	N/A	Not Hazardous	Inhalation, skin/eye contact, ingestion	N/A
G	TAE	> 1 L	N/A	Not Hazardous In accordance with European Directive 1999/45/EC and Regulations 1907/2006/EC and 1272/2008/EC	Inhalation, skin/eye contact, ingestion	N/A
H	SYBR-Safe	>10 ml	N/A	Not hazardous	Inhalation, skin/eye contact, ingestion	N/A
I	6X Blue sample buffer	>10 ml	10 mg m ⁻³	Not hazardous	Inhalation, skin/eye contact, ingestion	N/A
J	5X Flexi buffer	>10ml	N/A	Not hazardous	Inhalation, skin/eye contact, ingestion	N/A
K	LB agar	>2L	N/A	Not a hazardous substance or mixture according to Regulation (EC) No. 1272/2008.	Inhalation, skin/eye contact, ingestion	N/A
L	CaCl ₂	500g	N/A	Irritating to the eyes	Inhalation, skin/eye contact, ingestion	N/A

I: Health & Safety/Public/Hazardous Substances/COSHH assessment form 2003

Version 1 May 2

M	0.1M X-Gluc	>10ml	N/A	R20/21/22 - Harmful by inhalation, in contact with skin and if swallowed. H302 - Harmful if swallowed. H312 - Harmful in contact with skin. H332 - Harmful if inhaled. P261 - Avoid breathing dust/fume/gas/mist/vapours/spray. P280 - Wear protective gloves/protective clothing/eye protection/face protection.	Inhalation, skin/eye contact, ingestion	N/A
N	EDTA	>100g	N/A	Harmful by inhalation, contact with skin and eyes and if swallowed. R36/37/38 - Irritating to eyes, respiratory system and skin S26 - In case of contact with eyes, rinse immediately with plenty of water and seek medical advice S24/25 - Avoid contact with skin and eyes S36/37/39 - Wear suitable protective clothing, gloves and eye/face protection	Inhalation, skin/eye contact, ingestion	N/A
O	10mM P-nitrophenol	>10ml	N/A	Not a hazardous substance or mixture according to Regulation (EC) No. 1272/2008	Inhalation, skin/eye contact, ingestion	N/A
P	P-nitrophenyl glucuronide	>2g	N/A	Not a hazardous substance or mixture according to Regulation (EC) No. 1272/2008.	Inhalation, skin/eye contact, ingestion	N/A
Q	Ampicillin	10g	N/A	Irritating to the eyes, respiratory system and skin H334: May cause allergy or asthma symptoms or breathing difficulties if inhaled. H317: May cause an allergic skin reaction. R42/43: May cause sensitisation by inhalation and skin contact. P261 Avoid breathing dust/ fume/ gas/ mist/ vapours/ spray. P280 Wear protective gloves. P284 Wear respiratory protection	Inhalation, skin/eye contact, ingestion	N/A

I: Health & Safety/Public/Hazardous Substances/COSHH assessment form
2003

Version 1 , May

3

15. Control Measures												
1. Standard lab procedures: - no eating, drinking or other hand to mouth contact. Despite most chemicals being described as non-hazardous, gloves, safety spectacles and lab coats must be worn at all times. Nitrile (or equivalent) gloves, safety spectacle and lab coat must be worn. Wipe spillages and clean surfaces. Tie hair back. Wash hands before leaving lab. 2. 6X Blue sample buffer will be used in very small quantities and this should ensure that the OEL is not exceed. However, good general ventilation will be maintained throughout the procedure. Now mark in the letters from the list of 'Hazardous Ingredients' above to indicate potential danger:												
16. Indication of Danger			17. Route of Exposure		18. Chemical State		19. Flammability		20. Volatility		21. Dust rating	
Very Toxic		Irritant	C, G, Q, N	Inhalation	A-Q	Solid	F, P, N, L	Flammable		Low	All	Low
Toxic		Sensitiser		Skin Contact	A-Q	liquid	A-E & G-L, M, O, Q	Highly flammable		Medium		Medium
												F, P, N, L, F, Q
Corrosive		Carcinogen		Eye Contact	A-Q	Gas/vapour	N/A	Extremely flammable		High		High
Harmful	M, N, Q	Mutagenic		Swallowing	A-Q			Oxidising				
Biological Agent		Toxic to reproduction		Injection	N/A			Explosive				
22. First Aid Procedures (as advised from Material Safety Data Sheet)												
If inhaled			If skin contact			If eye contact			If swallowed		If injected	
Supply fresh air; consult doctor in case of complaints.			Immediately wash with soap and water and rinse thoroughly.			Rinse open eye for several minutes under running water. If symptoms persist, consult a doctor.			Washout mouth with water provided person is conscious. Call a physician		Seek urgent medical aid.	
23. Spillage Procedures: →			Adhering to the control measures in section 15 above, absorb with liquid-binding material (sand, diatomite, acid binders, universal binders, sawdust). Ensure adequate ventilation.									
24. Disposal Arrangements												
Collection			Swill down sink		Evaporation		In normal waste		Other			
✓												
25. Are the risks adequately controlled? (Write 'Yes' or 'No'):					YES							

If you decide that the controls in Section 15, are sufficient, skip to section 27. If you decide that the risks are NOT adequately controlled (or you're not sure), then you will need to give special instructions to control the risk.

26. Special Instructions to control the risk:	
27. Ensure those affected are informed of the Risks & Controls - Confirm how this will be done e.g. by issuing written instructions:	
1. Practical leader and any other person in the lab will be given a copy of this assessment.	

Appendix 4 GusB-XylE pairwise protein sequence alignment

```
#####
# Program: needle
# Rundate: Wed 13 Jun 2018 08:07:16
# Commandline: needle
#
# -auto
# -stdout
# -asequence emboss_needle-I20180613-080715-0259-57572104-plm.asequence
# -bsequence emboss_needle-I20180613-080715-0259-57572104-plm.bsequence
# -datafile EBLOSUM62
#
# -gapopen 10.0          # Aligned_sequences: 2
# -gapextend 0.5         # 1: GusB_(WP_023149767.1)
# -endopen 10.0          # 2: XylE_(WP_089616118.1)
# -endextend 0.5         # Matrix: EBLOSUM62
# -aformat3 pair         # Gap_penalty: 10.0
# -sprtein1              # Extend_penalty: 0.5
# -sprtein2              #
# Align_format: pair     #
# Report_file: stdout    # Length: 614
##### # Identity:      108/614 (17.6%)
# Similarity:      174/614 (28.3%)
# Gaps:            280/614 (45.6%)
# Score: 90.0

GusB_(WP_0231      1 ----- 0
XylE_(WP_0896      1 MNTQYNSSYIFSITLVATLGGLLFYDITAVISGTVESLNTVFVAPQNLSE 50

GusB_(WP_0231      1 --MNQQLSH-----RTIVGYSLGDVAMN-----FAFAM 26
   |..|.:      ..|:|:|..|:|
XylE_(WP_0896      51 SAANSLLGFCVASALIGCIIGGALGGYCSNRFGRDRLKIAAVLFFISGV 100

GusB_(WP_0231      27 GA-----LFLLSY-----YTDVAGVGAAAGTMLLLV 53
   |:      :|..|      |..|:|..| :|..|
XylE_(WP_0896      101 GSAWPELGFTSINPDNTVPVYLAGYVPEFYRIIGGIGVGLA-SMLSPM 149

GusB_(WP_0231      54 RVFDAFADFVAGRVVDSVNTRWGKFRPF-LLFGTAPLMIFSVLVF----- 97
   :.:.....|:|      .|..| :|..| :|:..|..|
XylE_(WP_0896      150 YIAELAPAHIRGKLV-----SFNQFAIIFG--QLLVYCVNYFIARSG 189

GusB_(WP_0231      98 ---WVPTD-WSH--GSKVYAYLTMYGLGLCYSLVNIPYGSLATAMTQQP 141
   |.:| |.: :|:..|..| |.:|..| :..
XylE_(WP_0896      190 DASHLNTDGRYMFASECIPALLFLM---LLYTVPEP-----RWL 227

GusB_(WP_0231      142 QSRARLGAAGIAASLTFVCLAFILGPSIKNSSPEEMSVYHFWITVLAI 191
   :|:..|..| |..|:|:..|:| :|:..|..|
XylE_(WP_0896      228 MSRGKQEQAEGI-----LRKIMGNTLATQAVQE---IKHSLDHGRKT 266

GusB_(WP_0231      192 AGMVLFIKFKSTRENVVIVAQPSLNISLQTLKRNPLFMLCIGALCVL 241
   .|:|..|      :.:..|:|:..|
XylE_(WP_0896      267 GGRLLMF-----GVGVIVIGVMSLSI 286

GusB_(WP_0231      242 ISTFAVSASSLFYVRYVNDTG-----LFTVLVLVQNLVGTVASAPLV 284
   ..|:..|..|:|..|..| |.:|:|..|..| |.
XylE_(WP_0896      287 FQQFVGINVLVYAPVFKTLGASTDIALQTIIVGVINLTFTV----LA 332

GusB_(WP_0231      285 PGMVARIGKKNLFLIGAL-----LGTCGYL-----LFFWWSV 316
   ..|:..|:|:..|:| |..|..| :|:|..
XylE_(WP_0896      333 IMTVDKFGRKPLQIIGALGMAIGMFLGTAFYTAQPGIVALLSMLFYVAA 382

GusB_(WP_0231      317 WSL---PVALV-----ALAIASIGQGV-TMTVMMALEADTVE 349
   :.: |..|      |..|..|:|:..|..| :.:
XylE_(WP_0896      383 FAMSWGPPCVLLSEIFPNAIRGKALAIATAAQLANFYVSWTF--PMWD 430

GusB_(WP_0231      350 YGEYLTGVRIEGLTYSLSFSTRKCGQAIGGSIPAFILGLSGYIANQVQTP 399
   :.:|:..|..|:|:..| |
XylE_(WP_0896      431 KNSHLVAHFHNGFSYWIYG---C----- 450

GusB_(WP_0231      400 EVIMGIRTISIALVPCGFMLLAFVII-FYPLTDKKFKEIVVEID-----N 443
   |:|      |:..| |.:|..| |:|
XylE_(WP_0896      451 ---MGV-----LAALFMKFKVPETKGKTL---ELEALNEPE 481

GusB_(WP_0231      444 RKKVQQQLISDITN 457
   .|:|..|..|
XylE_(WP_0896      482 TKKTQQTATL---- 491

#-----
#-----
```

Appendix 5 Constituents and Methods for routinely used Solutions

1M Tris-HCl (pH 8.0)

121.14g of Tris base was dissolved in 800 ml of dH₂O. The pH was adjusted to 8.0 by adding a total of 52 ml of concentrated HCl. The solution was supplemented with dH₂O to yield a total volume of 1 litre, and sterilised by autoclaving.

0.5M EDTA.Na₂ (pH 8.0)

93.06g of EDTA.Na₂ was dissolved in 400 ml of dH₂O. 11.7g of NaOH pellets were added to adjust the pH to 8.0. The solution was supplemented with dH₂O to yield a total volume of 500 ml, and sterilised by autoclaving.

10X Tris-EDTA.Na₂ buffer (T₁₀E₁)

10 ml of 1M Tris-HCl (pH 8.0), 2 ml of 0.5M EDTA.Na₂ (pH 8.0), and 988 ml of dH₂O were combined. The 1 litre solution was sterilised by autoclaving.

50X TAE buffer (stock solution)

100 ml of 0.5M EDTA.Na₂ (pH 8.0), 57.1 ml of glacial acetic acid, and 242g of Tris base were combined. The solution was supplemented with dH₂O to yield a total volume of 1 litre, and sterilised by autoclaving.

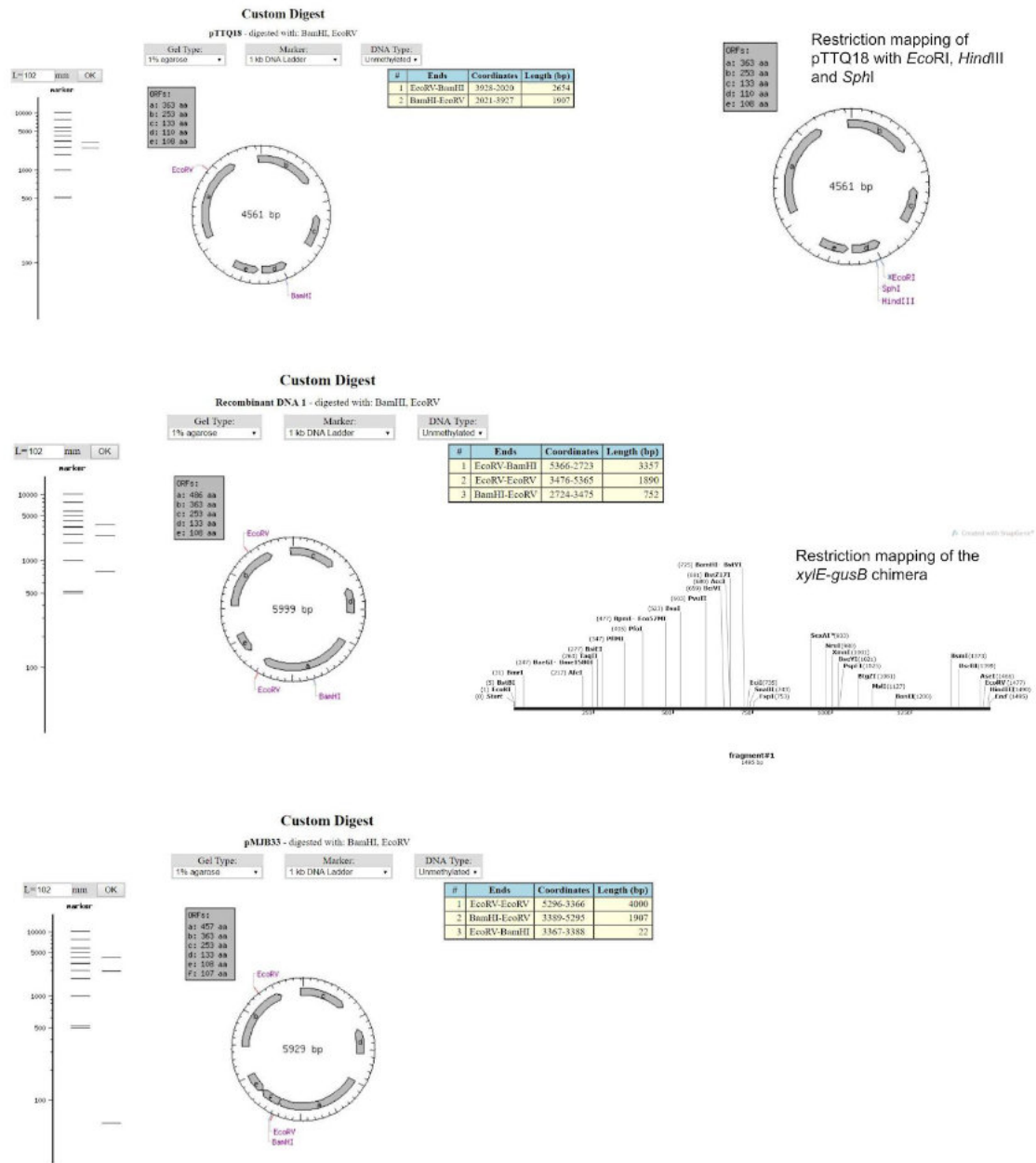
1kb DNA ladder

300 µl of T₁₀E₁, 100 µl of 5X green GoTaq® flexi buffer, and 100 µl of the defrosted DNA ladder were combined in a sterile eppendorf. The solution was vortexed to mix the constituents.

Sepharose beads

The bottled mixture was gently swirled until it was homogenous. Approximately 20 ml was poured into a sterile falcon tube. The falcon tube was centrifuged at 3000 rpm for 10 minutes within a refrigerated centrifuge. The supernatant was removed through pipetting and replaced with the same volume of fresh T₁₀E₁. This process was repeated three times to wash the sepharose beads and ensure that they were hydrated.

Appendix 6 Restriction mapping and fragment size predictions



Appendix 7 Primer Design within this project

Primer design for Xyle-GusB chimeras

XyleE (1st half) GusB (2nd half) (chosen helice to cut: ¼ GusB helice)

1. Before IC2 and IC3

231:MNTQYNSSYIFSITLVATLGGLLFGYDTAVISGTVESINTVFVAPQNLSESAANSLLGFCV
ASALIGCIIGGALGGYCSNRFGRDLSLIAAVLFFISGVGSAWPELGFTSINPDNTVPVYLAGYV
PEFVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGKLVSFNQFAIIFGQLLVYCVNYFIARSGDA
SWLNTDGWRVMFASECIPALLFLMLLYTVPESPRWLMSRG
NVVRIVAQPSLNISLQTLKRNRLFMLCIGALCVLISTFAVSASSLFYVRYVLNDTGLFTVLVLV
QNLVGTVASAPLVPGMVARIGKKNTEFLIGALLGTCGYLLFFWVSVWVSLPVALVALAIASIGQGV
MTVMWALEADTVEYGEYLTGVRIEGLTYSLSFSTRKCGQAIGGSIPAFILGLSGYIANQVQTPEV
IMGIRTSIALVPCGFMLLAFVLIWFYPLTDDKKFKEIVVEIDNRKKVQQQLISDITN

atgaataccagatataattccagttatatattttcgattaccttagtcgctacattaggt
ggtttattattttggctacgacaccgcggttatttcgggtactgttgagtcactcaatacc
gtctttgttgctccacaaaacttaagtgaatccgctgccaactccctgtaggggttttgc
gtggccagcgcctctgattggttgcatcatcggcgggtgccctcggtggttattgcagtaac
cgcttcggtcgctcgatgattcacttaagattgctgctgtcctgttttttatttctggtgta
ggttctgcctggccagaacttggttttacctctataaaccggacaaactgtgcctgtt
tatctggcaggttatgtcccggaatttggtatttatcgcatatttggcgggtattggcggt
ggttttagcctcaatgctctcgccaatgtatattgcggaactggctccagctcatattcgc
gggaaactgggtctcttttaaccagtttgcgattattttcgggcaacttttagtttactgc
gtaaaactattttattgcccgttcgggtgatgccagctgggtgaatactgacggctggcgt
tatatgtttgcctcggaatgtatccctgcactgctgttcttaattgctgctgtataccgtg
ccagaaaagtcctcgctggctgatgtcgcgcggaatgtggttacgtatcggtgcgcagccg
tcattgaatatcagtcctgcaaaccctgaaacggaatcgcccgctgtttatgttggtgcac
gggtgcgctgtgtgtgctgatttcgacctttgcggtcagcgccctcgctggttcttacgtg
cgctatgtgttaaatgataccgggctgttcactgtgctggtactgggtgcaaaacctgggt
gggtactgtggcatcggcaccgctgggtgccggggatgggtcgcgaggatcggtaaaaagaat
accttcctgattggcgctttgctgggaacctgcgggttatctgctgttcttctgggtttcc
gtctgggtcactgccggtggcggttggttgcggttgccatcgcttcaattgggtcagggcggt
accatgaccgtgatgtggcgctggaagctgataccgtagaatacgggtgaataacctgacc
ggcgtgcgaattgaagggtcacctattcactattctcatttaccgtaaatgcgggtcag
gcaatcggaggttcaattcctgcctttattttgggggttaagcggatataatcgccaatcag
gtgcaaacgccggaagttattatgggcatccgcacatcaattgccttagtaccttgcgga
tttatgctactggcattcggttattatctgggttttatccgctcacggataaaaaattcaaa
gaaatcggtggttgaaattgataatcgtaaaaaagtgcagcagcaattaatcagcgatatc
actaattaa

XyleE (1st half) GusB (2nd half) (chosen helice to cut: ¾ of the helice of GusB)

2. After IC2 and IC3

266:MNTQYNSSYIFSITLVATLGGLLFgydTAVISGTVESINTVFVAPQNLSESAANSLLGFCV
ASALIGCIIGGALGGYCSNRFGRDRLKIAAVLFFISGVGSAPWELGFTSINPDNTVPVYLAGYV
PEFVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGKLVSFNQFAIIFGQLLVYCVNYFIARSGDA
SWLNTDGWRVYMFASECIPALLFLMLLYTVPESPRWLMSRGKQEQAEGLRKIMGNTLATQAVQEI
KHSLDHGRKT
TLKRNRPLFMLCIGALCVLISTFAVSASSLFYVRYVLNDTG LFTVLVLVQNLVGTVASAPLVPGM
VARIGKKN TFLIGALLGTCGYLLFFWVSVW SLPVALVALAIASIGQGVMTVMWALEADTVEYGE
YLTGVRIEGLTYSLSFSTRKCGQAIGGSIPAFILGLSGYIANQVQTPEVIMGI RTSIALVPCGFM
LLAFVLIWFYPL TDKKFKEIWEIDNRKKVQQQLISDITN

atgaataccaggtataattccagttatatatattttcgattaccttagtcgctacattaggt
ggtttattattttggctacgacaccgcggttatttccggtactgttgagtcactcaatacc
gtctttgttgctccacaaaacttaagtgaatccgctgccaactccctgtaggggttttgc
gtggccagcgctctgattgggtgcatcatcggcgggtgccctcggtgggttattgcagtaac
cgcttcggtcgctcggtgattcacttaagattgctgctgctcctgttttttatttctgggtgta
ggttctgcctggccagaacttggttttacctctataaaccggacaaactgtgcctgtt
tatctggcaggttatgtcccggaaatttggtattttatcgcatatttggcgggtattggcggt
ggtttagcctcaatgctctcgccaatgtatattgcggaactgggtccagctcatattcgc
gggaaactgggtctcttttaaccagtttgcgattattttcgggcaacttttagtttactgc
gtaaaactattttattgcccgttcgggtgatgccagctgggtgaatactgacgggtggcggt
tataatgtttgctcggaatgtatccctgactgctgttcttaatgctgctgtataccgtg
ccagaaagtcctcgctgggtgatgtcgcgcggaagcaagaacaggcggaaggtatcctg
cgcaaaattatgggcaacacgcttgcaactcaggcagtaacaggaaattaaacactccctg
gatcatggccgcaaaaccaccctgaaacggaatcgcccgtgtttatgttggtgcatcggt
gcgctgtgtgtgctgatttcgacctttgcgggtcagcgctcgtcggtgttctacgtgcgc
tatgtgttaaataatgataccgggctgttcactgtgctggtactgggtgcaaaacctgggtgggt
actgtggcatcggcaccgctgggtgccggggatgggtcgcgaggatcggtaaaaagaatacc
ttcctgattggcgctttgctgggaacctgcgggttatctgctgttcttctgggtttccgtc
tggtcactgccgggtggcggttggttgcggtggccatcgcttcaattgggtcagggcggttacc
atgaccgtgatgtggggcgctggaagctgataccgtagaatacgggtgaataacctgaccggc
gtgcgaattgaagggtcacctattcactattctcatttaccgtaaatgcggtcaggca
atcggaggttcaattcctgcctttattttgggggttaagcggatataatcgccaatcaggtg
caaacgccggaagttattatgggcatccgcacatcaattgccttagtaccttgcggtatt
atgctactggcattcggttattatctgggtttatccgctcacggataaaaaattcaaagaa
atcgtgggttgaaattgataatcgtaaaaaagtgccagcagcaattaatcagcgatatcact
aattaa

XyleE (1st half) GusB (2nd half) (chosen helice to cut: To the left of the bundle in GusB)

3. Half way through IC2 before IC3

240:MNTQYNSSYIFSITLVATLGGLLFgydTAVISGTVESINTVFVAPQNLSESAANSLLGFCV
ASALIGCIIGGALGGYCSNRFGRDRLKIAAVLFFISGVGSAPWELGFTSINPDNTVPVYLAGYV
PEFVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGKLVSFNQFAIIFGQLLVYCVNYFIARSGDA
SWLNTDGWRVYMFASECIPALLFLMLLYTVPESPRWLMSRGKQEQAEGL
VAQPSLNISLQTLKRNRPLFMLCIGALCVLISTFAVSASSLFYVRYVLNDTG LFTVLVLVQNLVGT
TVASAPLVPGMVARIGKKN TFLIGALLGTCGYLLFFWVSVW SLPVALVALAIASIGQGVMTVMW

ALEA DTVEYGEYLTGVRIEGLTYSLFSFTRKCGQAIGGSIPAFILGLSGYIANQVQTPEVIMGIR
TSIALVPCGFMLLAFVIIWFYPLTDKKFKEIVVEIDNRKKVQQQLISDITN

atgaataccaggtataattccagttatatatattttcgattaccttagtcgctacattaggt
ggtttattattttggctacgacaccgcggttattttccgggtactggtgagtcactcaatacc
gtctttgttgctccacaaaacttaagtgaatccgctgccaactccctggttaggggttttgc
gtggccagcgctctgattggttgcatcatcggcgggtgcctcggtggttattgcagtaac
cgcttcggtcgctcggtgattcacttaagattgctgctgtcctgttttttatttctgggtgta
ggttctgcctggccagaacttggttttacctctataaaaccggacaacactgtgcctggt
tatctggcaggttatgtcccgggaatttggtatttatcgcatatttggcgggtattggcggt
ggtttagcctcaatgctctcgccaatgtatattgcggaactgggtccagctcatattcgc
gggaaactgggtctcttttaaccagtttgcgattattttcgggcaacttttagtttactgc
gtaaaactattttattgcccgttcgggtgatgccagctgggtgaatactgacggctggcggt
tatatgtttgcctcggaatgtatccctgcactgctgttcttaatgctgctgtataccgtg
ccagaaagtcctcgctgggtgatgtcgcgcggaagcaagaacaggcggaaggtatcctg
gttgcgcgagccgtcattgaatatcagtcctgcaaaccctgaaacggaatcgcccgtggtt
atgttgctgcatcggtgctgctgtgtgtgctgatttcgacctttgcgggtcagcgctcgctg
ttgttctacgtgctgctatgtgttaaatgataccggggtggtcactgtgctgggtactgggtg
caaaacctgggttggtactgtggcatcggcaccgctgggtgccccggatgggtcgcgaggatc
ggtaaaaagaataaccttctgattggcgctttgtctgggaacctgcggttatctgctgttc
ttctgggtttccgtctggtcactgccgggtggcggttggttgcggttgccatcgcttcaatt
ggtcagggcggttaccatgaccgtgatgtgggcgctggaagctgataccgtagaatacgggt
gaatacctgaccggcgctgcaattgaagggtcacctattcactattctcatttaccgt
aaatgcggtcaggcaatcggaaggttcaattcctgcctttattttggggtaagcggtat
atcgccaatcaggtgcaaacgcgggaagttattatgggcacatccgcacatcaattgcctta
gtaccttgcggtttatgtactggcattcggtatttatctggttttatccgctcacggat
aaaaaattcaaagaaatcggtggttgaaattgataatcgtaaaaaagtgcagcagcaatta
atcagcgatatcactaattaa

XyleE (1st half) GusB (2nd half) (chosen helice to cut: To the right of the bundle in GusB)

4. After IC2 half way through IC3
247:MNTQYNSSYIFSITLVATLGLLFGYDTAVISGTVESLNTVFVAPQNLSESAANSLLGFCV
ASALIGCIIGGALGICYSNRFGRRDSLKIAAVLFFISGVGSAPWPELGFTSINPDNTVPVYLAGYV
PEFVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGKLVSNQFAIIFGQLLVYCVNYFIARSGDA
SWLNTDGWRVYMFASECIPALLFLMLLYTVPESPRWLMSRGKQEQAEGLRKIMGNT
ISLQTLKRNRPLFMLCIGALCVLISTFAVSASSLFYVRYVLNDTGLFTVLVLVQNLVGTVASAPL
VPGMVARIGKKN TFLIGALLGTCGYLLFFWVSVWSLPVALVALAIASIGQVMTVMWALEADTV
EYGEYLTGVRIEGLTYSLFSFTRKCGQAIGGSIPAFILGLSGYIANQVQTPEVIMGIRTSIALVP
CGFMLLAFVIIWFYPLTDKKFKEIVVEIDNRKKVQQQLISDITN

atgaataccaggtataattccagttatatatattttcgattaccttagtcgctacattaggt
ggtttattattttggctacgacaccgcggttattttccgggtactggtgagtcactcaatacc
gtctttgttgctccacaaaacttaagtgaatccgctgccaactccctggttaggggttttgc
gtggccagcgctctgattggttgcatcatcggcgggtgcctcggtggttattgcagtaac
cgcttcggtcgctcggtgattcacttaagattgctgctgtcctgttttttatttctgggtgta

ggttctgcctggccagaacttggttttacctctataaaccggacaacactgtgcctgtt
 tatctggcagggttatgtcccggaatttggtatttatcgcatattggcggtattggcggtt
 ggttttagcctcaatgctctcgccaatgtatattgcggaactggctccagctcatattcgc
 gggaaactgggtctcttttaaccagtttgcgattattttcgggcaacttttagtttactgc
 gtaaaactattttattgcccgttcgggtgatgccagctggctgaatactgacggctggcggt
 tatatgtttgctcggaatgtatccctgcactgctgttcttaatgctgctgtataccgtg
 ccagaaagtcctcgctggctgatgtcgcgcggcaagcaagaacaggcggaaggatatcctg
 cgcaaaattatgggcaacacgatcagctctgcaaaccctgaaacggaatcgcccgtgttt
 atgttggtgcatcgggtgctgtgtgtgtgctgatttcgacctttgcgggtcagcgcctcgctg
 ttgttctacgtgctgtatgtgttaaatgataccggggtgttactgtgctgggtactgggtg
 caaaacctgggttggtactgtggcatcggcacccgtgggtgccgggggatgggtcgcgaggatc
 ggtaaaaagaataccttctgattggcgctttgctgggaacctgcgggttatctgctgttc
 ttctgggtttccgtctggtcactgccgggtggcggttggttgcggttgcccatcgcttcaatt
 ggtcagggcggttaccatgaccgtgatgtgggctggaagctgataccgtagaatacgggt
 gaatacctgaccggcggtgcgaattgaagggctcacctattcactattctcatttaccgt
 aaatgcgggtcaggcaatcggaggttcaattcctgcctttattttgggggttaagcggatat
 atcgccaatcaggtgcaaacgcggaagttattatgggcatccgcacatcaattgcctta
 gtaccttgcggtattatgctactggcattcggtattatctggttttatccgctcacggat
 aaaaaattcaaagaaatcggtggttgaaattgataatcgtaaaaaagtgcagcagcaatta
 atcagcgatatcactaattaa

XyleE (1st half) GusB (2nd half) (GusB chosen helice to cut: exactly halfway through)

5. Just after IC2. Between IC2 and IC3
 244:MNTQYNSSYIFSITLVATLGGLLFgyDTAVISGIVESINTVFVAPQNLSESAANSLLGFCV
 ASALIGCIIGGALGGYCSNRFGRDRLKIAAVLFFISGVGSAPWELGFTSINPDNTVPVYLAGYV
 PEFVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGKLVSEFNQFAIIFGQLLVYCVNYFIARSGDA
 SWLNTDGWRVYMFASECIPALLFLMLLYTVPESPRWLMSRGKQEQAEGLRKIM
 SLNISLQTLKRNRLFMLCIGALCVLISTFAVSASSLFYVRYVLNDTGLFTVLVLVQNLVGTVAS
 APLVPGMVARIGKKNFLIGALLGTCGYLLFFWVSVWSLPVALVALAIASIGQGVTMTVMWALEA
 DTVEYGEYLTGVRIEGLTYSLSFSTRKCGQAIGGSIPAFILGLSGYIANQVQTPEVIMGI RTSIA
 LVPCGFMLLAFVLIWFYPLTDKKFKEIVVEIDNRKKVQQQLISDITN

Atgaataccaggtataattccaggtatatattttcgattaccttagtcgctacattaggt
 ggtttattattttggctacgacacccggttattttcgggtactgttgagtcactcaatacc
 gtctttgttgctccacaaaacttaagtgaatccgctgccaactccctgttaggggttttgc
 gtggccagcgtctgattgggtgcatcatcggcggtgccctcggtgggttattgcagtaac
 cgcttcggtcgctggtgattcacttaagattgctgctgtcctgttttttattttctgggtga
 ggttctgcctggccagaacttggttttacctctataaaccggacaacactgtgcctgtt
 tatctggcagggttatgtcccggaatttggtatttatcgcatattggcggtattggcggtt
 ggttttagcctcaatgctctcgccaatgtatattgcggaactggctccagctcatattcgc
 gggaaactgggtctcttttaaccagtttgcgattattttcgggcaacttttagtttactgc
 gtaaaactattttattgcccgttcgggtgatgccagctggctgaatactgacggctggcggt
 tatatgtttgctcggaatgtatccctgcactgctgttcttaatgctgctgtataccgtg
 ccagaaagtcctcgctggctgatgtcgcgcggcaagcaagaacaggcggaaggatatcctg
 cgcaaaattatgtcattgaatatcagctctgcaaaccctgaaacggaatcgcccgtgtttt

atggttgatgcacggtgagctgtgtgtgctgatttcgacctttgcgggtcagcgcctcgctg
ttgttctacgtgagctatgtgttaaatagataccgggctgttcactgtgctgggtactgggtg
caaaacctggttggtactgtggcatcgacacgctggtgacggggatgggtcgcgaggatc
ggtaaaaagaataaccttcctgattggcgctttgctgggaacctgcggttatctgctgttc
ttctgggtttccgtctggtcactgccgggtggcggttggttgcggtggccatcgcttcaatt
ggtcagggcggttacctgacgctgatgtggcgctggaagctgataccgtagaatacgggt
gaataacctgacggcgctgcaattgaagggctcacctattcactattctcatttaccctg
aatgcgggtcaggcaatcgaggttcaattcctgcctttattttgggggttaagcggatat
atcgccaatcaggtgcaaacgcggaagttattatgggcacatcgacacatcaattgcctta
gtaccttgcggtttatgtctactggcattcggttattatctggttttatccgctcacggat
aaaaaattcaaagaaatcggtggttgaaattgataatcgtaaaaaagtgcagcagcaatta
atcagcgatatcactaattaa

Appendix 8 Factors Considered for Primer Design

Primer length

To promote primer annealing and PCR amplification, primers were designed to be between 18-30 bases. Primers shorter than this range can lead to nonspecific PCR product amplification.

Primer sequence

To avoid nonspecific primer annealing, poly-guanine and poly-cytosine nucleotide bases at the 3'-end of the primer were avoided. The 3'-end of the primer pairs was also assessed for complementarity to avoid the formation of primer-dimers. The primer sequences were assessed to avoid the production of internal secondary structures such as hairpin loops (also known as stem-loops).

Melting temperature

To ensure PCR specificity and successful primer annealing, the melting temperature (T_m) of the primers was considered. The T_m of a primer is the temperature in which half of the primer has annealed to the template DNA. Therefore, the T_m 's of both primers within a pair were designed to be similar (within 5°C). The following formula was used to estimate the T_m and annealing temperature (T_a) for all primers:

$$T_m = 81.5 + 41 \times \text{GC\%} - 675/\text{mers} \\ T_a = T_m - 5^\circ\text{C}$$

GC content

To ensure similar primer T_m 's and discourage the formation of internal secondary structures, the GC content of primers was kept between 40% and 60%, where possible. Due to the high GC content of the *xyIE* sequence, a GC content above 60% was sometimes required within primer design.

Fusion considerations

To increase the probability of successful fusion between multiple genes, several factors were considered for primer design:

1. For the XylE-GFP-GusB fusion, the stop codon of *gfp* (TAA) was replaced with TTA. The amino acid leucine was chosen to replace the stop codon for its codon similarity, side-chain hydrophathy and charge, size, and codon usage within *E. coli* (<http://www.kazusa.or.jp/codon/>). This did not have to be considered when designing primers for the fusion of the two domains of XylE and GusB.
2. All fusion primers were designed to incorporate a shared nucleotide bridge between the binding regions of each gene, to promote fusion. The bridge consisted of 12 nucleotides, and comprised the sequence 5'-GGAGGATCCGCA-3'. This sequence encodes the amino acid residues glycine, glycine, serine and alanine, respectively. Glycine and alanine were selected for their side-chain hydrophathy and charge, size, and codon usage within *E. coli* (<http://www.kazusa.or.jp/codon/>). The central hexanucleotide sequence

(5'-GGATCC-3') was incorporated as it is recognised by the restriction endonuclease *Bam*HI. This recognition sequence was included to enable the recombinant DNA products to be manipulated after transformation, in order to characterise the clones.

3. To ensure that the fusion products encoded one contiguous open reading frame (ORF), the DNA sequences were assessed using the online Expasy Translate tool (<https://web.expasy.org/translate/>).

Cloning considerations

1. For successful incorporation of the fusion fragment into the cloning vector, endonuclease restriction sites were added to the 5'-ends of the flanking primers. The restriction site for *Eco*RI was added to the 5'-end of the forward flanking primer. The restriction site for *Hind*III was added to the 5'-end of the reverse flanking primer. The cloning vector pTTQ18 possesses a multiple cloning site which features both of these restriction sites, only once. DNA cleavage with both *Eco*RI and *Hind*III generates sticky ends. Usage of the same restriction enzymes within the fusion fragment creates compatible sticky ends, permitting ligation.
2. To permit efficient endonuclease digestion at the recognition site, five random nucleotides were added to the 5'-end of the flanking primer molecules. These nucleotides ensure a large enough platform for the restriction endonucleases to bind, allowing for accurate cleavage of the adjacent recognition site. The sequence of nucleotides added was assessed to ensure that it was not palindromic, and that the 5'-ends of the primers could not form primer-dimers.
3. Internal restriction site analyses of the DNA sequence of each fusion fragment was performed using NEBcutter. The fusion fragments did not contain restriction sites for *Eco*RI, *Hind*III or *Bam*HI.

Appendix 9 Primer Designs for Future Experiments

Primer design and primers for Xyle-GFP-GusB chimeras

Xyle forward flanking primer with **EcoR1**

Name	Primer	Annealing temp
XyleF1	5'- ATACTGAATTCGAATGGTCTAAGGCAGGTC TGA -3'	66.3 (22 nt (11GC))

XyleE-GFP reverse fusion primer

Name	sequence	Reverse complement primer	Annealing temp
XyleGFPR1	5'- atgtcgcgcgg c -3' 5'- atgtctaaagg t -3'	5'- ACCTTTAGACATTGCGGAT CCGCCGCCGCGCGACAT -3'	75.6 (24 nt (13GC))
XyleGFPR2	5'- ggccgc aaaac c -3' 5'- atgtctaaagg t -3'	5'- ACCTTTAGACATTGCGGAT CCGCCGGTTTTGCGGCC -3'	73.9 (24 nt (12GC))
XyleGFPR3	5'- gaaggtatcct g -3' 5'- atgtctaaagg t -3'	5'- ACCTTTAGACATTGCGGAT CCGCCCAGGATACCTTC -3'	70.5 (24 nt (10GC))
XyleGFPR4	5'- atgggcaaac g -3' 5'- atgtctaaagg t -3'	5'- ACCTTTAGACATTGCGGAT CCGCCCGTGTGCCCAT -3'	72.2 (24 nt (11GC))
XyleGFPR5	5'- cgcaaaattat g -3' 5'- atgtctaaagg tg -3'	5'- CACCTTTAGACATTGCGGA TCCGCCCATAATTTTGCG -3'	69.3 (25 nt (9GC))

XylE-GFP forward fusion primers

Name	Primer	Annealing temp
XylE-GFPF1	5'-atgtcgcgcggcGGCGGATCCG CAatgtctaaaggt-3'	75.6 (24 nt (13GC))
XylE-GFPF2	5'-ggccgcgcaaaaccGGCGGATCCG CAatgtctaaaggt-3'	73.9 (24 nt (12GC))
XylE-GFPF3	5'-gaaggtatcctgGGCGGATCCG CAatgtctaaaggt-3'	70.5 (24 nt (10GC))
XylE-GFPF4	5'-atgggcaaacacgGGCGGATCCG CAatgtctaaaggt-3'	72.2 (24 nt (11GC))
XylE-GFPF5	5'-gcgcaaaattatgGGCGGATCC GCAatgtctaaaggt-3'	69.3 (25 nt (9GC))

GusB-GFP reverse fusion primers

Name	sequence	Reverse complement primer	Annealing temp
GusB-GF PR1	5'-ttgtacaaatTa- 3' 5'-aatgtggtacgta tc-3'	5'-GATACGTACCACATT TGCGGATCCGCCTAATTT GTACAA-3'	68.6 (27 nt (8GC))
GusB-GF PR2	5'-ttgtacaaatTa- 3' 5'-accctgaaacgga -3'	5'-TCCGTTTCAGGGTTG CGGATCCGCCTAATTTGT ACAA-3'	69.3 (25 nt (9GC))
GusB-GF PR3	5'-ttgtacaaatTa- 3' 5'-gttgcgagccg- 3'	5'-CGGCTGCGCAACTGC GGATCCGCCTAATTTGTA CAA-3'	72.2 (24 nt (11GC))
GusB-GF PR4	5'-gaattgtacaaat Ta-3' 5'-atcagtctgcaaa c-3'	5'-GTTTGCAGACTGATT GCGGATCCGCCTAATTTG TACAATTC-3'	70.9 (29 nt (9GC))

GusB-GF PR5	5'-gaattgtacaaat Ta-3' 5'-tcattgaatatca gtc-3'	5'-GACTGATATTCAATG ATGCGGATCCGCCTAATT TGTACAATTC-3'	70.3 (31 nt (8GC))
----------------	---	---	--------------------------

GusB-GFP forward fusion primer

Name	Primer	Annealing temp
GusB-GFPF1	5'-ttgtacaaatTaGGCGGATCCGCAaatg tggtacgtatc-3'	68.6 (27 nt (8GC))
GusB-GFPF2	5'-ttgtacaaatTaGGCGGATCCGCAaccc tgaaacgga-3'	69.3 (25 nt (9GC))
GusB-GFPF3	5'-ttgtacaaatTaGGCGGATCCGCAgttg cgcagccg-3'	72.2 (24 nt (11GC))
GusB-GFPF4	5'-gaattgtacaaatTaGGCGGATCCGCAat cagtctgcaaac-3'	70.9 (29 nt (9GC))
GusB-GFPF5	5'-gaattgtacaaatTaGGCGGATCCGCAt cattgaatatcagtc-3'	70.3 (31 nt (8GC))

GusB reverse flanking primer with HindIII

Name	sequence	Reverse complement primer	Annealing temp
GusBR 1	5'-GCAATTAA TCAGCGATATC ACTAATTAA-3' ,	5'-ATACTAAGCTTTTAATTAGTGAT ATCGCTGATTAATTGC-3'	69.1 (28 nt (8GC))

XyleE (1st half) GusB (2nd half) (chosen helice to cut: ¼ GusB helice)

1. Before IC2 and IC3

231:MNTQYNSSYIFSITLVATLGLLFGYDTAVISGTVESLNTVFVAPQNLSESAANSLLGFCV
ASALIGCIIGGALGGYCSNRFGRDLSLKIAAVLFFISGVGSAWPELGFTSINPDNTVPVYLAGYV

PE FVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGK LVSNQFAIFGQLLVYCVNYF IARSGDA
 SWLNTDGWR YMFASECIPALLFLMLLYT VPESPRWLMSRG
 MSKGEELFTGVVPILVELDGDVNGHKFSVSSEGEEDATYGKLT LKFICTTGKLPVPWPTLVTTFG
 YGVQC FARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDF
 KEDGNILGHKLEYNNSHNVIYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLL
 PDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYKL
 NVVRIVAQP SLNISLQTLKRNRP LFMLCIGALCVLISTFAVSASSLFYVRYVLNDTG LFTVLVLV
 QNLVGTVASAPLVPGMVARIGKKN TFLIGALLGTCGYLLFFWVS VWSLPVALVALAIASIGQGV
 MTVMWALEADTVEYGEYLTGVRIEGLTYSLSFSTRK CGQAIGGSIPAFILGLSGYIANQVQTPEV
 IMGIRTSIALVPCGFMLLAFVIIWFYPL TDKKFKEIVVEIDNRKKVQQQLISDITN

atgaataccaggtataattccagttatatatattttcgattaccttagtcgctacattaggt
 ggttttattattttggctacgacaccgcggttatttccggtactgttgagtcactcaatacc
 gtctttgttgctccacaaaacttaagtgaatccgctgccaactccctgtaggggttttgc
 gtggccagcgctctgattgggtgcatcatcggcgggtgcctcggtgggttattgcagtaac
 cgcttcggtcgctcgatgattcacttaagattgctgctgctcctgttttttatttctgggtga
 gggttctgcctggccagaacttggttttacctctataaaaccggacacactgtgcctgtt
 tatctggcaggttatgtcccggaatttggtatttatcgcattattggcggtattggcggtt
 gggttagcctcaatgctctcgccaatgtatattgcggaactggctccagctcatattcgc
 gggaaactgggtctcttttaaccagtttgcgattattttcgggcaacttttagtttactgc
 gtaaaactattttatttgcccggttcgggtgatgccagctgggtgaatactgacggctggcg
 tataatgtttgctcggaatgtatccctgcactgctgttcttaatgctgctgtataccgtg
 ccagaaagtccctcgctgggtgatgtcgcgcggtatgtctaaaggtgaagaattattcactgg
 tggtgtcccaattttgggtgaattagatgggtgatgtaaatgggcacaaattttctgtctccggtg
 aaggtgaaggtgatgctacttacggtaaattgacctaaaattttatttgtactactggtaaattg
 ccagttccatggccaaccttagtcactactttcggttatgggtgttcaatgttttgctagataccc
 agatcatatgaaacaacatgactttttcaagtctgccatgccagaaggttatgttcaagaagaa
 ctatttttttcaaagatgacggtaactacaagaccagagctgaagtcaagtttgaaggtgatacc
 ttagttaatagaatcgaattaaaagggtattgatttttaagaagatggtaacatttttaggtcaca
 attggaatacaactataactctcacaatgtttacatcatgggtgacaaacaaaagaatgggtatca
 aagtttaacttcaaaatttagacacacattgaagatgggttctgttcaattagctgaccattatcaa
 caaaatactccaattgggtgatgggtccagtcctgttaccagacacaccattacttatccactcaatc
 tgccttatccaaagatccaaacgaaaagagagaccacatgggtcttgtttagaattttgttactgctg
 ctgggtattaccatgggtatggatgaattgtacaaatTaaatgtgggtacgtatcggttgcgcagc
 cgtcattgaatatcagtcctgcaaaccctgaaacggaatcgcccgctgtttatgttgtgca
 tcggtgcgctgtgtgtgctgatttcgacctttgcggtcagcgctcctcgctgttgttctacg
 tgcgctatgtgttaaataatgataccgggctgttcaactgtgctgggtactgggtgcaaaacctgg
 ttgggtactgtggcatcggcaccgctgggtgccggggatgggtcgcgaggatcggtaaaaaga
 ataccttcctgattggcgctttgctgggaacctgcgggttatctgctgttcttctgggttt
 ccgtctgggtcactgccggtggcggttggttgcggttgccatcgcttcaattgggtcagggcg
 ttaccatgaccgtgatgtggcgctggaagctgataccgtagaatacgggtgaataacctga
 ccggcggtgcgaattgaagggtcacctattcactattctcatttaccgtaaatgcgggtc
 aggcaatcggaggttcaattcctgcctttattttgggggttaagcggatatatcgccaatc
 aggtgcaaacgccggaagttattatgggcatccgcacatcaattgccttagtaccttgcg
 gatttatgctactggcattcggttattatctgggttttatccgctcacggataaaaaattca
 aagaaatcggtggttgaattgataatcgtaaaaaagtgacgacgcaattaatcagcgata
 tcactaattaa

XyleE (1st half) GusB (2nd half) (chosen helice to cut: $\frac{3}{4}$ of the helice of GusB)

2. After IC2 and IC3

266:MNTQYNSSYIFSITLVATLGGLLFGYDTAVISGTVESINTVFVAPQNLSESAANSLLGFCV
 ASALIGCIIGGALGGYCSNRFGRRDSLKIAAVLFFISGVGSAPWPELGFTSINPDNTVPVYLAGYV
 PEFVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGKIVSFNQFAIIFGQLLVYCVNYFIARSGDA
 SWLNTDGWRVMFASECIPALLFLMLLYTVPESPRWLMSRGKQEQAEGLRKIMGNTLATQAVQEI
 KHSLDHGRKT
 MSKGEELFTGVVPILVELDGDVNGHKFSVSSEGEEDATYGKLTCLKFICTTGKLPVPWPPTLVTTFG
 YGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDF
 KEDGNILGHKLEYNNSHNVIYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLL
 PDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYKI
 TLKRNRPLFMLCIGALCVLISTFAVSASSLFYVRYVLNDTGLFTVLVLVQNLVGTVASAPLVPGM
 VARIGKKN TFLIGALLGTCGYLLFFWVSVW SLPVALVALAIASIGQGVTMTVMWALEADTVEYGE
 YLTGVRIEGLTYSLSFSTRKCGQAIGGSIPAFILGLSGYIANQVQTPEVIMGIRTSIALVPCGFM
 LLAFVIWIFYPLTDKKFKEIWEIDNRKKVQQQLISDITN

atgaataccaggtataattccagttatatatattttcgattaccttagtcgctacattaggt
 ggtttattattttggctacgacaccgcggttatttccggtagctgttgagtcactcaatacc
 gtctttgttgctccacaaaacttaagtgaatccgctgccaactccctgtaggggttttgc
 gtggccagcgctctgattgggtgcatcatcgggcggtgcccctcggtgggttattgcagtaac
 cgcttcggtcgctcgtagttcacttaagattgctgctgctcctgttttttatttctgggtga
 gggtctgcctggccagaacttggttttacctctataaaccgggacaacactgtgcctgtt
 tatctggcaggttatgtcccgaatttggtattttatcgcattattggcggtattggcggtt
 gggttagcctcaatgctctcgccaatgtatattgcggaactggctccagctcatattcgc
 gggaaactgggtctcttttaaccagtttgcgattattttcgggcaacttttagtttactgc
 gtaaaactattttattgcccgttccgggtgatgccagctgggtgaatactgacggctggcggt
 tataatgtttgctcggaatgtatccctgcaactgctgttcttaatgctgctgtataccgtg
 ccagaaaagtcctcgctgggtgatgtcgcgcggaagcaagaacaggcggaaggtatcctg
 cgcaaaaattatgggcaaacacgcttgcaactcaggcagtaacaggaaattaaacactccctg
 gatcatggccgcaaaaaccatgtctaaaggtgaagaattattcactgggtgttgctccaattttg
 gttgaattagatgggtgatgttaatgggtcacaaattttctgtctccgggtgaaggtgaaggtgatgc
 tacttacggtaaatgaccttaaaattttatttgtactactggtaaatggcagttccatggccaa
 ccttagtcactactttcggttatgggtgttcaatgttttgctagataccagatcatatgaacaa
 catgactttttcaagctctgccatgccagaaggttatgttcaagaaagaactatttttttcaaaga
 tgacggtaactacaagaccagagctgaagtcaagtttgaaggtgataccttagttaatagaatcg
 aattaaaaggtattgatttttaagaagatggtaacatttttaggtcacaaattggaatacaactat
 aactctcacaaatgtttacatcatgggtgacaaacaaaagaatgggtatcaaagttaacttcaaat
 tagacacaacattgaagatgggtctgttcaattagctgaccattatcaacaaaatactccaattg
 gtgatgggtccagctctgttaccagacaaccattacttatccactcaatctgccttatccaaagat
 ccaaacgaaaagagagaccacatgggtctgtttagaattttgttactgctgctgggtattaccatgg
 tatggatgaattgtacaaatTaaccctgaaacggaatcgcccgctgtttatgttggtgcatcg
 gtgcgctgtgtgtgctgatttcgacctttgcggtcagcgccctcgctggtgttctacgtgc
 gctatgtgttaaatgataccgggctgttccactgtgctggtaggtgcaaaacctgggttg
 gtactgtggcatcggcaccgctgggtgccggggatgggtcgcgaggatcggtaaaaagaata
 ccttctctgattggcgcttttgcgtgggaacctgcgggttatctgctgttcttctgggtttccg
 tctggtcactgccgggtggcggttggttgcggttgcccatcgcttcaattgggtcagggcggtta

ccatgaccgtgatgtggcgctggaagctgataccgtagaatacgggtgaatacctgaccg
gcggtgcgaattgaagggctcacctattcactattctcatttaccggtaaatgcggtcagg
caatcggaggttcaattcctgcctttattttgggggttaagcggatataatcgccaatcagg
tgcaaacgccggaagttattatgggcatccgcacatcaattgccttagtaccttgccgat
ttatgctactggcattcggttattatctggttttatccgctcacggataaaaaaattcaaag
aaatcgtggttgaaattgataatcgtaaaaaagtgacgagcaattaatcagcgatatca
ctaattaa

740

XyleE (1st half) GusB (2nd half) (chosen helice to cut: To the left of the bundle in GusB)

3. Half way through IC2 before IC3

240:MNTQYNSSYIFSITLVATLGGLLFGYDTAVISGTVESLNTVFVAPQNLSESAANSLLGFCV
ASALIGCIIGGALGGYCSNRFGRRDSLKIAAVLFFISGVGSAPWELGFTSINPDNTVPVYLAGYV
PEFVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGKI VSFNQFAIIFGQLLVYCVNYFI ARSGDA
SWLNTDGWR YMFASECIPALLFLMLLYTVPESPRWLMSRGKQEQAEIGIL
MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWPPTLVTTFG
YGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDF
KEDGNILGHKLEYNNSHNVIYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQONTPIGDGPVLL
PDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYKL
VAQPSLNISLQTLKRNRPLFMLCIGALCVLISTFAVSASSLFYVRYVLNDTG LFTVLVLVQNLVG
TVASAPLVPGMVARIGKKN TFLIGALLGTCGYLLFFWVSVW SLPVALVALAIASIGQGVMTVMW
ALEADTVEYGEYLTGVRIEGLTYSLSFSTRK CGQAIGGSI PAFILGLSGYIANQVQTPEVIMGIR
TSIALVPCGFMLLAFVIIWFYPL TDKKFKIEIVVEIDNRKKVQQQLISDITN

atgaataccaggtataattccagttatataatctcgattaccttagtcgctacattaggt
ggtttattatttggctacgacaccgccgttatttccggtagctgttgagtcactcaatacc
gtctttgttgctccacaaaacttaagtgaatccgctgccaaactccctgttagggttttgc
gtggccagcgctctgattggttgcatcatcggcgggtgccctcggtggttattgcagtaac
cgcttcggtcgctcgtagttcacttaagattgctgctgtcctgttttttatttctggtgta
ggttctgcctggccagaacttggttttacctctataaaccgggacaacactgtgcctgtt
tatctggcaggttatgtcccggaatttgttattttatcgcatatttgccggtattggcggt
ggtttagcctcaatgctctcgccaatgtatattgcggaactggctccagctcatattcgc
gggaaactgggtctcttttaaccagtttgcgattatttccgggcaacttttagtttactgc
gtaaaactattttattgcccgttccgggtgatgccagctggctgaatactgacggctggcgt
tatatgtttgctcggaatgtatccctgcactgctgttcttaatgctgctgtataccggtg
ccagaaagtcctcgctggctgatgtcgcgcggaagcaagaacaggcggaaggtatcctga
tgtctaaaggtgaagaattattcactgggtgtgtcccaattttggttgaattagatgggtgatgtt
aatggtcacaaattttctgtctccgggtgaaggtgaaggtgatgctacttacggtaaattgacctt
aaaattttatttgtactactggtaaattgccagttccatggccaaccttagtcactactttcggtt
atgggtgttcaatgttttgcctagataccagatcatatgaaacaacatgactttttcaagtctgcc
atgccagaaggttatgttcaagaaagaactatttttttcaaagatgacggtaactacaagaccag
agctgaagtcaagtttgaaggtgataccttagttaatagaatcgaattaaaaggtattgatttta
agaagatggtaacatttttaggtcacaaattggaatacaactataactctcacaatgtttacatc
atggctgacaaacaaaagaatgggtatcaaagttaacttcaaaattagacacaacattgaagatgg
ttctgttcaattagctgaccattatcaacaaaatactccaattggtgatgggtccagtcctgttac

cagacaaccattacttatccactcaatctgccttatccaaagatccaaacgaaaagagagaccac
atgggtcttggttagaatttggttactgctgctggtattacccatgggatggaattgtacaaatT
agttgcgcgagccgctcattgaatatcagtcctgcaaaccctgaaacggaatcgcccgctgtt
atggttgatgcacggtgcgctgtgtgtgctgatttcgacctttgcggtcagcgccctcgctcg
ttgttctacgtgcgctatgtgttaaatagataccgggctgttcactgtgctggtactggtg
caaaacctggttggtactgtggcatcggcaccgctggtgcccgggatggtcgcgaggatc
ggtaaaaagaataaccttcctgattggcgctttgctgggaacctgcggttatctgctgttc
ttctgggtttccgctctggtcactgccggtggcgcttggttgcgcttgcccatcgcttcaatt
ggtcagggcggttaccatgaccgtgatgtggcgctggaagctgataccgtagaatacgggt
gaataacctgaccggcggtgcgaattgaagggctcacctattcactattctcatttaccggt
aaatgcggtcagggcaatcgaggttcaattcctgcctttattttgggggttaagcggatat
atcgccaatcaggtgcaaacgcggaagttattatgggcacatcgccacatcaattgcctta
gtaccttgcggtatttatgctactggcattcggttattatctggttttatccgctcacggat
aaaaaattcaaagaaatcggtggttgaaattgataatcgtaaaaaagtgcagcagcaatta
atcagcgatatcactaattaa

725

XyleE (1st half) GusB (2nd half) (chosen helice to cut: To the right of the bundle in GusB)

4. After IC2 half way through IC3

247:MNTQYNSSYIFSITLVATLGGLLFgydTAVISGTVESLNTVFVAPQNLSESAANSLLGFCV
ASALIGCIIGGALGGYCSNRFGRDRLKIAAVLFFISGVGSAPWPELGFTSINPDNTVPVYLAGYV
PEFVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGKLVSFNQFAIIFGQLLVYCVNYFIARSGDA
SWLNTDGWRMYFASECIPALLFLMLLYTVPESPRWLMSRGKQEQAEGLRKIMGNT
MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGGDATYGKLTlKFICTTGKLPVPWPTLVTTFG
YGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDF
KEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLL
PDNHYLSTQSALSKDPNEKRDHMLVLEFVTAAGITHGMDELYKL
ISLQTLKRNRP LFMLCIGALCVLISTFAVSASSLFYVRYVLNDTGLFTVLVLVQNLVGTVASAPL
VPGMVARIGKKN TFLIGALLGTCGYLLFFWVSVW SLPVALVALAIASIGQGVMTVMWALEADTV
EYGEYLTGVRIEGLTYSLFSFTRKCGQAIGGSIPAFILGLSGYIANQVQTPVEVIMGI RTSIALVP
CGFMLLAFVIIWFYPLTDKKFKEIVVEIDNRKKVQQQLISDITN

atgaataccaggtataattccagttatatattttcgattaccttagtcgctacattaggt
ggtttattattttggctacgacaccgcggttatttccggtactgttgagtcactcaatacc
gtctttgttgctccacaaaacttaagtgaatccgctgccaactccctggttaggggttttgc
gtggccagcgctctgattggttgcatcatcggcggtgcccctcggtggttattgcagtaac
cgcttcggtcgctgattcacttaagattgctgctgtcctgtttttttatttctggtgta
ggttctgcctggccagaacttggttttacctctataaaaccggacaacactgtgcctgtt
tatctggcaggttatgtcccgaatttggtatttatcgcatattggcggtattggcggtt
ggttttagcctcaatgctctcgccaatgtatattgcggaactggctccagctcatattcgc
gggaaactggtctcttttaaccagtttgcgattattttcgggcaacttttagtttactgc
gtaaaactattttattgcccgttccggtgatgccagctggctgaatactgacggctggcggt
tatatggttgctcggaatgtatccctgcaactgctgttcttaatgctgctgtataccggtg
ccagaaagtcctcgctggctgatgtcgcgcggaagcaagaacaggcggaaggtatcctg
cgcaaaattatgggcaacacgatgtctaaagggtgaagaattattcactgggtggttgcctcaatt

ttggttgaattagatggtgatgttaatggtcacaaatcttctgtctccggtgaaggtgaaggtga
 tgctacttacggtaaattgaccttaaaatcttattgtactactggtaaattgccagttccatggc
 caaccttagtcactactttcgggttatggtgttcaatgttttgctagatacccagatcatatgaaa
 caacatgactttttcaagtctgccatgccagaaggttatgttcaagaaagaactatcttttttcaa
 agatgacggtaactacaagaccagagctgaagtcaagtttgaaggtgataccttagttaatagaa
 tcgaattaaaaggtattgattttaagaagatggtaacatttttaggtcacaaattggaatacaac
 tataactctcacaatgtttacatcatggctgacaaacaaaagaatgggtatcaaagttaacttcaa
 aattagacacacacattgaagatggttctgttcaattagctgaccattatcaacaaaataactcaa
 ttggtgatgggtccagtcctgttaccagacaaccattacttatccactcaatctgccttatccaaa
 gatccaaacgaaaagagagaccacatgggtcttgttagaattttgttactgctgctggtattacca
 tgggtatggatgaattgtacaaatTaatacagtcctgcaaaccctgaaacggaatcgcccgctgt
 ttatgtttgtgcatcggtgcgctgtgtgtgctgattttcgacctttgcggtcagcgccctcgt
 cggtgttctacgtgcgctatgtgttaaatagataccgggctgttccactgtgctggtactgg
 tgcaaaacctggttggtactgtggcatcggcaccgctggtgccggggatgggtcgcgagga
 tcggtaaaaagaataccttccctgattggcgctttgctgggaacctgcggttatctgctgt
 tcttctgggtttccgctctggtcactgccggtggcggttggttgcggtggccatcgcttcaa
 ttggtcagggcggttaccatgaccgtgatgtggcgctggaagctgataccgtagaatacgt
 gtgaatacctgaccggcggtgcgaattgaagggctcacctattcactatttctcatttacc
 gtaaatgcggtcaggcaatcgaggttcaattcctgcctttattttgggggttaagcggat
 atatcgccaatcaggtgcaaacgcggaagtattattatgggcacatccgcacatcaattgcct
 tagtaccttgcggtatttatgctactggcattcggtattatctgggttttatccgctcacgg
 ataaaaaattcaaagaaatcggtggttgaaattgataatcgtaaaaaagtgcagcagcaat
 taatcagcgatatcactaattaa

725

XyleE (1st half) GusB (2nd half) (GusB chosen helice to cut: exactly half way through)

5. Just after IC2. Between IC2 and IC3
 244:MNTQYNSSYIFSITLVATLGGLLFgydTAVISGTVESLNTVFVAPQNLSESAANSLLGFCV
 ASALIGCIIGGALGGYCSNRFGRDRLKIAAVLFFISGVGSAWPELGFTSINPDNTPVPVYLAGYV
 PEFVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGKLVSNQFAIIFGQLLVYCVNYFIARSGDA
 SWLNTDGWRVYMFASECIPALLEMLLYTVPESPRWLMSRGKQEQAEGLRKIM
 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYgKLTlKFICTTGKLPVPWPtLVTTFG
 YGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDF
 KEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQONTPIGDGPVLL
 PDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYKI
 SLNISLQTLKRNRLFMLCIGALCVLISTFAVSASSLFYVRYVLNDTGLFTVLVLVQNLVGTVAS
 APLVPGMVARIGKKNTFLLIGALLGTCGYLLFFWVSVWSLPVALVALAIASIGQGVMTVMWALEA
 DTVEYGEYLTGVRIEGLTYSLSFSTRKCGQAIGGSIPAFILGLSGYIANQVQTPEVIMGIRTSIA
 LVPCGFMLLAFVIIWFYPLTDKKFKEIVVEIDNRKKVQQQLISDITN

atgaataccaggtataattccagttatatattttcgattaccttagtcgctacattaggt
 ggtttattattttggctacgacacgcggttattttcggtactgttgagtcactcaatacc
 gtctttgttgcctccacaaaacttaagtgaatccgctgccaaactccctggttaggggttttgc
 gtggccagcgctctgattgggtgcatcatcggcgggtgccctcggtgggttattgcagtaac
 cgcttcggtcgctcggtgattcacttaagattgctgctgtcctgttttttattttctggtgta

gggtctgcctggccagaacttggttttacctctataaaccgggacaacactgtgcctgtt
 tatctggcagggttatgtcccgaatttggtatttatcgcatattggcggtattggcggtt
 gggttagcctcaatgctctcgccaatgtatattgcggaactggctccagctcatattcgc
 gggaaactgggtctcttttaaccagtttgcgattattttcgggcaacttttagttactgc
 gtaaaactattttattgcccgttcgggtgatgccagctggctgaatactgacggctggcgt
 tataatggttgctcggaatgtatccctgcactgctgttcttaaatgctgctgtataccgtg
 ccagaaagtcctcgctgggtgatgtcgcgcggaagcaagaacaggcggaagggtatcctg
 cgcaaaattatgatgtctaaagggtgaagaattattcactgggtgttgcctcaattttgggtgaat
 tagatgggtgatgttaatggtcacaaattttctgtctccggtgaagggtgaagggtgatgctacttac
 ggtaaatgaccttaaaattttatttgactactggtaaatggcagttccatggccaaccttagt
 cactactttcgggttatgggtgttcaatgttttgctagataccagatcatatgaacaacatgact
 ttttcaagtctgccatgccagaagggttatgttcaagaaagaactattttttcaaagatgacgggt
 aactacaagaccagagctgaagtcaagtttgaagggtgataccttagttaatagaatcgaattaaa
 aggtattgatttttaagaagatggtaacatttttaggtcacaaattggaatacaactataactctc
 acaatgtttacatcatggctgacaaacaaaagaatgggtatcaaaggttaacttcaaaattagacac
 aacattgaagatgggtctgttcaattagctgaccattatcaacaaaatactccaattgggtgatgg
 tccagtcctgttaccagacaaccattacttatccactcaatctgccttatccaaagatccaaacg
 aaaagagagaccacatgggtcttgttagaattttgttactgctgctgggtattacccatgggtatggat
 gaattgtacaaatTatcattgaatatcagtcctgcaaaccctgaaacggaatcgcccgtgt
 ttatgttgtgcatcggtgcgctgtgtgtgctgatttcgacctttgcggtcagcgccctcgt
 cggtgttctacgtgcgctatgtgttaaatacaccgggctgttccactgtgctgggtactgg
 tgcaaaacctgggttggtactgtggcatcggcaccgctgggtgccgggggatgggtcgcgagga
 tcggtaaaaagaataaccttcctgattggcgctttgctgggaacctgcgggttatctgctgt
 tcttctgggtttccgctctgggtcactgccgggtggcggttggttgcgttggccatcgcttcaa
 ttgggtcaggggcggttaccatgacgctgatgtgggcgctggaagctgataccgtagaatacgt
 gtgaatacctgaccggcggtgcgaattgaagggtcacctattcactatttctcatttacc
 gtaaatgcgggtcaggcaatcggaagggttcaattcctgcctttattttgggggttaagcggat
 atatcgccaatcagggtgcaaacgccggaagttattatgggcacatccgcacatcaattgcct
 tagtaccttgccgatttatgctactggcattcggttattatctgggttttatccgctcacgg
 ataaaaaattcaaagaaatcgtggttgaaattgataatcgtaaaaaagtgcagcagcaat
 taatcagcgatatcactaattaa

725

Fusion of N-domain of Xyle with GFP

Xyle forward flanking primer with **EcoRI**:

Name	Primer	Annealing temp
XyleF 1	5'- ATACTGAATTC GAATGGTCTAAGGCAGGTC TGA- 3'	66.3 (22 nt (11GC))

XyleE-GFP reverse fusion primer:

Name	sequence	Reverse complement primer	Annealing temp
XylE-GFP R1	5'-atgtcgcgcg gc-3' 5'-atgtctaaag gt-3'	5'-ACCTTTAGACATTGCGGATC CGCCGCCGCGCGACAT-3'	75.6 (24 nt (13GC))
XylE-GFP R2	5'-ggccgcaaaa cc-3' 5'-atgtctaaag gt-3'	5'-ACCTTTAGACATTGCGGATC CGCCGGTTTTGCGGCC-3'	73.9 (24 nt (12GC))
XylE-GFP R3	5'-gaaggtatcc tg-3' 5'-atgtctaaag gt-3'	5'-ACCTTTAGACATTGCGGATC CGCCAGGATACCTTC-3'	70.5 (24 nt (10GC))
XylE-GFP R4	5'-atgggcaaca cg-3' 5'-atgtctaaag gt-3'	5'-ACCTTTAGACATTGCGGATC CGCCGTGTTGCCCAT-3'	72.2 (24 nt (11GC))
XylE-GFP R5	5'-cgcaaaatta tg-3' 5'-atgtctaaag gtg-3'	5'-CACCTTTAGACATTGCGGAT CCGCCATAATTTTGCG-3'	69.3 (25 nt (9GC))

XylE-GFP forward fusion primers

Name	Primer	Annealing temp
XylE-GFPF1	5'-atgtcgcgcgggcGGCGGATCCGCAa tgtctaaaggt-3'	75.6 (24 nt (13GC))
XylE-GFPF2	5'-ggccgcaaaaaccGGCGGATCCGCAa tgtctaaaggt-3'	73.9 (24 nt (12GC))
XylE-GFPF3	5'-gaaggtatcctgGGCGGATCCGCAa tgtctaaaggt-3'	70.5 (24 nt (10GC))
XylE-GFPF4	5'-atgggcaaacacgGGCGGATCCGCAa tgtctaaaggt-3'	72.2 (24 nt (11GC))
XylE-GFPF5	5'-gcgcaaaattatgGGCGGATCCGCA atgtctaaaggt-3'	69.3 (25 nt (9GC))

GFP reverse flanking primer with HindIII

Name	sequence	Reverse complement primer	Annealing temp
GFPR1	5' - catggta tggatgaatt gtacaaataa -3'	5' - ATACTAAGCTT TTATTGT ACAATTCATCCATACCATG-3'	68.6 (27 nt (8GC))

Xyle (1st half) - GFP

1. Before IC2 and IC3

231:MNTQYNSSYIFSITLVATLGGLLFGYDTAVISGTVESLNTVFVAPQNLSESAANSLLGFCV
 ASALIGCIIGGALGGYCSNRFGRDRLKIAAVLFFISGVGSAWPELGFTSINPDNTVPVYLAGYV
 PEFVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGKLVSENFQFAIFGQLLVYCVNYFIARSGDA
 SWLNTDGWRVYMFASECIPALLFLMLLYTVPESPRWLMSRG
 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWPPTLVTTFG
 YGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEEDTLVNRIELKGIDF
 KEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLL
 PDNHLYSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK

atgaataccagataattccagttatatatatttcgattaccttagtcgctacattaggt
 ggtttattatttggctacgacaccgcggttatttcggtactgttgagtcactcaatacc
 gtctttgttgctccacaaaacttaagtgaatccgctgccaaactccctggttaggggttttgc
 gtggccagcgctctgattggttgcacatcggcgggtgccctcggtggttattgcagtaac
 cgcttcggtcgctcgatcacttaagattgctgctgctcctgttttttatttctggtgta
 ggttctgcctggccagaacttggttttacctctataaaacccggacaacactgtgcctgtt
 tatctggcagggttatgtcccggaatttggtatttatcgcatatttggcgggtattggcggtt
 ggttttagcctcaatgctctcgccaatgtatattgcggaactggctccagctcatattcgc
 gggaaactggtctcttttaaccagtttgcgattatttccgggcaacttttagtttactgc
 gtaaaactattttattgcccgttcgggtgatgccagctggctgaatactgacggctggcgt
 tataatggttgctcggaatgtatccctgcactgctgttcttaatgctgctgtataccgtg
 ccagaaaagtcctcgctggctgatgtcgcgcggtcatgtctaaaggtgaagaattattcactgg
 tgttggtcccaattttggttgaattagatggtgatgttaatggtcacaaattttctgtctccggtg
 aaggtgaaggtgatgctacttacggtaaaattgacctaaaatttatttgtactactggtaaatg
 ccagttccatggccaaccttagtcactactttcggttatggtggttcaatgttttgctagatacc
 agatcatatgaacaacatgactttttcaagtctgccatgccagaagggttatgttcaagaaagaa
 ctatttttttcaaagatgacggtaactacaagaccagagctgaagtcaagtttgaaggtgatacc
 ttagttaatagaatcgaattaaaagggtattgatttttaagaagatggtaacatttttaggtcaca
 attggaatacaactataactctcacaatgtttacatcatggctgacaaacaaaagaatggtatca
 aagttaacttcaaaattagacacaacattgaagatggttctgttcaattagctgaccattatcaa
 caaaatactccaattggtgatggtccagtcctgttaccagacaaccattacttatccactcaatc
 tgccttatccaaagatccaaacgaaaagagagaccacatgggtcttgttagaattttgttactgctg
 ctggtattaccatgggtatggatgaattgtacaaataa

XyleE (1st half) - GFP

2. After IC2 and IC3

266:MNTQYNSSYIFSITLVATLGGLLFGYDTAVISGTVESLNTVFVAPQNLSESAANSLLGFCV
ASALIGCIIGGALGGYCSNRFGRDRSLKIAAVLFFISGVGSAWPELGFTSINPDNTVPVYLAGYV
PEFVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGKLVSNQFAIIFGQLLVYCVNYFIARSGDA
SWLNTDGWRVMFASECIPALLFLMLLYTVPESPRWLMSRGKQEQAEGLRKIMGNTLATQAVQEI
KHSLDHGRKT
MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWPPTLVTTFG
YGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDF
KEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQONTPIGDGPVLL
PDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYK

atgaataccagataattccagttatatatatttcgattaccttagtcgctacattaggt
ggtttattatttggctacgacaccgcggttatttcggtactggtgagtcactcaatacc
gtctttggtgctccacaaaacttaagtgaatccgctgccaactccctggttaggggttttgc
gtggccagcgctctgattggttgcatcatcggcggtgccctcggtggttattgcagtaac
cgcttcggtcgctcgatcacttaagattgctgctgctcctgttttttatttctggtgta
ggttctgcctggccagaacttggttttacctctataaaccggacaacactgtgcctggt
tatctggcaggttatgtcccggaatttggtatttatcgcatatttggcggtattggcggt
ggttttagcctcaatgctctcgccaatgtatattgcggaactggctccagctcatattgc
gggaaactggtctcttttaaccagtttgcgattatttccggcaacttttagtttactgc
gtaaacatttttattgcccgttcgggtgatgccagctggctgaatactgacggctggcgt
tatatgtttgccctcggaatgtatccctgcactgctggttcttaatgctgctgtataccgtg
ccagaaagtcctcgctggctgatgtcgcgcggaagcaagaacaggcggaaggtatcctg
cgcaaaattatgggcaacacgcttgcaactcaggcagtacaggaaattaaacactccctg
gatcatggccgcaaaaccatgtctaaaggtgaagaattattcactgggtgttgcctaattttg
gttgaattagatggtgatgttaatggtcacaaattttctgtctccggtgaaggtgaaggtgatgc
tacttacggtaaatgaccttaaaattttatttgtactactggtaaatgccaagttccatggcaa
ccttagtcactactttcggttatggtgttcaatgttttgctagataccagatcatatgaacaa
catgactttttcaagtctgccatgccagaaggttatgttcaagaaagaactattttttcaaaga
tgacggtaactacaagaccagagctgaagtcaagtttgaaaggtgataccttagttaatagaatcg
aattaaaaggtattgatttttaagaagatggtaacatttttaggtcacaaattggaatacaactat
aactctcacaaatgtttacatcatggctgacaaacaaaagaatggatatcaagtttaacttcaaat
tagacacaacattgaagatggttctgttcaattagctgaccattatcaacaaaatactccaattg
gtgatggtccagtcctgttaccagacaaccattacttatccactcaatctgccttatccaaagat
ccaaacgaaaagagagaccacatggtcctgttagaatttggttactgctgctggtattaccatgg
tatggatgaattgtacaaataa

504

XyleE (1st half) - GFP

3. Half way through IC2 before IC3

240:MNTQYNSSYIFSITLVATLGGLLFGYDTAVISGTVESLNTVFVAPQNLSESAANSLLGFCV
ASALIGCIIGGALGGYCSNRFGRDRSLKIAAVLFFISGVGSAWPELGFTSINPDNTVPVYLAGYV
PEFVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGKLVSNQFAIIFGQLLVYCVNYFIARSGDA
SWLNTDGWRVMFASECIPALLFLMLLYTVPESPRWLMSRGKQEQAEGL

MSKGEELFTGVVPILVELDGDVNGHKFSVSSEGEEDATYGKLTCLKFICTTGKLPVPWPTLVTTFG
YGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEEDTLVNRIELKGIDF
KEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLL
PDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYK

atgaataccaggtataattccagttatatatattttcgattaccttagtcgctacattaggt
ggttttattattttggctacgacaccgcggttattttccggtactgttgagtcactcaataacc
gtctttgttgctccacaaaacttaagtgaatccgctgccaactccctggttaggggttttgc
gtggccagcgctctgattgggtgcatcatcggcggtgccctcggtgggttattgcagtaac
cgcttcggtcgctcggtgattcacttaagattgctgctgctcctgttttttatttctgggtga
ggttctgcctggccagaacttggttttacctctataaaccgggacaacactgtgcctgtt
tatctggcaggttatgtcccgaatttggtattttatcgcatatttggcggtattggcggt
ggtttagcctcaatgctctcgccaatgtatattgcggaactggctccagctcatattcgc
gggaaactgggtctcttttaaccagtttgcgattattttcgggcaacttttagtttactgc
gtaaactattttattgcccgttcgggtgatgccagctggctgaatactgacggctggcggt
tatatgtttgcctcggaatgtatccctgcactgctgttcttaatgctgctgtataccgtg
ccagaaagtcctcgctgggtgatgtcgcgcggaagcaagaacaggcggaagggtatcctga
tgtctaaagggtgaagaattattcactgggtgtgtcccaattttggttgaattagatgggtgatgtt
aatgggtcaciaaattttctgtctccggtgaagggtgaagggtgatgctacttacggtaaatgacctt
aaaattttatttgtactactggtaaatggcagttccatggccaacttagtcactactttcggtt
atgggtgttcaatgttttgcctagataaccagatcatatgaacaacatgactttttcaagtctgcc
atgccagaagggttatgttcaagaagaactatttttttcaaagatgacggtaactacaagaccag
agctgaagtcaagtttgaagggtgataccttagttaatagaatcgaattaaaagggtattgatttta
aagaagatggtaacatttttaggtcacaaattggaatacaactataactctcacaatgtttacatc
atgggtgacaaaacaaaagaatgggtatcaaagttaacttcaaaattagacacaacattgaagatgg
ttctgttcaattagctgaccattatcaacaaaatactccaattgggtgatgggtccagctctgtttac
cagacaaccattacttatccactcaatctgccttatccaaagatccaaacgaaaagagagaccac
atgggtcttgtttagaattttgttactgctgctgggtattacccatgggtatgggtgaattgtacaaata
a

478

XyleE (1st half) - GFP

4. After IC2 half way through IC3

247:MNTQYNSSYIFSITLVATLGGLLFGYDTAVISGTVESLNTVFVAPQNLSESAANSLLGFCV
ASALIGCIIGGALGGYCSNRFGRDRLKIAAVLFFISGVGSAPWPELGFTSINPDNTVPVYLAGYV
PEFVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGKLVSFNQFAIIFGQLLVYCVNYFIARSGDA
SWLNTDGWRVYMFASECIPALLEFLMLLYTVPEsprwLMSRGKQEQAEgILRKIMGNT
MSKGEELFTGVVPILVELDGDVNGHKFSVSSEGEEDATYGKLTCLKFICTTGKLPVPWPTLVTTFG
YGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEEDTLVNRIELKGIDF
KEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLL
PDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYK

atgaataccaggtataattccagttatatatattttcgattaccttagtcgctacattaggt
ggttttattattttggctacgacaccgcggttattttccggtactgttgagtcactcaataacc
gtctttgttgctccacaaaacttaagtgaatccgctgccaactccctggttaggggttttgc
gtggccagcgctctgattgggtgcatcatcggcggtgccctcggtgggttattgcagtaac

cgcttcggtcgtcgtgattcacttaagattgctgctgtcctgttttttatttctggtgta
 ggttctgcctggccagaacttggttttacctctataaaccggacaacactgtgcctgtt
 tatctggcaggttatgtcccgaatttggtatttatcgcattattggcgggtattggcgtt
 ggttttagcctcaatgctctcgccaatgtatattgcggaactgggtccagctcatattcgc
 gggaaactgggtctcttttaaccagtttgcgattattttcgggcaacttttagttactgc
 gtaaaactattttattgcccgttcgggtgatgccagctgggtgaatactgacggctggcgt
 tataatgtttgctcggaatgtatccctgcactgctgttcttaatgctgctgtataccgtg
 ccagaaagtcctcgtggtgatgtcgcgcggcaagcaagaacaggcggaagggtatccg
 cgcaaaattatgggcaacacgatgtctaaagggtgaagaattattcactgggtgttgcctcaatt
 ttggttgaattagatgggtgatgttaatgggtcacaattttctgtctccgggtgaagggtga
 tgctacttacggtaaattgaccttaaaatttatttgtactactggtaaattgccagttccatggc
 caaccttagtcactacttttcgggttatgggtgttcaatgttttgctagataccagatcatatgaaa
 caacatgactttttcaagtctgccatgccagaagggttatgttcaagaaagaactatttttttcaa
 agatgacggtaactacaagaccagagctgaagtcgaagtttgaagggtgataccttagttaatagaa
 tcgaattaaaagggtattgattttaagaagatggtaacatttttaggtcacaatttgaatacaac
 tataactctcacaatgtttacatcatggctgacaaacaaaagaatggtatcaaagttaacttcaa
 aattagacacacattgaagatgggtctgttcaattagctgaccattatcaacaaaatactccaa
 ttgggtgatgggtccagtccttgttaccagacacaccattacttatccactcaatctgccttatccaaa
 gatccaaacgaaaagagagaccacatgggtcttgttagaatttgttactgctgctggtattacca
 tgggtatggatgaattgtacaaataa

485

XyleE (1st half) - GFP

5. Just after IC2. Between IC2 and IC3

244:MNTQYNSSYIFSITLVATLGGLLFGYDTAVISGTVESLNTVFVAPQNLSESAANSLLGFCV
 ASALIGCIIGGALGGYCSNRFGRDLSLIAAVLFFISGVGSAPPELGFTSINPDNTVPVYLAGYV
 PEFVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGKLVSFNQFAIIFGQLLVYCVNYFIARSGDA
 SWLNTDGWRYMFASECIPALLFLMLLYTVPEsprwLMSRGKQEQAEgILRKIM
 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYgKLTlKFICTTGKLPVPWPtLVTTfG
 YGVQCfARYPDHMKQHDFfKSAMPEgYVQERTIFFKDDGNYKTRAEVKfEGDtlVNRIELKGIDf
 KEDGNILGHKLEYNynSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQONTPIGDGPVLL
 PDNHylSTQSALSKDPNEKRDHmVLLefVTAAGIThGMDELYK

atgaataccagttataattccagttatatattttcgattaccttagtcgctacattaggt
 ggtttattattttggctacgacaccgcggttattttccgggtactgttgagtcactcaatacc
 gtctttgttgctccacaaaacttaagtgaatccgctgccaactccctgttagggttttgc
 gtggccagcgtctgattgggtgcatcatcggcgggtgccctcggtgggttattgcagtaac
 cgcttcggtcgtcgtgattcacttaagattgctgctgtcctgttttttatttctggtgta
 ggttctgcctggccagaacttggttttacctctataaaccggacaacactgtgcctgtt
 tatctggcaggttatgtcccgaatttggtatttatcgcattattggcgggtattggcgtt
 ggttttagcctcaatgctctcgccaatgtatattgcggaactgggtccagctcatattcgc
 gggaaactgggtctcttttaaccagtttgcgattattttcgggcaacttttagttactgc
 gtaaaactattttattgcccgttcgggtgatgccagctgggtgaatactgacggctggcgt
 tataatgtttgctcggaatgtatccctgcactgctgttcttaatgctgctgtataccgtg
 ccagaaagtcctcgtggtgatgtcgcgcggcaagcaagaacaggcggaagggtatccg
 cgcaaaattatgatgtctaaagggtgaagaattattcactgggtgttgcctcaattttggttgaat
 tagatgggtgatgttaatgggtcacaattttctgtctccgggtgaagggtgaagggtgatgctacttac

ggtaaattgaccttaaaattttatgtgtactactggtaaattgccagttccatggccaaccttagt
 cactacttttcggttatggtgttcaatgttttgctagataccagatcatatgaaacaacatgact
 ttttcaagtctgccatgccagaagggttatgttcaagaaagaactatTTTTTCAAAGATGACGGT
 aactacaagaccagagctgaagtcaagtttgaagggtgataccttagttaatagaatcgaattaaa
 aggtattgatttttaaagaagatggtaacatttttaggtcacaaattggaatacaactataactctc
 acaatgtttacatcatggctgacaaacaaaagaatggtatcaaagttaacttcaaaaattagacac
 aacattgaagatgggttctgttcaattagctgaccattatcaacaaaatactccaattgggtgatgg
 tccagtccttgttaccagacaaccattacttatccactcaatctgccttatccaaagatccaaacg
 aaaagagagaccacatgggtcttggttagaatttggttactgctgctggtattaccatgggtatggat
 gaattgtacaaataa

482

Fusion of GFP with C-domain of GusB

GFP forward flanking primer with **EcoR1**

Name	sequence	Annealing temp
GFPF1	5'- ATACTGAATTC atgtctaaagg tgaagaattattcactgg-3'	72.4 (29 nt (10GC))

GusB-GFP reverse fusion primer

Name	sequence	Reverse complement primer	Annealing temp
GusB GFPR1	5'- ttgtacaaatTa -3' ' 5'- aatgtggtacgtat c -3'	5'- GATACGTACCACATTTGC GATCCGCCTAATTTGTACAA -3' '	68.6 (27 nt (8GC))
GusB GFPR2	5'- ttgtacaaatTa -3' ' 5'- accctgaaacgga - 3'	5'- TCCGTTTCAGGGTTGCGGA TCCGCCTAATTTGTACAA -3'	69.3 (25 nt (9GC))
GusB GFPR3	5'- ttgtacaaatTa -3' ' 5'- gttgcgagccg -3' '	5'- CGGCTGCGCAACTGCGGAT CCGCCTAATTTGTACAA -3'	72.2 (24 nt (11GC))

GusB GFPR4	5'-gaattgtacaaatT a-3' 5'-atcagtctgcaaac -3'	5'-GTTTGCAGACTGATTGCGG ATCCGCCTAATTTGTACAATT -3'	70.9 (29 nt (9GC))
GusB GFPR5	5'-gaattgtacaaatT a-3' 5'-tcattgaatatcag tc-3'	5'-GACTGATATTCAATGATGC GGATCCGCCTAATTTGTACAAT TC-3'	70.3 (31 nt (8GC))

GusB-GFP forward fusion primer

Name	Primer	Annealing temp
GusB-GFPF1	5'-ttgtacaaatTaGGCGGATCCGCAaa tgtggtacgtatc-3'	68.6 (27 nt (8GC))
GusB-GFPF2	5'-ttgtacaaatTaGGCGGATCCGCAac cctgaaacgga-3'	69.3 (25 nt (9GC))
GusB-GFPF3	5'-ttgtacaaatTaGGCGGATCCGCAgt tgcgagccg-3'	72.2 (24 nt (11GC))
GusB-GFPF4	5'-gaattgtacaaatTaGGCGGATCCGCA atcagtctgcaaac-3'	70.9 (29 nt (9GC))
GusB-GFPF5	5'-gaattgtacaaatTaGGCGGATCCGC Atcattgaatatcagtc-3'	70.3 (31 nt (8GC))

GusB reverse flanking primer with HindIII

Name	sequence	Reverse complement primer	Annealing temp
GusBR 1	5'-GCAATTAATCA GCGATATCACTAAT TAA-3'	5'-ATACTAAGCTTTTAATTAGT GATATCGCTGATTAATTGC-3'	69.1 (28 nt (8GC))

GFP - GusB (2nd half)

1. Before IC2 and IC3

MSKGEELFTGVVPILVELDGDVNGHKFSVSSEGEEDATYGKLTCLKFICTTGKLPVPWPPTLVTTFG
YGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDF
KEDGNILGHKLEYNNSHNVIYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLL
PDNHYLSTQSALS KDPNEKRDH MVLLFEVTAAGITHGMDELYKL

NVVRIVAQPSLNISLQTLKRNRLFMFLCIGALCVLISTFAVSASSLFYVRYVLNDTGLFTVLVLV
 QNLVGTVASAPLVPGMVARIGKKNLFLIGALLGTCGYLLFFWVSVWSLPVALVALAIASIGQGVT
 MTVMWALEADTVEYGEYLTGVRIEGLTYSLSFSTRKCGQAIGGSIPAFILGLSGYIANQVQTPEV
 IMGIRTSIALVPCGFMLLAFVLIWFYPLTDDKKFKEIWEIDNRKKVQQQLISDITN

atgtctaaaggtgaagaattattcactgggtggtgtcccaatTTTgggtgaattagatgggtgatgt
 taatgggtcacaaatTTTctgtctccggtgaaggtgaaggtgatgtacttacggtaaattgacct
 taaaatTTTatttgtactactggtaaattgccagttccatggccaaccttagtcactactTTTcgggt
 tatgggtgttcaatgTTTTgttagataccagatcatatgaaacaacatgactTTTTcaagctctgc
 catgccagaaggttatgttcaagaaagaactatTTTTTcaaagatgacggtaactacaagacca
 gagctgaagtcaagTTTgaaggtgataccttagttaatagaatcgaattaaaaggtattgattTT
 aaagaagatggtaacatTTTtaggtcacaaattggaatacaactataactctcacaaatgTTTcat
 catggctgacaaacaaaagaatggatatcaaagttaacttcaaaattagacacaaacattgaagatg
 gttctgttcaattagctgaccattatcaacaaaatactccaattgggtgatgggtccagctcttgtta
 ccagacaaaccattacttatccactcaatctgccttatccaaagatccaaacgaaaagagagacca
 catggctcttgttagaattTgttactgctgctgggtattacccatgggtatggatgaattgtacaaat
 TaaatgtgggtacgtatcgTTtgcgcgagccgtcattgaatatcagctctgcaaaccctgaaacg
 gaatcgcccgctgTTtatgttgtgcatcggtgcgctgtgtgtgctgatttcgacctTTgc
 ggtcagcgctcgTcgTTgttctacgtgcgctatgtgttaaatgataccgggctgTTcac
 tgtgctgggtactgggtgcaaaacctgTTTgggtactgtggcatcggcaccgctgggtgccggg
 gatgggtcgcgaggatcggtaaaaagaataccttctctgattggcgctTTTgctgggaacctg
 cggttatctgctgTTcttctgggtTTccgctctgggtcactgccgggtggcgctgggtgcgTT
 ggccatcgcttcaattgggtcagggcgTTaccatgaccgtgatgtgggcgctggaagctga
 taccgtagaatacgggtgaatacctgaccggcggtgcgaattgaagggctcacctattcact
 attctcatttacccgtaaatgcggtcaggcaatcgaggttcaattcctgcctttatTTT
 ggggttaagcggatatatcgccaatcaggtgcaaacgccggaagttattatgggcatcgg
 cacatcaattgccttagtaccttgcggtattatgctactggcattcgTTattatctggtt
 ttatccgctcacggataaaaaattcaaagaaatcgTggttgaaattgataatcgtaaaaa
 agtgcagcagcaattaatcagcgatatcactaattaa

490

GFP - GusB (2nd half)

2. After IC2 and IC3

MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGGDATYGKLTCLKFICTTGKLPVPWPTLVTTFG
 YGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDF
 KEDGNILGHKLEYNNSHNVIYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLL
 PDNHYLSTQSALS KDPNEKRDH MVLLFVTAAGITHGMDELYKL
 TLKRNRLFMFLCIGALCVLISTFAVSASSLFYVRYVLNDTGLFTVLVLVQNLVGTVASAPLVPGM
 VARIGKKNLFLIGALLGTCGYLLFFWVSVWSLPVALVALAIASIGQGVTMTVMWALEADTVEYGE
 YLTGVRIEGLTYSLSFSTRKCGQAIGGSIPAFILGLSGYIANQVQTPEVIMGIRTSIALVPCGFMLLAFVLIWFYPLTDDKKFKEIWEIDNRKKVQQQLISDITN

atgtctaaaggtgaagaattattcactgggtggtgtcccaatTTTgggtgaattagatgggtgatgt
 taatgggtcacaaatTTTctgtctccggtgaaggtgaaggtgatgtacttacggtaaattgacct
 taaaatTTTatttgtactactggtaaattgccagttccatggccaaccttagtcactactTTTcgggt
 tatgggtgttcaatgTTTTgttagataccagatcatatgaaacaacatgactTTTTcaagctctgc
 catgccagaaggttatgttcaagaaagaactatTTTTTcaaagatgacggtaactacaagacca

gagctgaagtcaagtttgaaggtgataccttagttaatagaatcgaattaaaggtattgatttt
aaagaagatggtaacatttttaggtcacaaattggaatacaactataactctcacaaatgtttacat
catggctgacaaacaaaagaatggatcaaaagttaacttcaaaattagacacaacattgaagatg
gttctgttcaattagctgaccattatcaacaaaatactccaattggatggatccagtccttgta
ccagacaaccattacttatccactcaatctgccttatccaaagatccaaacgaaaagagagacca
catggctcttgtagaatttggtactgctgctgggtattacccatgggtatggatgaattgtacaaat
Taaccctgaaacggaatcgcccgctgtttatgtttgtgcatcggtgcgctgtgtgtgctgat
ttcgacctttgcgggtcagcgccctcgctgtttgttctacgtgcgctatgtgttaaatagatac
cgggctgttcaactgtgctgggtactgggtgcaaacctgggttggtactgtggcatcggcacc
gctgggtgcccgggatgggtcgcgaggatcggtaaaaagaataccttcttgattggcgcttt
gctgggaacctgcgggttatctgctgttcttctgggtttccgtctgggtcactgcccgggtggc
gttggttgcggttgcccatcgcttcaattgggtcagggcggttaccatgaccgtgatgtgggc
gctggaagctgataccgtagaatacgggtgaatacctgaccggcggtgcgaattgaagggct
cacctattcactatttctcatttaccgcgtaaatgcgggtcaggcaatcggaggttcaattcc
tgcctttattttgggggttaagcggatatatcgccaatcaggtgcaaacgccggaagttat
tatgggcacccgcacatcaattgccttagtaccttgcggtatttatgctactggcattcgt
tattatctggttttatccgctcacggataaaaaattcaaagaaatcgtggttgaaattga
taatcgtaaaaaagtgagcagcaattaatcagcgatatcactaattaa

474

GFP - GusB (2nd half)

3. Half way through IC2 before IC3

MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGGDATYGKLT LKFICTTGKLPVPWPTLVTTFG
YGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDF
KEDGNILGHKLEYNNSHNVIYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLL
PDNHYLSTQSALS KDPNEKRDH MVLLFVTAAGITHGMDELYKL
VAQPSLNISLQTLKRNRLFMLCIGALCVLISTFAVSASSLFFVRYVLNDTGLFTVLVLVQNLVG
TVASAPLVPGMVARIGKKN TFLIGALLGTCGYLLFFWVSVW SLPVALVALAIASIGQGVMTVMW
ALEADTVEYGEYLTGVRIEGLTYSLFSFTRKCGQAIGGSIPAFILGLSGYIANQVQTPEVIMGIR
TSIALVPCGFMLLAFVLIWFYPLTDKKFKEIVVEIDNRKKVQQQLISDITN

atgtctaaaggtgaagaattattcactgggtgttgtcccaattttgggtgaattagatgggtgatgt
taatgggtcacaaattttctgtctccgggtgaaggtgaaggtgatgctacttacggtaaaattgacct
taaaattttatttgtactactggtaaaattgccagttccatggccaaccttagtcactactttcggt
tatgggtgttcaatgttttgc tagataccagatcatatgaaacaacatgactttttcaagtctgc
catgccagaaggttatgttcaagaaagaactattttttcaaagatgacggtaactacaagacca
gagctgaagtcaagtttgaaggtgataccttagttaatagaatcgaattaaaggtattgatttt
aaagaagatggtaacatttttaggtcacaaattggaatacaactataactctcacaaatgtttacat
catggctgacaaacaaaagaatggatcaaaagttaacttcaaaattagacacaacattgaagatg
gttctgttcaattagctgaccattatcaacaaaatactccaattggatggatccagtccttgta
ccagacaaccattacttatccactcaatctgccttatccaaagatccaaacgaaaagagagacca
catggctcttgtagaatttggtactgctgctgggtattacccatgggtatggatgaattgtacaaat
Tagttgcgcgagccgtcattgaatatcagtcctgcaaacctgaaacggaatcgcccgctgtt
tatgttgtgcatcggtgcgctgtgtgtgctgatttcgacctttgcgggtcagcgccctcgto
gttgttctacgtgcgctatgtgttaaatagataccgggctgttcaactgtgctgggtactgggt
gcaaacctgggttggtactgtggcatcggcaccgctgggtgccgggggatgggtcgcgaggat

cggtaaaaagaataaccttcctgattggcgcttttgctgggaacctgcgggttatctgctgtt
cttctggggtttccgtctggtcactgccggtggcgcttggttgcggtggccatcgcttcaat
tggtcagggcggttaccatgaccgtgatgtgggcgctggaagctgataccgtagaatacgg
tgaataacctgaccggcggtgcgaattgaagggctcacctattcactattctcatttacccg
taaatacgcggtcaggcaatcggaggttcaattcctgcctttattttgggggttaagcggata
tatcgccaatcaggtgcaaacgccggaagtattatgggcacatcgccacatcaattgcctt
agtaccttgcggtatttatgctactggcattcggttattatctgggttttatccgctcacgga
taaaaaattcaaagaaatcgtgggttgaaattgataatcgtaaaaaagtgcagcagcaatt
aatcagcgatatcactaattaa

485

GFP - GusB (2nd half)

4. After IC2 half way through IC3

MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWPTLVTTFG
YGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDF
KEDGNILGHKLEYNNSHNVIYIMADKQKNGIKVNEKIRHNIEDGSVQLADHYQQNTPIGDGPVLL
PDNHYLSTQSALS KDPNEKRDH MVLLFVTAAGITHGMDELYKL
ISLQTLKRNRF LFM LCIGALCVLISTFAVSASSLF YVRYVLNDTG LFTVLVLVQNLVGTVASAPL
VPGMVARIGKKN TFLIGALLGTCGYLLFFWVSVW SLPVALVALAIASIGQGVMTVMWALEADTV
EYGEYLTGVRIEGLTYSLSFSTRK CGQAIGGSIPAFILGLSGY IANQVQTPREVIMGI RTSIALVP
CGFMLLAFVIIWFYPL TDKKFKEIVVEIDNRKKVQQQLISDITN

atgtctaaaggtgaagaattattcactgggtgttgcccaattttgggtgaattagatgggtgatgt
taatgggtcacaaattttctgtctccggtgaaggtgaaggtgatgctacttacggtaaatgacct
taaaatttatgttactactggtaaatgcccaggtccatggccaaccttagtcactactttcgggt
tatgggtgttcaatgttttgcctagatacccagatcatatgaaacaacatgactttttcaagtcctgc
catgccagaaggttatgttcaagaaagaactatttttttcaagatgacggtaactacaagacca
gagctgaagtcaagtttgaaggtgataccttagttaatagaatcgaattaaaaggtattgatttt
aaagaagatggtaacattttaggtcacaaattggaatacaactataactctcacaaatgtttacat
catgggtgacaaacaaaagaatggatcaaaagttaacttcaaaattagacacaaacattgaagatg
gttctgttcaattagctgaccattatcaacaaaatactccaattgggtgatgggtccagtccttgta
ccagacaaccattacttatccactcaatctgccttatccaaagatccaaacgaaaagagagacca
catgggtcttggttagaatttggttactgctgctgggtattaccatgggtatggatgaattgtacaaat
Taatcagtcctgcaaaccctgaaacggaatcgcccgcctgtttatgtttgtgcatcggtgcgct
gtgtgtgctgatttcgacctttgcgggtcagcgcctcgctcggtgttctacgtgcgctatgt
gttaaataatgataccgggctgttactgtgctgggtactgggtgcaaaacctgggtgggtactgt
ggcatcggcaccgctgggtgccggggatgggtcgcgaggatcggttaaaaagaataccttcct
gattggcgcttttgctgggaacctgcgggttatctgctgttcttctggggtttccgtctggtc
actgccggtggcgcttggttgcggttgccatcgcttcaattgggtcagggcggttaccatgac
cgtgatgtgggcgctggaagctgataccgtagaatacgggtgaataacctgaccggcggtgcg
aattgaagggctcacctattcactattctcatttacccgtaaatgcgggtcaggcaatcgg
aggttcaattcctgcctttattttgggggttaagcgggatatatcgccaatcaggtgcaaac
gccggaagttattatgggcacatcgccacatcaattgccttagtaccttgcggtatttatgct
actggcattcggttattatctgggttttatccgctcacgggataaaaaattcaaagaaatcgt
gggttgaaattgataatcgtaaaaaagtgcagcagcaattaatcagcgatatcactaatta
a

GFP - GusB (2nd half)

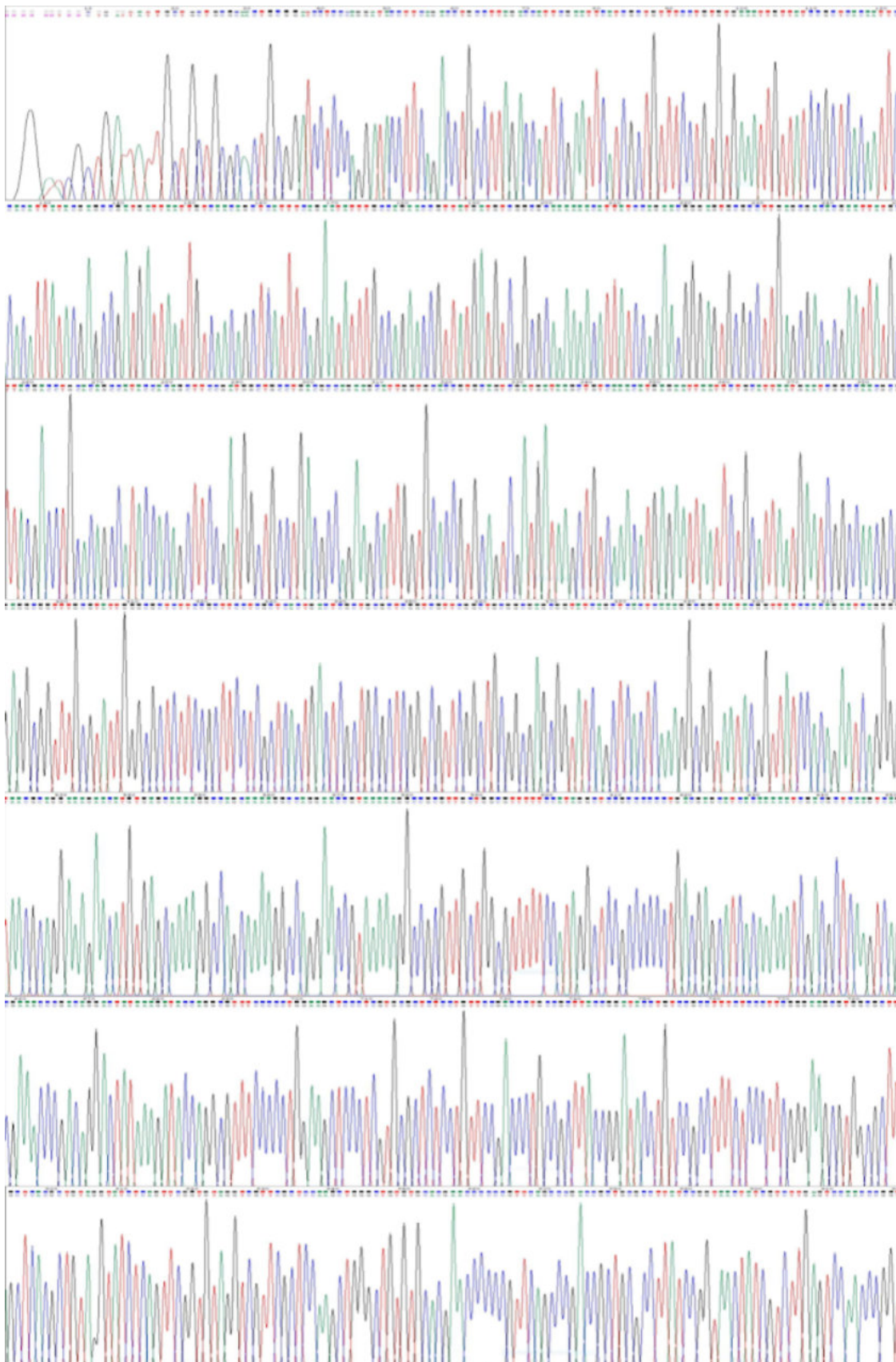
5. Just after IC2. Between IC2 and IC3 244:

MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWPPTLVTTFG
YGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDF
KEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLL
PDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYKL

SLNISLQTLKRNRLFMFLCIGALCVLISTFAVSASSLFYVRYVLNDTGLFTVLVLVQNLVGTVAS
APLVPGMVARIGKKN TFLIGALLGTCGYLLFFWVSVW SLPVALVALAIASIGQGVMTVMWALEA
DTVEYGEYLTGVRIEGLTYSLSFSTRK CGQAIGGSIPAFILGLSGYIANQVQTPEVIMGI RTSIA
LVPCGFMLLAFVLIWFYPL TDKKFKEIVVEIDNRKKVQQQLISDITN

atgtctaaaggtgaagaattattcactggtggttgcctcaatttttggttgaattagatggtgatgt
taatggtcacaaattttctgtctccggtgaaggtgaaggtgatgctacttacggtaaattgacct
taaaattttatttgtactactggttaaattgccagttccatggccaaccttagtcactactttcgggt
tatggtggtcaatgttttgcctagataccagatcatatgaaacaacatgactttttcaagtctgc
catgccagaaggttatgttcaagaaagaactatttttttcaaagatgacggtaactacaagacca
gagctgaagtcaagtttgaaggtgataccttagttaatagaatcgaattaaaagggtattgatttt
aaagaagatggttaacatttttaggtcacaaattggaatacaactataactctcacaaatgtttacat
catggctgacaaacaaaagaatggtatcaaagttaacttcaaaattagacacaaacattgaagatg
gttctgttcaattagctgaccattatcaacaaaatactccaattggtgatggtccagtccttggtta
ccagacaaaccattacttatccactcaatctgccttatccaaagatccaaacgaaaagagagacca
catggtccttggttagaatttggttactgctgctggtattacccatggtatggatgaattgtacaaat
Tatcattgaatatcagtcctgcaaaccctgaaacggaatcgcccgcctgtttatgttgtgcat
cgggtgcgctgtgtgtgctgatttgcacctttgcggtcagcgcctcgtcgttgttctacgt
gcgctatgtgttaaatgataccgggctgttcaactgtgctggtactggtgcaaaacctggt
tggtactgtggcatcggcaccgctggtgccggggatggtcgcgaggatcggtaaaaagaa
taccttcctgattggcgctttgctgggaacctgcggttatctgctgttcttctgggtttc
cgtctggtcactgccggtggcgcttggttgcgcttgccatcgcttcaattggtcagggcgt
taccatgaccgtgatgtggcgctggaagctgataccgtagaatacgggtgaataacctgac
cggcgtgcgaattgaagggctcacctattcactatttctcatttaccgtaaatgcggtca
ggcaatcgaggttcaattcctgcctttattttgggggttaagcggtatatcgccaatca
ggtgcaaacgccggaagtattatgggcacatcgacacatcaattgccttagtaccttgcg
atattatgctactggcattcggtattatctggttttatccgctcacggataaaaaattcaa
agaaatcggtggttgaaattgataatcgtaaaaaagtgacgacgcaattaatcagcgatat
cactaattaa

Appendix 10 Sequencing Analysis for clone 31B



Appendix 11 Laboratory Book Notes