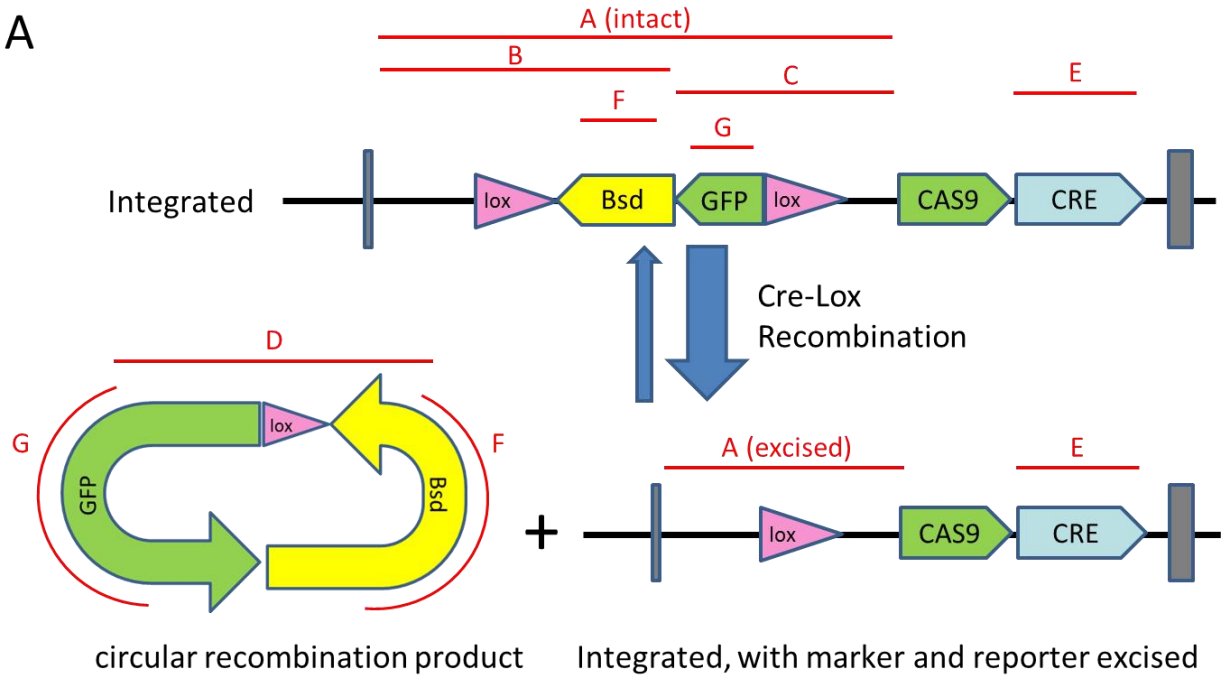
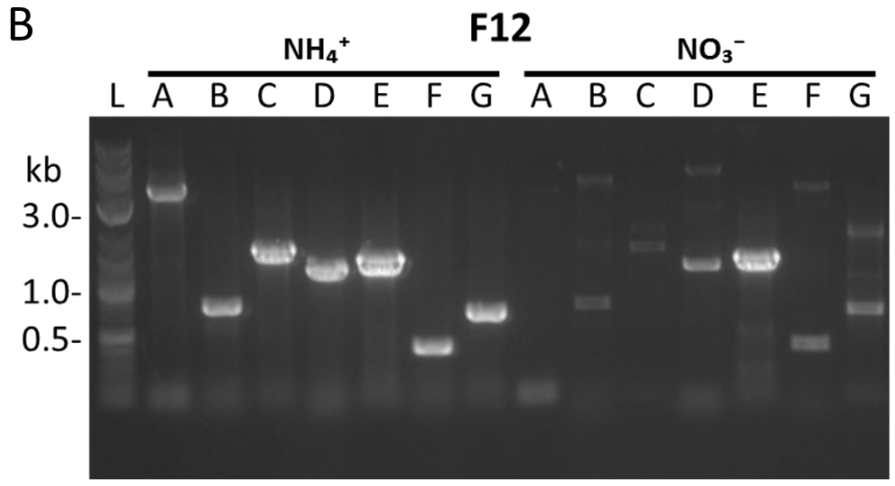
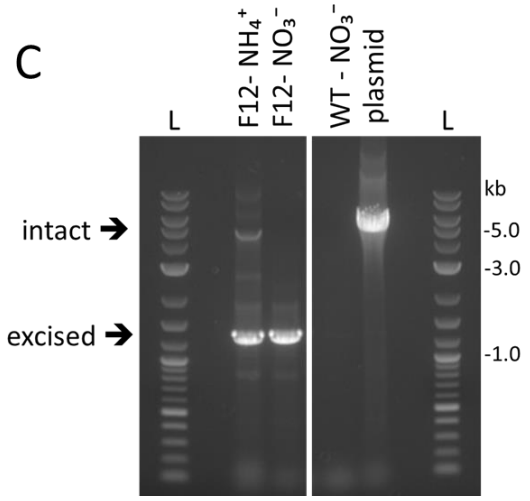
SI appendix:

Supplementary figures


A
A (intact)
B
C
E
F
G
Integrated
lox
Bsd
GFP
lox
CAS9
CRE
Cre-Lox
Recombination
D
G
F
GFP
lox
Bsd
A (excised)
E
lox
CAS9
CRE
circular recombination product
Integrated, with marker and reporter excised


B
F12
NH4+
NO3−
L A B C D E F G A B C D E F G
kb
3.0-
1.0-
0.5-


C
L
F12- NH4+
F12- NO3−
WT - NO3−
plasmid
L
kb
-5.0
-3.0
-1.0
intact
excised

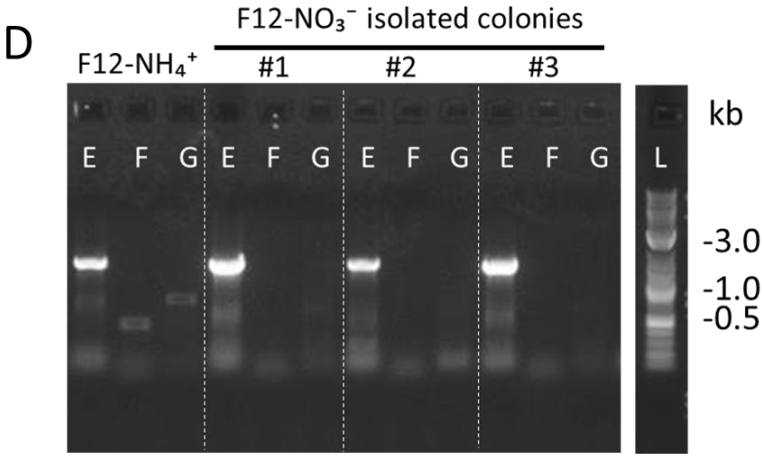
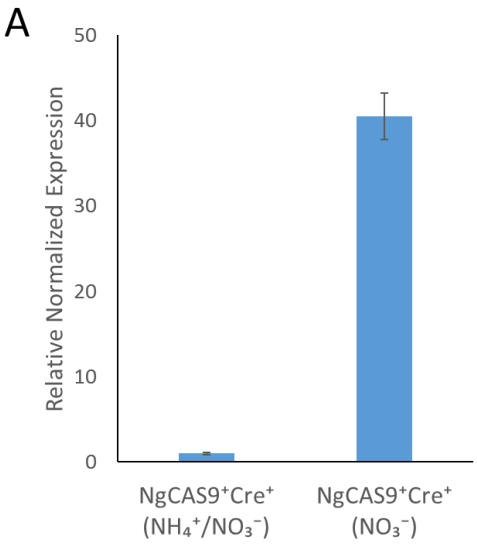
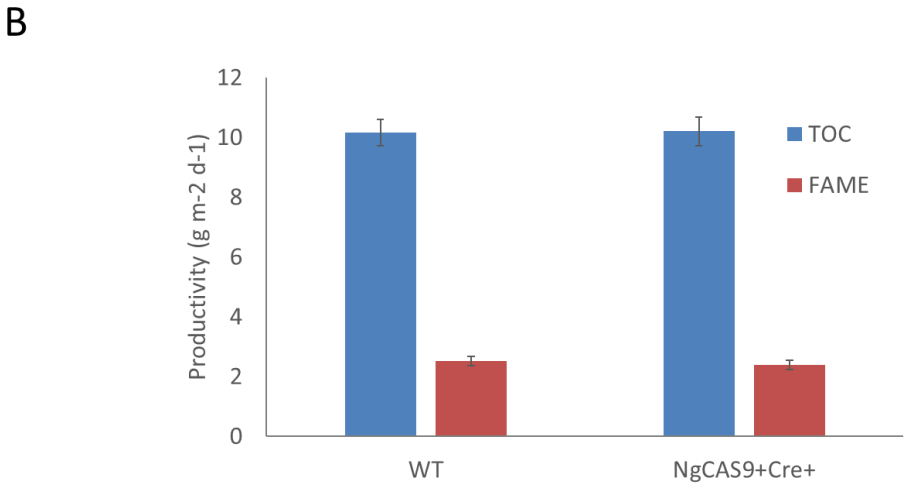

D
F12-NO3− isolated colonies
F12-NH4+
#1
#2
#3
E F G E F G E F G E F G
L
kb
-3.0
-1.0
-0.5

Figure S1. Molecular analysis of the progression from transformation to isolation of NgCAS9^{+}Cre^{+}. (A) Illustration of plasmid integration and subsequent excision of the floxed *GFP/Bsd* region, and expected amplicons from primer sets A-G (amplicons shown as red lines, not to scale). Expected amplicon sizes from the four main states of progression can be found in Table S3 (i.e. fully intact with no recombination, partially excised with circular recombination product remaining, fully excised with circular recombination product remaining, and fully excised with no circular recombination product remaining). Primer sequences can be found on Supplemental Table S1. (B) PCR assessment of Cre-mediated marker/reporter excision progress in NgCAS9^{+}Cre+ in CRM or CIM media. Primer set "A" spans the entire floxed region to detect the presence of intact floxed region (3.4 kb band), and/or a smaller 0.19 kb band indicating that recombination and excision have started to occur. Primer set "B" spans the 5' lox site, where an 0.85 kb band indicates intact floxed region and lack of a PCR amplicon indicates excision. Primer set "C" spans the 3' lox site where a 1.6 kb band indicates intact floxed region and lack of PCR amplicon indicates excision. Primer set "D" detects the presence of the circular recombination product; in the intact integrated state these primers are oriented away from each other, whereas in the circular recombination product they are oriented to produce an expected 1.3 kb band by PCR. Primer set "E" amplifies the Cre gene with an expected 1.4 kb band. Primer set "F" amplifies the Bsd gene with an expected band size of 0.4 kb. Primer set "G" amplifies the GFP gene with an expected 0.7kb gene. (C) Primer set "H" (Table S1) which produces a 1.3kb band for the excised form versus a 4.9kb band, included to provide better resolution of the excised form observed using primer set-A.(D) Analysis of single cell colony isolates from F12-NO_3^- culture screened with primers to detect Cre, Bsd, and GFP (Primer sets "E", "F", "G" respectively). Isolate #2 was carried forward as the final NgCas9^{+}Cre^{+} strain. The Lane labeled "L" in all gels refers to the NEB 2-Log DNA Ladder (0.1-10.0 kb).





C

Strain	σ_{PSII} (Å², 450nm)	Chl a/TOC (mg/g)
WT	454±11	64 ± 1
NgCAS9⁺Cre⁺	454±9	62 ± 2

Figure S2. qRT-PCR analysis of Cre expression and comparison of NgCAS9⁺Cre⁺ [22]and wild type (WT) strain performance under semi-continuous growth. (A) qRT-PCR results demonstrate that Cre expression is greatly repressed in the presence of NH_4^+. Expression levels in NH_4^+/NO_3^- are given a value of 1, while expression levels in NO_3^- are reported as relative to NH_4^+/NO_3^-. Expression levels were observed to increase by 40-fold in NO_3^-. (B) Semi-continuous growth mode TOC (blue) and FAME (red) areal productivities for wild type *Nannochloropsis* (WT) and NgCAS9⁺Cre⁺. (C) Chlorophyll content and functional absorption cross section of PSII (σ_{PSII}). Shown are the average and standard deviation of technical triplicates from cell culture duplicates (N = 6).

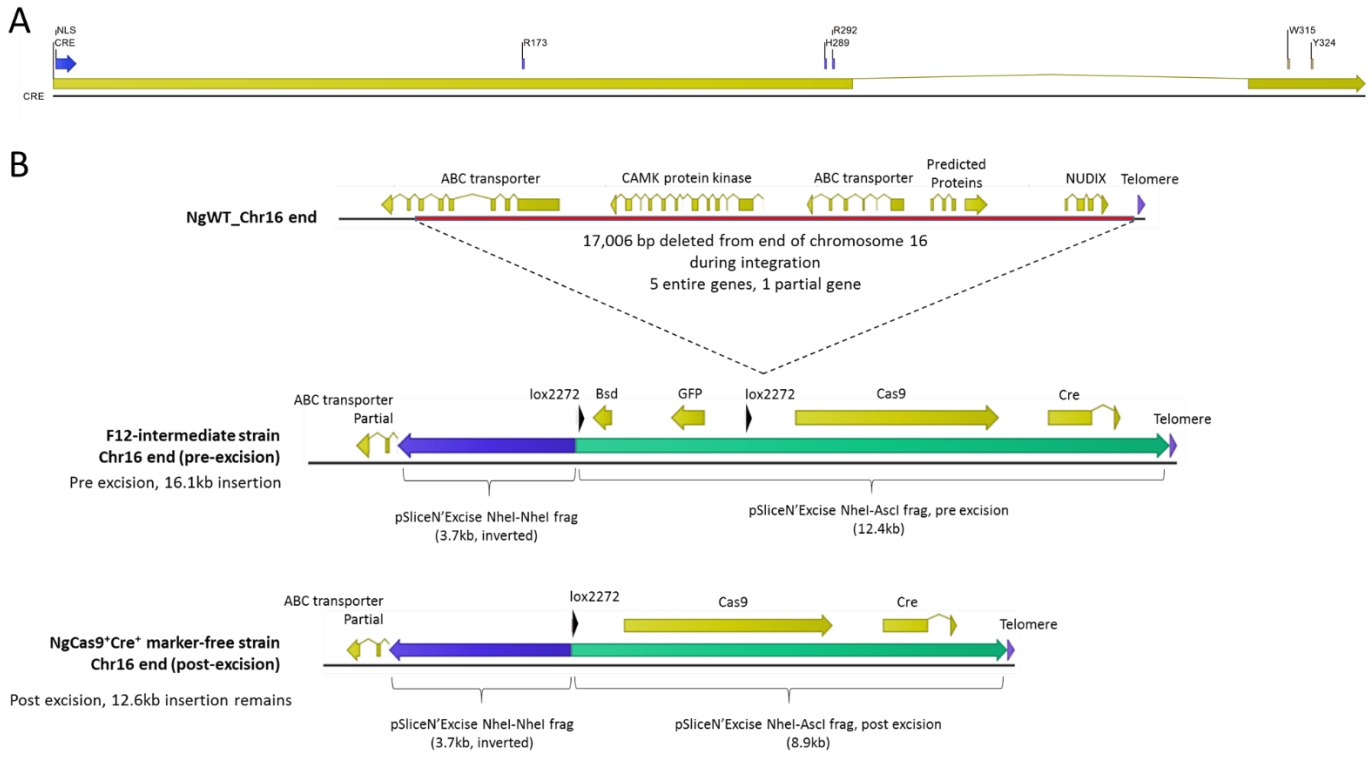
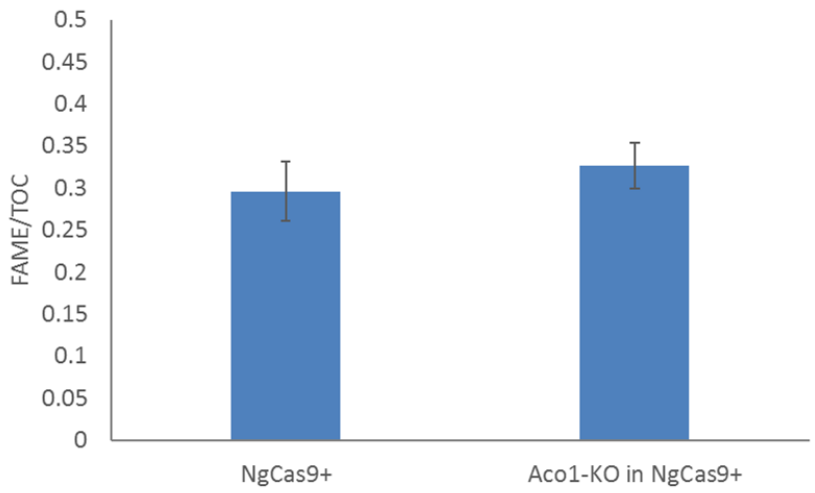


Figure S3. Architecture of the intronated recombinant Cre gene and locus of stable inserted Cas9 and Cre transgene in NgCAS9^{+}Cre^{+}. (A) Recombinant Cre gene has an N terminal NLS (SV40, shown in blue) and an intron from the native RPS4x gene transplanted between catalytic (R173, R289, and R292) and nucleophilic residues (W315 and Y324) to defunctionalize basal level CRE expression in *E.coli*. (B) Scheme depicting the integration of the Cas9 and Cre expression construct as deduced from PacBio genome sequencing of NgCas9^{+}Cre^{+}.

A



B

Strain	gene target	†Selectable marker	‡On-target marker insertion detected	Total transformants analyzed by PCR
NgCas9+	*Aco1*	HygR	90	160
NgCas9+Cre+	*Aco1*	Bsd-GFP	9	19

†Selectable markers co-transformed with the gRNA targeting Aco1 for Cas9-mediated insertional mutagenesis

‡A PCR band at the expected mass was observed (xx kb for HygR; yy kb for BSD-GFP).

Figure S4. Comparison of Cas9 mediated insertional mutagenesis in NgCas9+ (22) and NgCas9+Cre+ generated in this study. (A) Characterization of carbon partitioning to lipid (FAME/TOC) in an acyl-CoA oxidase gene knockout (*Aco1*) generated in the previously published *Nannochloropsis* Cas9 editor line (NgCas9+). Cultures were grown in biological duplicate in batch mode as previously described (22), and FAME (fatty acid methyl ester) and TOC (total organic carbon) measurements were taken daily for three days. Shown is the average and standard deviation of the FAME/TOC ratio used as a proximate indicator of partitioning of carbon to lipid. (B) Summary of results for targeting the *Aco1* gene in the original Cas9 editor line NgCas9+ and in the NgCas9+Cre+ line described in this study. The number of on-target insertions vs the total number of transformants analyzed by PCR is shown.

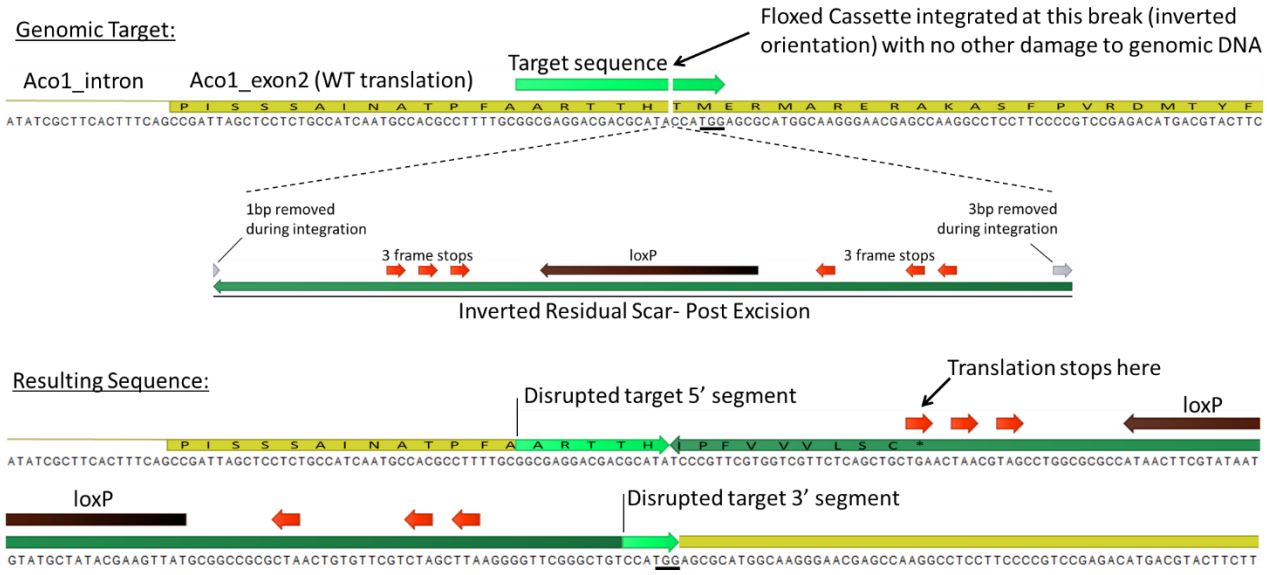


Figure S5. Diagram depicting PCR sequencing results at the *Aco1* knockout locus in isolates 1 to 3, where exon 2 was disrupted by the *lox* "scar" in an inverted orientation. The location and orientation of the insertion are indicated, as well as 1 and 3 bp deletions observed on the 5' and 3' ends of the integrated construct, respectively. All three isolates showed identical sequence results.

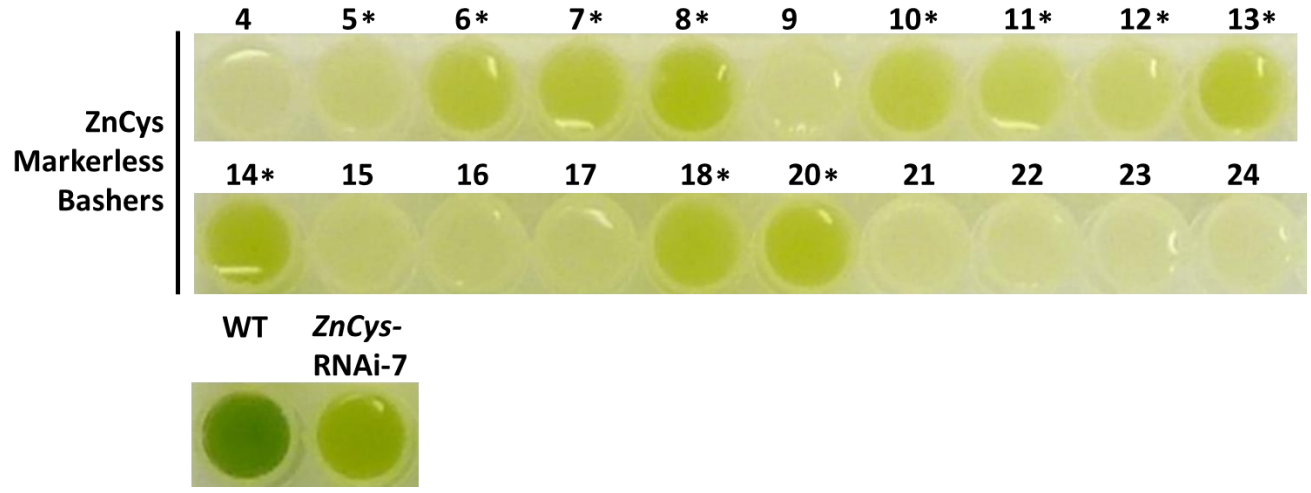


Figure S6. Range of pale phenotypes observed when *ZnCys* "5'-Bash" strains were cultured with medium containing NO_3^- as the sole N source, compared to wild type *Nannochloropsis gaditana.* A *ZnCys* RNAi strain from (22) was cultured in the same media, and strains exhibiting

visual pigmentation level comparable to the RNAi line were carried forward for isolation and genotyping (indicated here by asterisks).

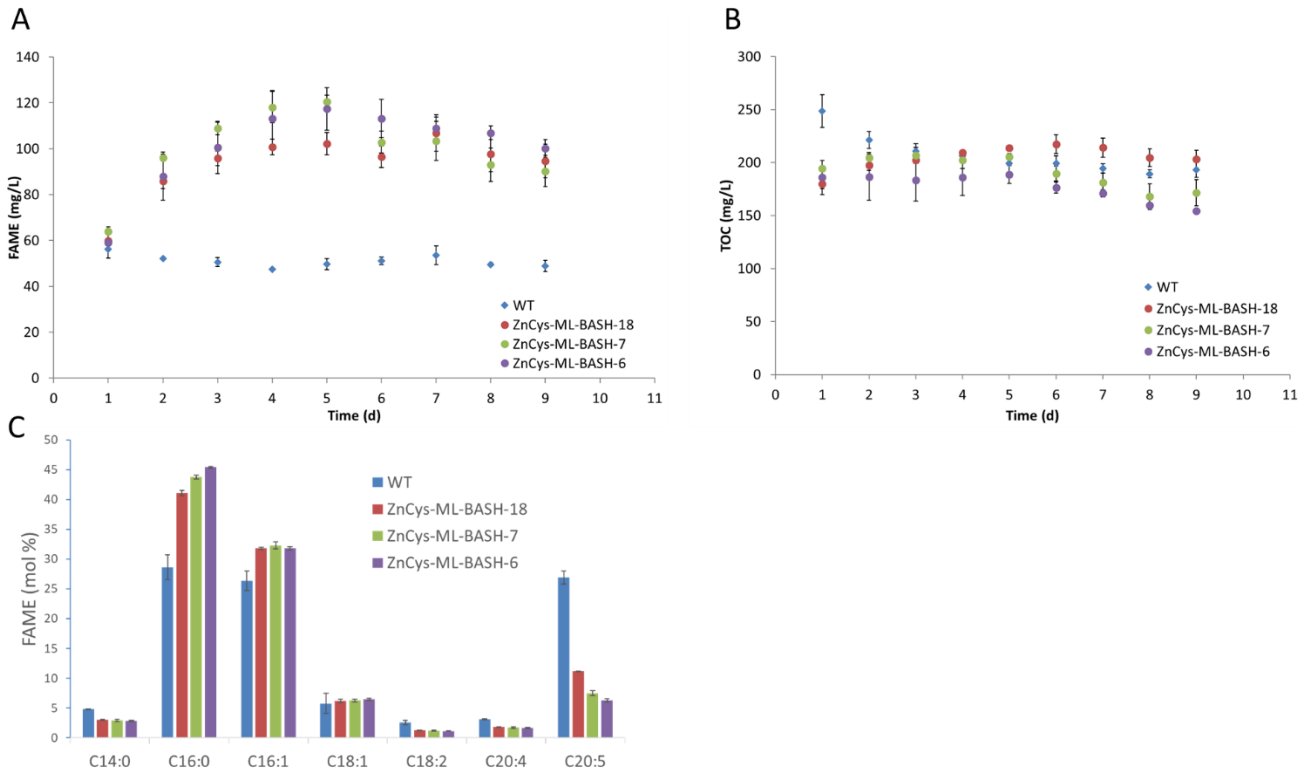


Figure S7. Volumetric FAME (A) and TOC (B) levels observed for wild-type and *Zncys*-ML-BASH lines under semi-continuous growth mode, where a 30 % daily dilution rate was applied. Shown are the average and standard deviation for biological triplicate cultures. Data from days 3 to 9 were used in the calculation of summary aerial productivities shown in Figure 3C. Total fatty acid profiles (C) were calculated from FAME data collected on day 5 of the run depicted in panels A and B.

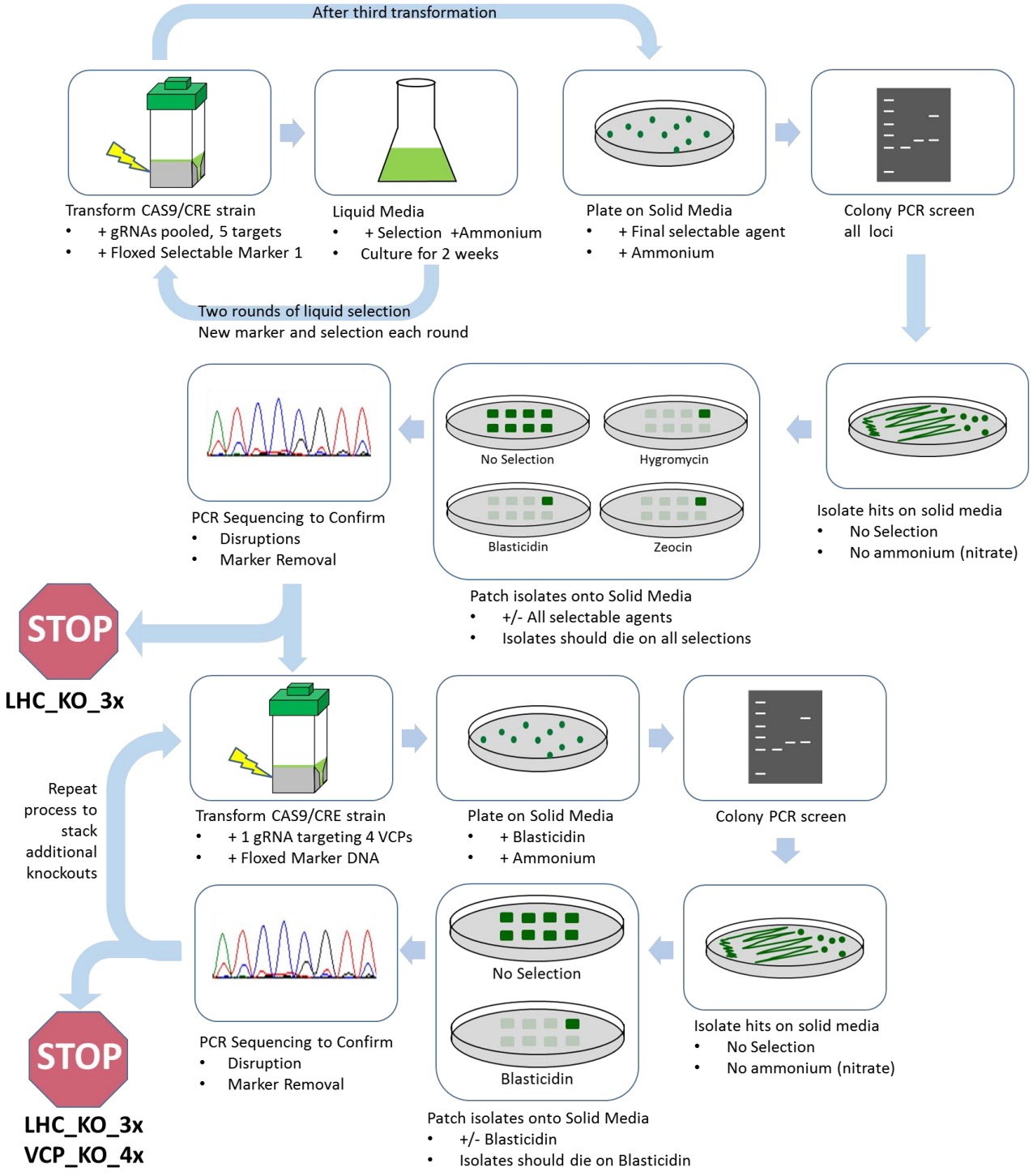


Figure S8. Rapid generation of strains harboring stacked knockout mutations in the light harvesting complex (LHC, non-VCP) and violaxanthin carrier protein (VCP) encoding gene families via a multiplexed iterative strategy in NgCAS9^{+}Cre^{+}. The diagram illustrates the method employed for the rapid stacking of knockouts, where briefly, the NgCAS9^{+}Cre^{+} strain was

transformed with 5 separate and unique gRNA molecules which target 5 highly expressed Light Harvesting Complex genes, and a floxed blasticidin resistance marker, and the transformed cells were inoculated into liquid media with blasticidin and NH_4^+ and cultured for 2 weeks. This culture was then transformed with the same 5 gRNAs again, but this time with a floxed hygromycin resistance marker, and the transformed cells were inoculated into liquid media with hygromycin and NH_4^+ and again cultured for 2 weeks. This culture was then transformed with the same 5 gRNAs, and a floxed zeocin resistance marker, and was plated onto solid media containing NH_4^+ and zeocin. Transformants were screened for inserts at the 5 loci by colony PCR with locus specific primers, and potential hits were streaked out for isolation on solid media with no selection and no NH_4^+. Isolated lines were analyzed by colony PCR at the 5 loci using locus specific primers which yielded one line, LHC-3x-, that was shown to have residual lox scars inserted in three of the five target loci and was sensitive to all antibiotics used. LHC-3x- was then transformed with a single gRNA which targets four separate violaxanthin-chlorophyll-a genes (VCP 1-4) and a floxed blasticidin resistance marker. Transformants were screened by colony PCR using unique locus specific primers. After cre-mediated marker excision and plating to isolation, one line, LHC_KO_3x/VCP_KO_4x, was found to have all four loci disrupted.

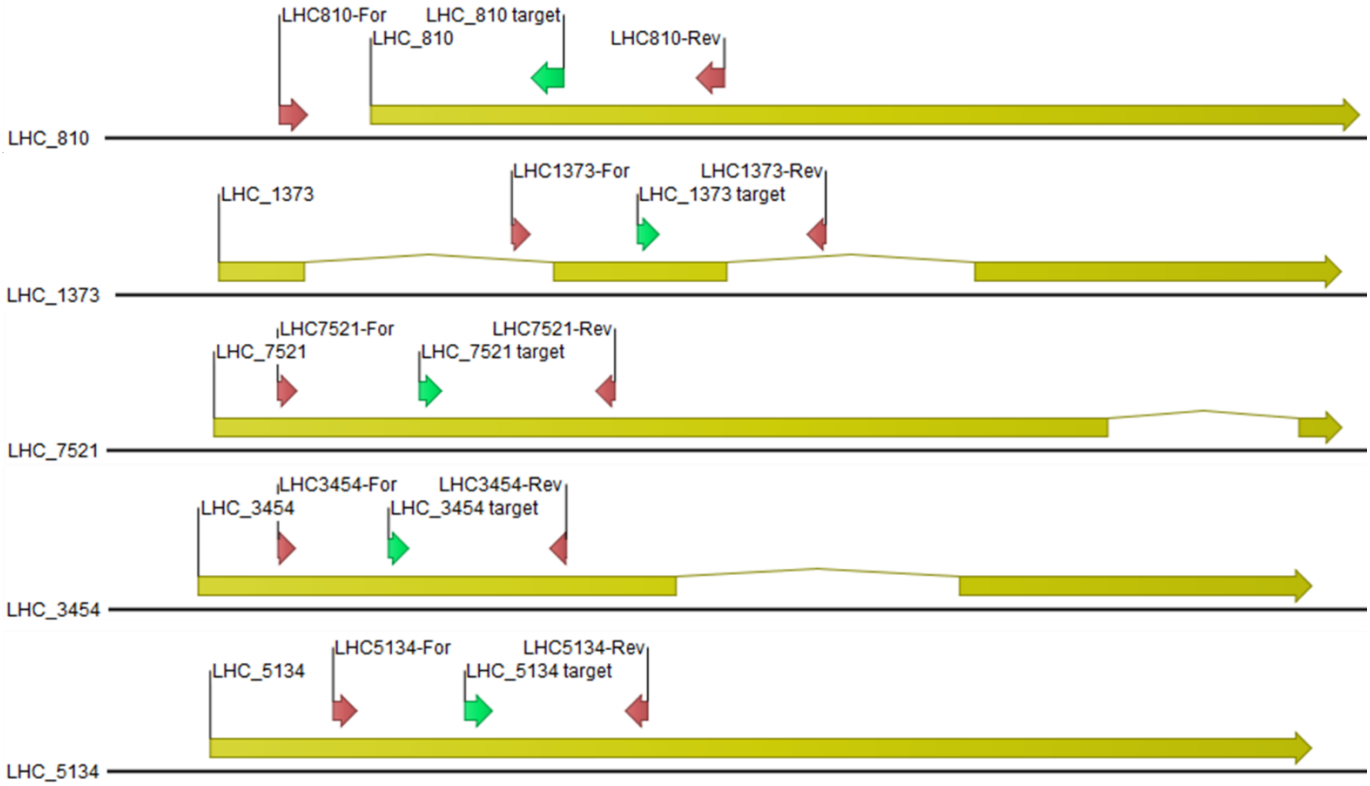


Figure S9. Five light harvesting complex (LHC, non-VCP) genes targeted for knockout and marker recycling. The coding sequences are shown in yellow, CRISPR targets in green, and screening primers in red.

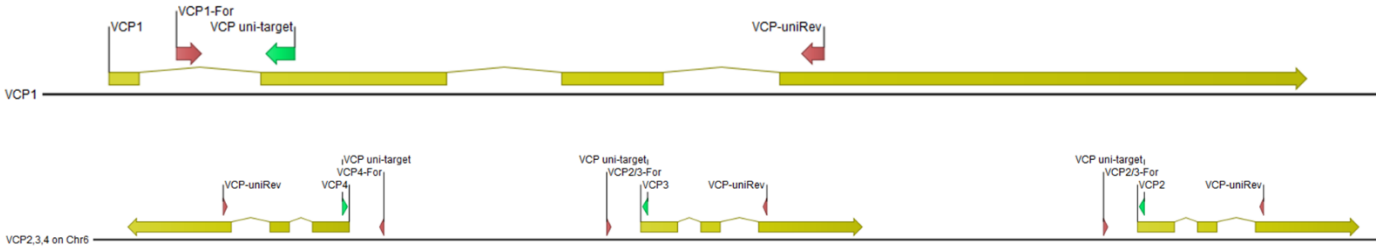


Figure S10. Light harvesting genes of the violaxanthin chlorophyll-a protein (VCP) family targeted for knockout and marker recycling. The coding sequences are shown in yellow, CRISPR targets in green, and screening primers in red.

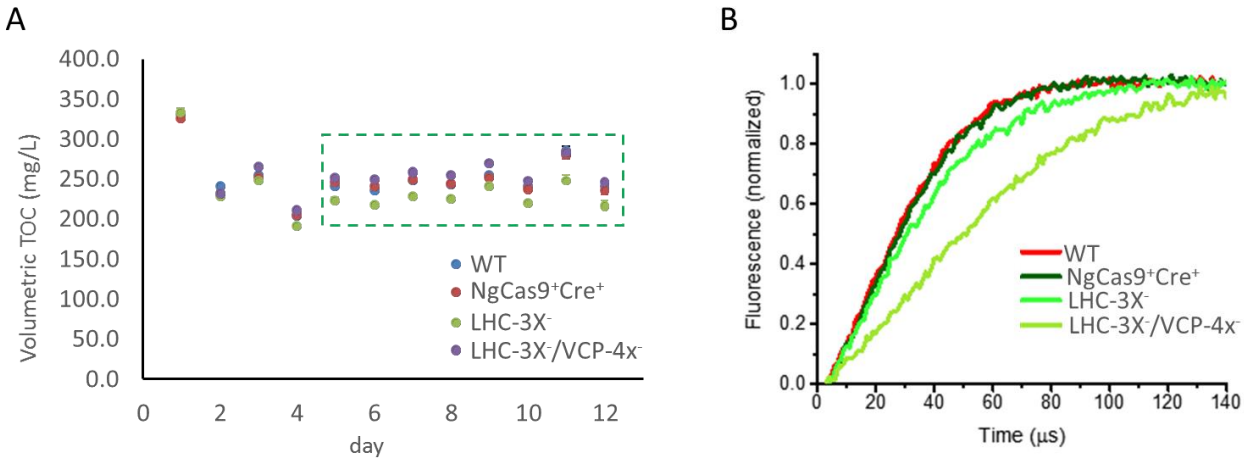


Figure S11. Characterization of LHC knockout strains for TOC productivity (A) and fluorescence induction (B) using the FIRe fluorimeter. Volumetric TOC for wild type (WT), NgCas9^{+}Cre^{+}, and the LHC mutants LHC-3X^{-} and LHC-3X^{-}/VCP-4X^{-} represent the average and standard deviation of cell culture triplicates for cultures diluted daily at 30 %; the data points within the dashed green box were used in determining areal TOC productivities summarized in Figure 4.

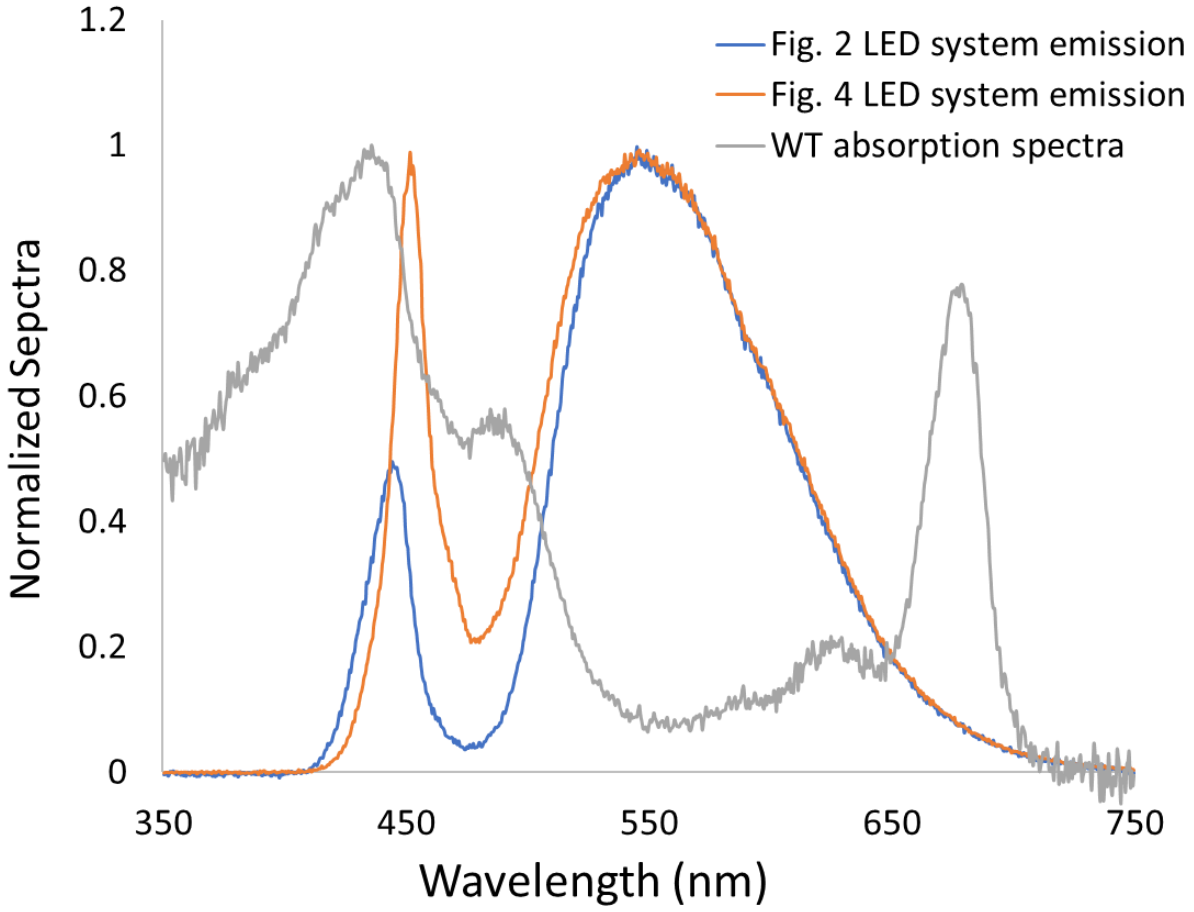


Figure S12. Peak normalized absorbance and emission spectra. Absorption spectra for wild type *N. gaditana* culture is shown in gray (measured using the Perkin Elmer Lambda 650 spectrophotometer as described in Methods). Emission spectra for the LED system used to assess strain productivity described in Figure 2 is shown in blue; Emission spectra for the updated LED system used to assess strain productivity in Figure 4 is shown in orange (both spectra were measured using Ocean Optics Jax Spectrometer). Spectra were normalized by setting the maxima to 1.

Supplemental tables:

Table S1. Primers used in this study.

I. Primers sequences used in analysis of the progression Cre-mediated marker removal leading to NgCAS9+Cre+.	
A/B/H-For	ctagcggccgcataac
A/C-Rev	AAACCATAATATGCCGTACCAACTG
F-For/B-Rev	atggccaagcctttgtcc
C-For/G-Rev	ttattcttcaccggcatctgc
D-For	CACAGCGGTTGGCATCAGGGAGTTGC
D-Rev	ggtgatgcggcactcgatctccatgg
E-For	ggtgcatcagaacttgcctgccttgcc
E-Rev	GCACCATTGCTCCCGTCTCGGAGTCC
F-Rev	tcaaccctcccagacgtagc
G-For	atgttggagagcgacgagagc
G-Rev	ttattcttcaccggcatctgc
H-Rev	CTTGAACTTCTTGGAAGGCACC
II. Screening Primers for CAS9 knockouts	
Aco1-For	TCAAAGATCATTTAGCAGAGACGGGGGCTGTGGCGCG
Aco1-Rev	AGTCGAGAGATGGTGCGTTCACGGATCTAGCCAGAG
ZnCys-5'Bash-For	TAGCAGAGCAGGCTCATCAC
ZnCys-5'Bash-Rev	GAATATGTGGTCTAGCTCGT
LHC810-For	CCCTTACTCCATCCCCAGAT
LHC810-Rev	ACTCGTCCCACTTGGTAGGC
LHC1373-For	TGGCTATGCTTCTCGTTCCT
LHC1373-Rev	ATTCCCCACACGACATCTCT
LHC7521-For	CTTCATGGGCAAGAACTTCG
LHC7521-Rev	GCGAAATCAGGGTTGGATAG
LHC3454-For	CTCTTCCCTCCCGAAAAACT
LHC3454-Rev	GCACCGACTTCGAGAACAC
LHC5134-For	TCTTCCTCCACCAGCTTTTT
LHC5134-Rev	GGAGGGTTGTTCGAGAAGC
VCP1-For	GTCTTTCTGGCTGGACTTGC
VCP-uniRev	GGTCACCAGGGTGTGGAG
VCP2/3-For	GACCCCGGAGGATCTTAGAC
VCP4-For	GGTCACCAGGGTGTGGAG

Table S2. Cas9 sites for loci targeted in this study.

CAS9 target sites	Sequence (PAM underlined)
Aco1 target (direct strand)	GGCGAGGACGACGCATACCATGG
ZnCys_5'UTR target (direct strand)	GGGACTGTCCCATTGTGCCGG
LHC_810 target (complementary strand)	GGGGGCTTCCAGGAAAGGGACGG
LHC_1373 target (direct strand)	GGCGAGCTTGCTGGGGATGTAGG

LHC_7521 target (direct strand)	GGTCTCTATCCCCAAACCCGCGG
LHC_3454 target (direct strand)	GGCCCTTCTCTGCTCCTCCGCGG
LHC_5134 target (direct strand)	GGCCCGTCGTGCCACCGGCGCGG
VCP uni-target (direct strand, present in all 4 VCP genes)	GGAGACGGTGAGCAGAGCGGCGG

Table S3. Description of expected PCR results for primer sets A-H.

Primer Set	Amplifies	Expected PCR bands (kb) from four states of progression			
		Fully Intact, no recombination	Partially excised, circular recombination product remains	Fully excised, circular recombination product remains	Fully excised, no circular recombination product remains
A	entire floxed region	3.7	0.185/3.7 (mix)	0.185	0.185
B	5' lox site	0.85	0.85 (faint)	no band	no band
C	3' lox site	1.6	1.6 (faint)	no band	no band
D	circular, excised recombination product	no band	1.3	1.3	no band
E	Cre only	1.4	1.4	1.4	1.4
F	Bsd only	0.4	0.4	0.4	no band
G	GFP only	0.7	0.7	0.7	no band
H	entire floxed region	4.9	1.3/4.9 (mix)	1.3	1.3

Table S4. Photophysiological phenotypes of LHC knockout lines grown under low incident irradiance as described in methods.

Data represent the average and standard deviation of biological duplicates.

Strain	σ_{PSII} (Å^2, 450nm)	σ_{PSII} (Å^2, 530nm)	N_{chl}/PSII	a_{chl} (m^2/g Chl)	Chl a/TOC (mg/g)
Wild Type	402 ± 2	163 ± 1	304 ± 9	4. 0± 0.1	79 ± 1
LHC-3X$^-$	374 ± 1	143 ± 1	255 ± 3	4.1 ± 0.1	66 ± 1
LHC-3X$^-$/VCP-4X$^-$	197 ± 2	92 ± 1	154 ± 6	4.9 ± 0.2	45 ± 1

Table S5. Description of DNA construct elements.

DNA element abbreviation	Function	Source/Description
BSD	Blasticidin resistance gene	codon optimized for N. gaditana
HygR	Hygromycin resistance gene	codon optimized for N. gaditana
Ble	Bleomycin/Zeocin resistance gene	codon optimized for N. gaditana
CAS9	CAS9 harboring N-terminal NLS and Flag epitope tag	codon optimized for N. gaditana
CRE	Cre recombinase harboring N-terminal NLS and intron from N.gaditana	codon optimized for N. gaditana
GFP	Turbo green fluorescent protein (Evrogen)	amplified from pTurboGFP-C (Evrogen)
TCT_P	Promoter	Amplified from N. gaditana gDNA
RPL24_P	Promoter	Amplified from N. gaditana gDNA
4AIII_P	Promoter	Amplified from N. gaditana gDNA
EIF3_P	Promoter	Amplified from N. gaditana gDNA
NIR_P	Promoter (NH_4^+ repressible)	Amplified from N. gaditana gDNA
EIF3_T	Terminator	Amplified from N. gaditana gDNA
FRD_T (bi-directional)	Terminator	Amplified from N. gaditana gDNA
GNPDA_T	Terminator	Amplified from N. gaditana gDNA
NIR_T	Terminator	Amplified from N. gaditana gDNA
loxP	lox site for recombination by Cre	synthesized
lox2272	lox site for recombination by Cre	synthesized
loxN	lox site for recombination by Cre	synthesized

.GB formatted sequence for pSliceN'Excise

```
 LOCUS       pSliceN'Excise     21360 bp    DNA     circular UNA 10-SEP-2014
DEFINITION  Made by John Verruto.
COMMENT     This file is created by Vector NTI
            http://www.invitrogen.com/
FEATURES             Location/Qualifiers
     CDS             174..1340
                     /vntifkey="4"
                     /label=sopA/parA
     CDS             1340..2311
                     /vntifkey="4"
                     /label=sopB/parB
     Site            3533..3540
                     /site_type="restriction site"
                     /vntifkey="21"
                     /label="NotI site"
     misc_feature    3541..3574
                     /feature_type="lox sites"
                     /vntifkey="21"
                     /label=lox2272
     terminator      complement(3575..3892)
                     /vntifkey="43"
                     /label=EIF3_Term
                     /label="EIF3 Term"
     CDS             complement(3893..4291)
                     /vntifkey="4"
                     /label=BSD-ORF
                     /label=BSD
     promoter        complement(4313..5314)
```

```
                /vntifkey="21"
                /label="TCT_Pro (Sequence corrected)"
                /label="TCT Prom"
Site            5315..5327
                /site_type="restriction site"
                /vntifkey="21"
                /label="SfiI site"
terminator      complement(5328..5527)
                /vntifkey="43"
                /label=Term5
                /label="GPI2 Term"
CDS             complement(5528..6229)
                /vntifkey="4"
                /label=GFP
promoter        complement(6230..7051)
                /vntifkey="21"
                /label="Promoter 4AIII sequence corrected"
                /label="4AIII Prom"
misc_feature    7052..7085
                /feature_type="lox sites"
                /vntifkey="21"
                /label=lox2272
Site            7086..7093
                /site_type="restriction site"
                /vntifkey="21"
                /label=SbfI
promoter        7119..8118
                /vntifkey="29"
                /label=WE3730_RPL24_Promoter
                /label="RPL24 Prom"
CDS             8119..12369
                /vntifkey="4"
                /label=NLS-CAS9version2reducedPyStretches
                /label=CAS9
misc_feature    8122..8145
                /feature_type="Nuclear localization signals"
                /vntifkey="21"
                /label=NLS
misc_feature    8146..8169
                /feature_type="epitope tags"
                /vntifkey="21"
                /label="FLAG tag"
terminator      12370..12686
                /vntifkey="43"
                /label=Term2
                /label="FRD Term"
promoter        12712..13394
                /vntifkey="29"
                /label=nir9236_Promoter
                /label="NIR Prom"
CDS             join(13395..14315,14771..14905)
                /vntifkey="4"
                /label=CREopt3730_FullTranslation
                /label=CRE
misc_feature    13398..13421
                /feature_type="Nuclear localization signals"
                /vntifkey="21"
                /label=NLS2
misc_feature    13935..13937
                /feature_type="Catalytic Residues of CRE"
```

```
                     /label=R173
     misc_feature    14283..14285
                     /feature_type="Catalytic Residues of CRE"
                     /label=H289
     misc_feature    14292..14294
                     /feature_type="Catalytic Residues of CRE"
                     /label=R292
     intron          14316..14770
                     /vntifkey="15"
                     /label=Intron
     misc_feature    14816..14818
                     /feature_type="Nucleophillic  Residues of CRE"
                     /label=W315
     misc_feature    14843..14845
                     /feature_type="Nucleophillic  Residues of CRE"
                     /label=Y324
     terminator      14906..15895
                     /vntifkey="43"
                     /label=nir9236_Terminator-PACBioCurated
                     /label="NIR Term"
     Site            15921..15928
                     /site_type="restriction site"
                     /vntifkey="21"
                     /label="PacI site"
     Site            15954..15961
                     /site_type="restriction site"
                     /vntifkey="21"
                     /label="AscI site"
     CDS             complement(16193..16852)
                     /vntifkey="4"
                     /label=CmR
     CDS             17071..17418
                     /vntifkey="4"
                     /label=redF
     rep_origin      17813..18430
                     /vntifkey="33"
                     /label=oriV
                     /note="Unknown feature type:Replication_ori  From
                     Epicentre annotation obtained by e-mail."
     rep_origin      18431..18485
                     /vntifkey="33"
                     /label=oriS/ori2
                     /note="Unknown feature type:Replication_ori"
     CDS             18814..19569
                     /vntifkey="4"
                     /label=repE/repA
                     /note="Initiation factor for ori2"
     promoter        20489..20649
                     /vntifkey="29"
                     /label="HIS3 promoter (Struhl 1982)"
                     /note="Unknown feature type:Promoter"
     CDS             20676..21335
                     /vntifkey="4"
                     /label=HIS3
                     /note="HIS3 imidazoleglycerolphosphate dehydratase"
ORIGIN
        1 GCTAGTGATA ATAAGTGACT GAGGTATGTG CTCTTCTTAT CTCCTTTTGT AGTGTTGCTC
       61 TTATTTTAAA CAACTTTGCG GTTTTTTGAT GACTTTGCGA TTTTGTTGTT GCTTTGCAGT
      121 AAATTGCAAG ATTTAATAAA AAAACGCAAA GCAATGATTA AAGGATGTTC AGAATGAAAC
      181 TCATGGAAAC ACTTAACCAG TGCATAAACG CTGGTCATGA AATGACGAAG GCTATCGCCA
```

```
 241 TTGCACAGTT TAATGATGAC AGCCCGGAAG CGAGGAAAAT AACCCGGCGC TGGAGAATAG
 301 GTGAAGCAGC GGATTTAGTT GGGGTTTCTT CTCAGGCTAT CAGAGATGCC GAGAAAGCAG
 361 GGCGACTACC GCACCCGGAT ATGGAAATTC GAGGACGGGT TGAGCAACGT GTTGGTTATA
 421 CAATTGAACA AATTAATCAT ATGCGTGATG TGTTTGGTAC GCGATTGCGA CGTGCTGAAG
 481 ACGTATTTCC ACCGGTGATC GGGGTTGCTG CCCATAAAGG TGGCGTTTAC AAAACCTCAG
 541 TTTCTGTTCA TCTTGCTCAG GATCTGGCTC TGAAGGGGCT ACGTGTTTTG CTCGTGGAAG
 601 GTAACGACCC CCAGGGAACA GCCTCAATGT ATCACGGATG GGTACCAGAT CTTCATATTC
 661 ATGCAGAAGA CACTCTCCTG CCTTTCTATC TTGGGGAAAA GGACGATGTC ACTTATGCAA
 721 TAAAGCCCAC TTGCTGGCCG GGGCTTGACA TTATTCCTTC CTGTCTGGCT CTGCACCGTA
 781 TTGAAACTGA GTTAATGGGC AAATTTGATG AAGGTAAACT GCCCACCGAT CCACACCTGA
 841 TGCTCCGACT GGCCATTGAA ACTGTTGCTC ATGACTATGA TGTCATAGTT ATTGACAGCG
 901 CGCCTAACCT GGGTATCGGC ACGATTAATG TCGTATGTGC TGCTGATGTG CTGATTGTTC
 961 CCACGCCTGC TGAGTTGTTT GACTACACCT CCGCACTGCA GTTTTTCGAT ATGCTTCGTG
1021 ATCTGCTCAA GAACGTTGAT CTTAAAGGGT TCGAGCCTGA TGTACGTATT TTGCTTACCA
1081 AATACAGCAA TAGCAATGGC TCTCAGTCCC CGTGGATGGA GGAGCAAATT CGGGATGCCT
1141 GGGGAAGCAT GGTTCTAAAA AATGTTGTAC GTGAAACGGA TGAAGTTGGT AAAGGTCAGA
1201 TCCGGATGAG AACTGTTTTT GAACAGGCCA TTGATCAACG CTCTTCAACT GGTGCCTGGA
1261 GAAATGCTCT TTCTATTTGG GAACCTGTCT GCAATGAAAT TTTCGATCGT CTGATTAAAC
1321 CACGCTGGGA GATTAGATAA TGAAGCGTGC GCCTGTTATT CCAAAACATA CGCTCAATAC
1381 TCAACCGGTT GAAGATACTT CGTTATCGAC ACCAGCTGCC CCGATGGTGG ATTCGTTAAT
1441 TGCGCGCGTA GGAGTAATGG CTCGCGGTAA TGCCATTACT TTGCCTGTAT GTGGTCGGGA
1501 TGTGAAGTTT ACTCTTGAAG TGCTCCGGGG TGATAGTGTT GAGAAGACCT CTCGGGTATG
1561 GTCAGGTAAT GAACGTGACC AGGAGCTGCT TACTGAGGAC GCACTGGATG ATCTCATCCC
1621 TTCTTTTCTA CTGACTGGTC AACAGACACC GGCGTTCGGT CGAAGAGTAT CTGGTGTCAT
1681 AGAAATTGCC GATGGGAGTC GCCGTCGTAA AGCTGCTGCA CTTACCGAAA GTGATTATCG
1741 TGTTCTGGTT GGCGAGCTGG ATGATGAGCA GATGGCTGCA TTATCCAGAT TGGGTAACGA
1801 TTATCGCCCA ACAAGTGCTT ATGAACGTGG TCAGCGTTAT GCAAGCCGAT TGCAGAATGA
1861 ATTTGCTGGA AATATTTCTG CGCTGGCTGA TGCGGAAAAT ATTTCACGTA AGATTATTAC
1921 CCGCTGTATC AACACCGCCA AATTGCCTAA ATCAGTTGTT GCTCTTTTTT CTCACCCCGG
1981 TGAACTATCT GCCCGGTCAG GTGATGCACT TCAAAAAGCC TTTACAGATA AAGAGGAATT
2041 ACTTAAGCAG CAGGCATCTA ACCTTCATGA GCAGAAAAAA GCTGGGGTGA TATTTGAAGC
2101 TGAAGAAGTT ATCACTCTTT TAACTTCTGT GCTTAAAACG TCATCTGCAT CAAGAACTAG
2161 TTTAAGCTCA CGACATCAGT TTGCTCCTGG AGCGACAGTA TTGTATAAGG GCGATAAAAT
2221 GGTGCTTAAC CTGGACAGGT CTCGTGTTCC AACTGAGTGT ATAGAGAAAA TTGAGGCCAT
2281 TCTTAAGGAA CTTGAAAAGC CAGCACCCTG ATGCGACCAC GTTTTAGTCT ACGTTTATCT
2341 GTCTTTACTT AATGTCCTTT GTTACAGGCC AGAAAGCATA ACTGGCCTGA ATATTCTCTC
2401 TGGGCCCACT GTTCCACTTG TATCGTCGGT CTGATAATCA GACTGGGACC ACGGTCCCAC
2461 TCGTATCGTC GGTCTGATTA TTAGTCTGGG ACCACGGTCC CACTCGTATC GTCGGTCTGA
2521 TTATTAGTCT GGGACCACGG TCCCACTCGT ATCGTCGGTC TGATAATCAG ACTGGGACCA
2581 CGGTCCCACT CGTATCGTCG GTCTGATTAT TAGTCTGGGA CCATGGTCCC ACTCGTATCG
2641 TCGGTCTGAT TATTAGTCTG GGACCACGGT CCCACTCGTA TCGTCGGTCT GATTATTAGT
2701 CTGGAACCAC GGTCCCACTC GTATCGTCGG TCTGATTATT AGTCTGGGAC CACGGTCCCA
2761 CTCGTATCGT CGGTCTGATT ATTAGTCTGG GACCACGATC CCACTCGTGT TGTCGGTCTG
2821 ATTATCGGTC TGGGACCACG GTCCCACTTG TATTGTCGAT CAGACTATCA GCGTGAGACT
2881 ACGATTCCAT CAATGCCTGT CAAGGGCAAG TATTGACATG TCGTCGTAAC CTGTAGAACG
2941 GAGTAACCTC GGTGTGCGGT TGTATGCCTG CTGTGGATTG CTGCTGTGTC CTGCTTATCC
3001 ACAACATTTT GCGCACGGTT ATGTGGACAA AATACCTGGT TACCCAGGCC GTGCCGGCAC
3061 GTTAACCGGG CTGCATCCGA TGCAAGTGTG TCGCTGTCGA CGAGCTCGCG AGCTCGGACA
3121 TGAGGTTGCC CCGTATTCAG TGTCGCTGAT TTGTATTGTC TGAAGTTGTT TTTACGTTAA
3181 GTTGATGCAG ATCAATTAAT ACGATACCTG CGTCATAATT GATTATTTGA CGTGGTTTGA
3241 TGGCCTCCAC GCACGTTGTG ATATGTAGAT GATAATCATT ATCACTTTAC GGGTCCTTTC
3301 CGGTGATCCG ACAGGTTACG TGGCGGCGAC CTCGCGGGTT TTCGCTATTT ATGAAAATTT
3361 TCCTGTTTAA GGCGTTTCCG TTCTTCTTCG TCATAACTTA ATGTTTTTAT TTAAAATACC
3421 CTCTGAAAAG AAAGGAAACG ACAGGTGCTG AAAGCGAGCT TTTTGGCCTC TGTCGTTTCC
3481 TTTCTCTGTT TTTGTCCGTG GAATGAACAA TGGAAGTCCG AGCTCATCGC TAGCGGCCGC
3541 ATAACTTCGT ATAGGATACT TTATACGAAG TTATCGTATG GTCGACGGTT GCTCGGATGG
3601 GGGGGGCGGG GAGCGATGGA GGGAGGAAGA TCAGGTAAGG TCTCGACAGA CTAGAGAAGC
3661 ACGAGTGCAG GTATAAGAAA CAGCAAAAAA AAGTAATGGG CCCAGGCCTG GAGAGGGTAT
3721 TTGTCTTGTT TTTCTTTGGC CAGGAACTTG TTCTCCTTTC TTCGTTTCTA GGACCCCGAT
3781 CCCCGCTCGC ATTTCTCTCT TCCTCAGCCG AAGCGCAGCG GTAAAGCATC CATTTTATCC
```

```
3841 CACCGAAAGG GCGCTCCCAG CCTTCGTCGA GCGGAACCGG GGTTACAGTG CCTCAACCCT
3901 CCCAGACGTA GCCAGAGGGA AGCAACTCCC TGATGCCAAC CGCTGTGGGC TGCCCATCGG
3961 AATCTTTGAC AATTGCCTTG ATCCCCGGGT GCAAGTCAAG CAGCACCTGC CGACATCGCC
4021 CGCACGGAGA CAGAATGCCG CGGTTTTCGT TCCCGATGGC CACTATGCAC GTCAGATTTC
4081 CGGCAGCAGC CGCAGCGGCC GTTCCGAGGA CCACGAGCTC CGCGCATGGC CCTCCGGTGA
4141 AATGATATAC ATTCACGCCG GTAAAGATCC GACCGTCGGA CGAGAGGGCT GCACTGGCCA
4201 CCGAGTAGTC CTCGCTAATA GGTATGCTGT TGATGGTCGC AGTTGCACGT TCGATCAGCG
4261 TGGATTCCTC TTGGGATAAA GGCTTGGCCA TCGAGCTCGG TACCCGGGGA TCCATGATTG
4321 TTGTATTATG TACCTATGTT TGTGATGAGA CAATAAATAT GAGAAGAGAA CGTTGCGGCC
4381 ACTTTTTTCT CCTTCCTTCG CGTGCTCATG TTGGTGGTTT GGGAGGCAGA AGATGCATGG
4441 AGCGCCACAC ATTCGGTAGG ACGAAACAGC CTCCCCCACA AAGGGACCAT GGGTAGCTAG
4501 GATGACGCAC AAGCGAGTTC CCGCTCTCGA AGGGAAACCC AGGCATTTCC TTCCTCTTTT
4561 CAAGCCACTT GTTCACGTGT CAACACAATT TTGGACTAAA ATGCCCCTCG GAACTCGGCA
4621 GGCCTCCCTC TGCTCCGTTG TCCTGGTCGC CGAGAACGCG AGACCGTGCC GCATGCCATC
4681 GATCTGCTCG TCTGTACTAC TAATCGTGTG CGTGTTCGTG CTTGTTTCGC ACGAAATTGT
4741 CCTCGTTCGG CCCTCACAAC GGTGGAAATC GGTGCTAGAA TAAAGTGAGG TGGCTTATTT
4801 CAATGGCGGC CGTCATCATG CGGGATCAAC TGAAGTACGG CGGGTTCTCG AGATTTCATC
4861 GTGCTCGTCC AGAGCAGGTG TTTTGCCTGC AGCTCTTCAT GTTTAGGGGT CATGATTTCA
4921 TCTGATATGC CGTAAGAAAA CCAATATTCA CTTCTCAATT TTCCATGGAA AGGTGAAGGC
4981 CTAGGTTGTG TGCGAGGCAA CGACTGGGGA GGGATCGCAA CATTCTTGCT AACCTCCCCT
5041 CTATCTTGGC CGCTGTGAAT CGGCATATTT ACCGGGCTGA ATTGAGAAAG TGTTTTGAGG
5101 GAATTAAAAG GTGGCTGTCT TGCAAGCTTG GCTTCAGTGC CTGCTTAATT CGAACCGATC
5161 CAGCTTGTGA TGAGGCCTTC CTAAGCCTGG TAGTCAGAAG CGACATGGCG CTATAAATTT
5221 CGTCTCAGTT GGAGAGTAGA AAAGCATGAT TCGAACACGG TTTTCAACTG CCAAAGATAT
5281 CTCCATTGTT TCCTTCAATC TGTACACCTG CACGGGCCAG TGAGGCCAGG AAATAAAGAT
5341 GGACAGACGG CATGCTAGTA GACTTTGTTG AGATTAGTGT TTGTGTTCGT CTTTATGGCT
5401 TTGAGTGGGC CCCCTTAACC TATACACACA TGACAATCAG GTGACGAGGA AGCTCTCGAC
5461 TCTCCAGGTC TCCAACACAT CATGAGGACG CCGCTCTGCC AGGACCCTCC CCGACTCCTT
5521 CCCACCCTTA TTCTTCACCG GCATCTGCAT CCGGGGTCTT GAAGGCGTGC TGGTACTCCA
5581 CGATGCCCAG CTCGGTGTTG CTGTGATCCT CCTCCACGCG GCGGAAGGCG AACATGGGGC
5641 CCCCGTTCTG CAGGATGCTG GGGTGGATGG CGCTCTTGAA GTGCATGTGG CTGTCCACCA
5701 CGGAGCTGTA GTAGCCGCCG TCGCGCAGGC TGAAGGTGCG GGTGAAGCTG CCATCCAGAT
5761 CGTTATCGCC CATGGGGTGC AGGTGCTCCA CGGTGGCGTT GCTGCGGATG ATCTTGTCGG
5821 TGAAGATCAC GCTGTCCTCG GGGAAGCCGG TGCCCATCAC CTTGAAGTCG CCGATCACGC
5881 GGCCGGCCTC GTAGCGGTAG CTGAAGCTCA CGTGCAGCAC GCCGCCGTCC TCGTACTTCT
5941 CGATGCGGGT GTTGGTGTAG CCGCCGTTGT TGATGGCGTG CAGGAAGGGG TTCTCGTAGC
6001 CGCTGGGGTA GGTGCCGAAG TGGTAGAAGC CGTAGCCCAT CACGTGGCTC AGCAGGTAGG
6061 GGCTGAAGGT CAGGGCGCCT TTGGTGCTCT TCATCTTGTT GGTCATGCGG CCCTGCTCGG
6121 GGGTGCCCTC TCCGCCGCCC ACCAGCTCGA ACTCCACGCC GTTCAGGGTG CCGGTGATGC
6181 GGCACTCGAT CTCCATGGCG GGCAGGCCGC TCTCGTCGCT CTCCAACATG TAAGCTAGGC
6241 TTTTGGTGAG AGAATGGGAA AGAAGTTAGA TGTAAAATTG AACTTCGGTT GTCGAATTTC
6301 AGAGGTAGTG CGCGGTGCGT GCGCAACGAA GGACCGTCTG CGACAGTCGG AGAGAATTGG
6361 GGTAGCCACT AGAGTAGAAA ACCTTCACTT TCCCGCCTGA GCACCGTTTC TGGAAAGGAT
6421 CTGAAGATTG AGATATGATT TTTCGAACTT GCACCGATGT GGCCCTCGTG TAGAAGACGA
6481 GGCAGAGTGG ATATAGTGCC ACTGAAGACA TGCAGCAAGC TACCGAACAA CGCGATAATG
6541 GAGACTAGCG CGTCTGCCAT TGGCAACCGT GCTCGCCTTC TCGTGATCTT ACGTGTCGCG
6601 TCTCTTCATC TCCGTACACG AAAAATATTG GTATGCGCGT GCATTATGCT TTCAGTACGT
6661 GTAAATGAGA GACAGGCAAT GCCACACTAC TGGCGCAGGA CATGTTATCC TCATCCGGGT
6721 CGCTTTTCTT GCTCTATGCA AGGAAAGGGG CGGAAATGAT AGAGATTGAT AAATTGATCG
6781 ACGCGGAAGA GTTATTACTC TGCATGACAA TGAAGTGTGC TTTTAAAGTT TTGTTTATCG
6841 AGAGGCCTCG TGCGAGAAAT TTTTGTCGCA GCATGATTGA CTTGTAGGAT AGATACTAGC
6901 TGGACTGGTC TTCGACATCC CTACACCTCC TGCCAAACGG AAAAAAAAAG CATCTGTCGG
6961 CTGCACACAG ATTGCGACTA CTTATAACTT CAAACTATGC TATAAGTGTC CTTTTCTTTC
7021 TTTCTTTTCT TTCCTTGCCG TCCTTTATGC CATAACTTCG TATAGGATAC TTTATACGAA
7081 GTTATCCTGC AGGCAGTTGG TACGGCATAT TATGGTTTAA TAAGCATACA TCATATGAAT
7141 ACAATTCAGC TTAAATTTAT CATACAAAGA TGTAAGTGCA GCGTGGGTCT GTAACGATCG
7201 GGCGTAATTT AAGATAATGC GAGGGACCGG GGGAGGTTTT GGAACGGAAT GAGGAATGGG
7261 TCATGGCCCA TAATAATAAT ATGGGTTTGG TCGCCTCGCA CAGCAACCGT ACGTGCGAAA
7321 AAGGAACAGA TCCATTTAAT AAGTTGAACG TTATTCTTTC CTATGCAATG CGTGTATCGG
7381 AGGCGAGAGC AAGTCATAGG TGGCTGCGCA CAATAATTGA GTCTCAGCTG AGCGCCGTCC
```

```
 7441 GCGGGTGGTG TGAGTGGTCA TCCTCCTCCC GGCCTATCGC TCACATCGCC TCTCAATGGT
 7501 GGTGGTGGGG CCTGATATGA CCTCAATGCC GACCCATATT AAAACCCAGT AAAGCATTCA
 7561 CCAACGAACG AGGGGCTCTT TTGTGTGTGT TTTGAGTATG ATTTTACACC TCTTTGTGCA
 7621 TCTCTCTGGT CTTCCTTGGT TCCCGTAGTT TGGGCATCAT CACTCACGCT TCCCTCGACC
 7681 TTCGTTCTTC CTTTACAACC CCGACACAGG TCAGAGTTGG AGTAATCAAA AAAGGGGTGC
 7741 ACGAATGAGA TACATTAGAT TTTGACAGAT ATCCTTTTAC TGGAGAGGGT TCAAGGGATC
 7801 AAATGAACAG CGGGCGTTGG CAATCTAGGG AGGGATCGGA GGTTGGCAGC GAGCGAAAGC
 7861 GTGTCCATCC TTTTGGCTGT CACACCTCAC GAACCAACTG TTAGCAGGCC AGCACAGATG
 7921 ACATACGAGA ATCTTTATTA TATCGTAGAC CTTATGTGGA TGACCTTTGG TGCTGTGTGT
 7981 CTGGCAATGA ACCTGAAGGC TTGATAGGGA GGTGGCTCCC GTAAACCCTT TGTCCTTTCC
 8041 ACGCTGAGTC TCCCCCGCAC TGTCCTTTAT ACAAATTGTT ACAGTCATCT GCAGGCGGTT
 8101 TTTCTTTGGC AGGCAAAGAT GCCCAAGAAA AAGCGGAAGG TCGGCGACTA CAAGGATGAC
 8161 GATGACAAGT TGGAGCCTGG AGAGAAGCCC TACAAATGCC CTGAGTGCGG AAAGAGCTTC
 8221 AGCCAATCTG GAGCCTTGAC CCGGCATCAA CGAACGCATA CACGAGACAA GAAGTACTCC
 8281 ATCGGGCTGG ACATCGGGAC GAACTCCGTG GGATGGGCCG TGATCACAGA CGAATACAAG
 8341 GTGCCTTCCA AGAAGTTCAA GGTGCTGGGG AACACGGACA GACACTCCAT CAAGAAGAAC
 8401 CTCATCGGGG CCTTGCTCTT CGACTCCGGA GAAACCGCCG AAGCAACGCG ATTGAAAAGA
 8461 ACCGCCAGAA GACGATACAC ACGACGGAAG AACCGCATCT GCTACCTCCA GGAGATCTTC
 8521 AGCAACGAGA TGGCCAAGGT GGACGACTCG TTCTTTCATC GCCTGGAGGA GAGCTTCCTG
 8581 GTGGAGGAAG ACAAGAAACA TGAGCGCCAC CCGATCTTCG GGAACATCGT GGACGAAGTG
 8641 GCCTACCACG AGAAATACCC CACGATCTAC CACTTGCGCA AGAAACTCGT GGACTCCACG
 8701 GACAAAGCGG ACTTGCGGTT GATCTACTTG GCCTTGGCCC ACATGATCAA ATTTCGGGGC
 8761 CACTTCCTGA TCGAGGGCGA CTTGAATCCC GACAATTCCG ACGTGGACAA GCTCTTCATC
 8821 CAGCTGGTGC AGACCTACAA CCAGCTCTTC GAGGAGAACC CCATCAATGC CTCCGGAGTG
 8881 GACGCCAAAG CCATCTTGTC CGCCCGATTG TCCAAATCCA GACGCTTGGA GAACTTGATC
 8941 GCACAACTTC CTGGCGAGAA GAAGAACGGC CTCTTCGGCA ACTTGATCGC GCTGTCGCTG
 9001 GGATTGACGC CTAACTTCAA GTCCAACTTC GACTTGGCCG AGGACGCCAA GTTGCAACTG
 9061 TCCAAGGACA CCTACGACGA CGACCTCGAC AACCTGCTGG CCCAAATTGG CGACCAATAC
 9121 GCGGACTTGT TTTTGGCGGC CAAGAACTTG AGCGACGCCA TCTTGTTGAG CGACATCTTG
 9181 CGCGTGAATA CGGAGATCAC CAAAGCCCCT TTGTCCGCCT CTATGATCAA GCGGTACGAC
 9241 GAGCACCACC AAGACTTGAC CCTGTTGAAA GCCCTCGTGC GGCAACAATT GCCCGAGAAG
 9301 TACAAGGAGA TCTTCTTCGA CCAGTCCAAG AACGGGTACG CCGGCTACAT CGACGGAGGA
 9361 GCCTCCCAAG AAGAGTTCTA CAAGTTCATC AAGCCCATCC TGGAGAAGAT GGACGGCACC
 9421 GAGGAGTTGC TCGTGAAGCT GAACCGCGAA GACTTGTTGC GAAAACAGCG GACGTTCGAC
 9481 AATGGCAGCA TCCCCCACCA AATCCATTTG GGAGAGTTGC ACGCCATCTT GCGACGGCAA
 9541 GAGGACTTCT ACCCGTTCCT GAAGGACAAC CGCGAGAAAA TCGAGAAGAT CCTGACGTTC
 9601 AGAATCCCCT ACTACGTGGG ACCCTTGGCC CGAGGCAATT CCCGGTTTGC ATGGATGACG
 9661 CGCAAAAGCG AAGAGACGAT CACCCCCTGG AACTTCGAAG AAGTGGTCGA CAAAGGAGCA
 9721 TCCGCACAGA GCTTCATCGA GCGAATGACG AACTTCGACA AGAACCTGCC CAACGAGAAG
 9781 GTGTTGCCCA AGCATTCGCT GCTGTACGAG TACTTCACGG TGTACAACGA GCTGACCAAG
 9841 GTGAAGTACG TGACCGAGGG CATGCGCAAA CCCGCGTTCC TGTCGGGAGA GCAAAAGAAG
 9901 GCCATTGTGG ACCTGCTGTT CAAGACCAAC CGGAAGGTGA CCGTGAAACA GCTGAAAGAG
 9961 GACTACTTCA AGAAGATCGA GTGCTTCGAC TCCGTGGAGA TCTCCGGCGT GGAGGACCGA
10021 TTCAATGCCT CCTTGGGAAC CTACCATGAC CTCCTGAAGA TCATCAAGGA CAAGGACTTC
10081 CTGGACAACG AGGAGAACGA GGACATCCTG GAGGACATCG TGCTGACCCT GACCCTGTTC
10141 GAGGACCGAG AGATGATCGA GGAACGGTTG AAAACGTACG CCCACTTGTT CGACGACAAG
10201 GTGATGAAGC AGCTGAAACG CCGCCGCTAC ACCGGATGGG GACGATTGAG CCGCAAACTG
10261 ATTAATGGAA TTCGCGACAA GCAATCCGGA AAGACCATCC TGGACTTCCT GAAGTCCGAC
10321 GGGTTCGCCA ACCGCAACTT CATGCAGCTC ATCCACGACG ACTCCTTGAC CTTCAAGGAG
10381 GACATCCAGA AGGCCCAAGT GTCCGGACAA GGAGACTCCT TGCACGAGCA CATCGCCAAT
10441 TTGGCCGGAT CCCCCGCAAT CAAAAAAGGC ATCTTGCAAA CCGTGAAAGT GGTCGACGAA
10501 CTGGTGAAGG TGATGGGACG GCACAAGCCC GAGAACATCG TGATCGAAAT GGCCCGCGAG
10561 AACCAAACCA CCCAAAAAGG ACAGAAGAAC TCCCGAGAGC GCATGAAGCG GATCGAAGAG
10621 GGCATCAAGG AGTTGGGCTC CCAGATCCTG AAGGAGCATC CCGTGGAGAA TACCCAATTG
10681 CAAAACGAGA AGCTCTACCT CTACTACCTC CAGAACGGGC GGGACATGTA CGTCGACCAA
10741 GAGCTGGACA TCAACCGCCT CTCCGACTAC GATGTGGATC ATATTGTGCC CCAGAGCTTC
10801 CTCAAGGACG ACAGCATCGA CAACAAGGTC CTGACGCGCA GCGACAAGAA CCGGGGCAAG
10861 TCTGACAATG TGCCTTCCGA AGAAGTCGTG AAGAAGATGA AGAACTACTG GCGGCAGCTG
10921 CTCAACGCCA AGCTCATCAC CCAACGGAAG TTCGACAACC TGACCAAGGC CGAGAGAGGA
10981 GGATTGTCCG AGTTGGACAA AGCCGGCTTC ATTAAACGCC AACTCGTGGA GACCCGCCAG
```

```
11041 ATCACGAAGC ACGTGGCCCA AATCTTGGAC TCCCGGATGA ACACGAAATA CGACGAGAAT
11101 GACAAGCTGA TCCGCGAGGT GAAGGTGATC ACGCTGAAGT CCAAGCTGGT GAGCGACTTC
11161 CGGAAGGACT TCCAGTTCTA CAAGGTGCGG GAGATCAACA ACTACCATCA CGCCCATGAC
11221 GCCTACCTGA ACGCCGTGGT CGGAACCGCC CTGATCAAGA AATACCCCAA GCTGGAGTCC
11281 GAATTCGTGT ACGGAGATTA CAAGGTCTAC GACGTGCGGA AGATGATCGC GAAGTCCGAG
11341 CAGGAGATCG GCAAAGCCAC CGCCAAGTAC TTCTTTTACT CCAACATCAT GAACTTCTTC
11401 AAGACCGAGA TCACGCTCGC CAACGGCGAG ATCCGCAAGC GCCCCCTGAT CGAGACCAAC
11461 GGCGAGACGG GAGAGATTGT GTGGGACAAA GGAAGAGATT TTGCCACAGT GCGCAAGGTG
11521 CTGTCCATGC CTCAGGTGAA CATCGTGAAG AAGACCGAGG TGCAAACAGG AGGGTTTTCC
11581 AAAGAGTCCA TTTTGCCTAA GAGGAATTCC GACAAGCTCA TCGCCCGCAA GAAGGACTGG
11641 GACCCCAAGA AGTACGGGGG CTTCGACTCC CCCACGGTGG CCTACTCCGT GTTGGTGGTG
11701 GCCAAAGTGG AGAAAGGGAA GAGCAAGAAG CTGAAATCCG TGAAGGAGTT GCTCGGAATC
11761 ACGATCATGG AACGATCGTC GTTCGAGAAA AACCCCATCG ACTTCCTCGA AGCCAAAGGG
11821 TACAAAGAGG TGAAGAAGGA CCTGATCATC AAGCTGCCCA AGTACTCCCT GTTCGAGCTG
11881 GAGAACGGCC GCAAGCGGAT GCTGGCCTCC GCCGGGGAAC TGCAGAAAGG GAACGAATTG
11941 GCCTTGCCCT CCAAATACGT GAACTTCCTC TACTTGGCCT CCCATTACGA AAAGCTCAAA
12001 GGATCCCCTG AGGACAATGA GCAGAAGCAA CTCTTCGTGG AACAACACAA GCACTACCTG
12061 GACGAGATCA TCGAGCAGAT CAGCGAGTTC TCCAAGCGCG TGATCCTCGC CGACGCCAAC
12121 CTGGACAAGG TGCTCTCCGC CTACAACAAG CACCGCGACA AGCCTATCCG CGAGCAAGCC
12181 GAGAATATCA TTCACCTGTT TACCCTGACG AATTTGGGAG CCCCTGCCGC CTTTAAATAC
12241 TTTGACACCA CCATCGACCG CAAAAGATAC ACCTCCACCA AGGAAGTCTT GGACGCCACC
12301 CTCATCCACC AGTCCATCAC GGGCCTCTAC GAGACGCGCA TCGACCTCTC CCAATTGGGC
12361 GGCGACTAAA GTGATGCGGC CTTTAGGAAA CACCACAAAA GTAATTGACA ATCTCAGGAA
12421 CGATCTGCGT GTTTACAGCT TCCCAAATAA CAATTATACC ACGTACCAAA AGGGGTTTAA
12481 TGTATCTCAC AAATTCTTCT AATAGGTACA GCTTCTCAAA TTGGGTGTAT GATGTGACAC
12541 TTCGTCTCAC ACACGTCACG ATAATTCAGC GTATGGCTTC CCTTCATCAC ATTCACGCAA
12601 ACTTCTACAC AACCCTGGGC ATATTTCTTG TGTTGGCAAC ACTCCCGAAA TCGATTCTGC
12661 ACACAATGGT TCATTCAATG ATTCAAGTAC GTTTTAGACG GACTAGGCAG TCTGGTGTCG
12721 TCAACAGCCA GCTGCCACAA GAAAGTGAAC ATGCGTCTAT TTATGACGTC ATTCATCAAC
12781 CACCCCGTTT CCAAACACCG TCCCACGCGC TGTTGAGAGA TGATTTTTTG AATGCCATAT
12841 GGTGCTCAAA CATGTGCATC GACGCTGTCG CACAAGCAGG AGCGGGCTTG CCCACTCGTT
12901 CTTGTTAACG GCTTGATTCA AAATCCCCGC CCGGAACAAA ATATGCCGGA GCGATCCAAC
12961 GAAGCAAAAG TCAACCAGAG CCTCTCTTTC CGTCCAACAC CCGTGTTGGT GCCATGTTAA
13021 CAATAGATTC ATGCATGGAT AGGCGAAGAC GTGAGAAGTT ACGGAGTTTG GGTCATGCTT
13081 GCGTACATCA CTCAACCCTT TTCCCCAAAA AAAAATCCCG CCATGCGATT GCCTTCGTTG
13141 CACCGCAAAA CGGAAATTAG TTATGGCGTC ATTGCTCAAG ATTACTGTTT TTCGACAAGG
13201 TGCTGCACAA CCTTGGAAGA AAACTCTGCA AATCCGTCAA TCACATGAGT TGTAGTTTTT
13261 TTCGGCAAGG CGGGTGAGCG TAGTGAATTA TATTCCTTGT AAGGCAAAGC GGATACTAAT
13321 TTTCACGTAG TTGCCCTGAC CTCCTATGCT CGGAAACGCC GCCGTACTGC CCCACCCGAA
13381 CTCAGATCAC CAGTATGCCG AAAAAGAAAC GCAAGGTGGG GTCCAACCTG TTGACGGTGC
13441 ATCAGAACTT GCCTGCCTTG CCTGTGGATG CCACATCCGA TGAAGTGCGG AAGAACCTGA
13501 TGGACATGTT CCGAGACAGA CAAGCCTTCA GCGAGCACAC CTGGAAGATG CTGCTGTCCG
13561 TGTGTAGATC TTGGGCAGCA TGGTGCAAGC TCAATAACCG GAAGTGGTTC CCAGCCGAAC
13621 CTGAGGACGT GAGAGACTAC CTGCTGTACC TGCAAGCCAG AGGATTGGCA GTGAAAACCA
13681 TCCAGCAGCA CTTGGGCCAG CTGAACATGT TGCATCGACG ATCCGGGTTG CCTAGACCTA
13741 GCGACTCTAA TGCCGTGTCT CTGGTGATGC GCCGAATCAG AAAGGAGAAC GTGGATGCCG
13801 GAGAACGGGC CAAACAAGCA TTGGCCTTTG AGCGAACCGA CTTCGACCAA GTGAGATCCT
13861 TGATGGAGAA CTCCGACCGG TGCCAAGACA TCCGGAATCT GGCGTTCTTG GGAATCGCCT
13921 ACAACACGTT GTTGAGAATA GCCGAGATCG CCCGGATCCG CGTGAAAGAC ATCTCCAGAA
13981 CAGACGGAGG ACGGATGTTG ATCCATATCG GACGGACGAA GACCCTGGTG TCTACAGCTG
14041 GAGTGGAAAA GGCCCTGTCC TTGGGAGTGA CGAAATTGGT GGAGCGATGG ATCTCCGTGT
14101 CTGGAGTGGC CGATGATCCC AACAACTACC TGTTCTGCAG AGTGCGGAAG AATGGAGTGG
14161 CAGCCCCTAG TGCCACGTCC CAATTGTCCA CAAGAGCCTT AGAGGGAATC TTCGAAGCCA
14221 CACATCGCCT GATCTACGGC GCCAAGGACG ATTCCGGACA ACGGTATTTG GCCTGGTCTG
14281 GACATTCTGC AAGAGTGGGA GCAGCCCGAG ATATGGTAAG TGTTTGCAAG AGGTCGTGCG
14341 GAGGATGAAG AGGTGCCTGA GAACGATAGA TGGAAAGGGT CGGGTGGCCT TGGTGATGGC
14401 ATTCTTTTCA GAGCTTTCCG AACACAGTCT TGTATCTGCA GTATTAATTG ATGTATGCAG
14461 TGTGTATGAT CCCACCCAGT GCCTTTATGC AGCATGGGAT TGTTAAATAG ATATGAAAGC
14521 ATAACCGGTA GAAAAGAAAG AGAGATGAGA CGCTTGGTAG AACGCCATAA TCTATGCGTT
14581 ATATGAGGAG ATACAAGCAT AGGCTGTCAC TCAATATGTA AATGGGAGAA GAAGCGTATG
```

```
14641 TTACTTGTAG ATCAGGGAGA CGTGTGGATA AAGCGCGCAG CGATTTGTCT TCCCCTCTCC
14701 GTCTCGATAC CTTTCTGCTC GGTAACAAAC TGACATGGAC TCTATCTTAT ATAAATCACA
14761 ACGTTTGTAG GCGCGCGCTG GAGTGTCCAT TCCCGAGATC ATGCAAGCTG GAGGATGGAC
14821 CAACGTGAAC ATCGTGATGA ACTACATCCG GAACCTGGAC TCCGAGACGG GAGCAATGGT
14881 GCGGCTGTTG GAAGATGGAG ATTAAAGGCA GGGTCCCCGC CAAAAAGGGT GGCGAGGACA
14941 AGAAGAAGAA ACAGGAAGGG GGGGGGCACG ACGGAGGACT TGTCGAGTCC ATCAGGGAGG
15001 TAGGGGCGGG AAGCCTCGAA TCTGCTAGTT GGTAGGGATA AATAGAGTTC AAGGACCGAA
15061 GGAGGAGGCG CCAGGATCAG CGAAAGCCTG GATTAAGAGC GAGACTCCTT GCGCTGCAGT
15121 CAAGGCGATT ACAGGACCCC CGGTGTCTGG GTTTGGAGAT GACCTCTTGG AGGACGGCTT
15181 GATGCGGGTT TTTGAGGAAG GTTGTACATT TTTGTTTGAA ATTTGCAAGG AAAGCGTCGC
15241 GCTCCGGCAT AGAGGGATAG GGGGAGGAAA GGGCACTTGT GCCCGCTCCG TCTCTGTACG
15301 GGTCTTTGAA GAAAAGATTC GAGAAACCAC CCAAAGGGCA TCAAATGCGA AACCTCCCTG
15361 AAAAAAGTTT CGATTTTCTT TATTTGTTGA GGAGGAGAGG GAAGAGTGGT ATCCAATGTG
15421 GGGTGTATTC ACGCCAACAA AGCGGGGGGA GCTGACCCAG AGGCCACCTG CCACAGGCTC
15481 CATCCAAACA AGCTTTCAGG GCTGATTCCA GAATTAGGGT TAGAGTAAGA ATGAGGGCTA
15541 CGCCAGCAGT CATCCTTTGC GGGCGTCTTG AGTCGCAAGA AGCTCTCCAA GGAAAGCGAA
15601 GGCGAATTTT CCCCAAAAAC AAAGGCAGTG GCGAGCTCCT TGTCCCTCTT TGAGCACCCC
15661 TCCTCGCTAA TTTTCTTACT CTGATTTTTT GGGGAAGTGT TTCTCCTTCT TTCGGAGACG
15721 TGGCCTTATG CTCCATCGCC TTCGCGCACC GACTCGACCA TGCCCACACA CTCTCCGTGC
15781 CCCCCTTCCC TCTGCCACCC TTCCCTCTCC CCCCCTCCCT TCCTCCCTCC CTCCCTCCCT
15841 CCCTCCCTCC CTCCTCCCAG GCACACCCCT ATTGTCCACT TCGCGCCCCA GGCTCAATAT
15901 AGGGCCGTCG TTAAGGATAC TTAATTAAAA ACATCTATCC TCCAGATCAC CAGGGCGCGC
15961 CTATGCTTGT AAACCGTTTT GTGAAAAAAT TTTTAAAATA AAAAAGGGGA CCTCTAGGGT
16021 CCCCAATTAA TTAGTAATAT AATCTATTAA AGGTCATTCA AAAGGTCATC CAGTCGACCA
16081 ATTCTCATGT TTGACAGCTT ATCATCGAAT TTCTGCCATT CATCCGCTTA TTATCACTTA
16141 TTCAGGCGTA GCAACCAGGC GTTTAAGGGC ACCAATAACT GCCTTAAAAA AATTACGCCC
16201 CGCCCTGCCA CTCATCGCAG TACTGTTGTA ATTCATTAAG CATTCTGCCG ACATGGAAGC
16261 CATCACAAAC GGCATGATGA ACCTGAATCG CCAGCGGCAT CAGCACCTTG TCGCCTTGCG
16321 TATAATATTT GCCCATGGTG AAAACGGGGG CGAAGAAGTT GTCCATATTG GCCACGTTTA
16381 AATCAAAACT GGTGAAACTC ACCCAGGGAT TGGCTGAGAC GAAAAACATA TTCTCAATAA
16441 ACCCTTTAGG GAAATAGGCC AGGTTTTCAC CGTAACACGC CACATCTTGC GAATATATGT
16501 GTAGAAACTG CCGGAAATCG TCGTGGTATT CACTCCAGAG CGATGAAAAC GTTTCAGTTT
16561 GCTCATGGAA AACGGTGTAA CAAGGGTGAA CACTATCCCA TATCACCAGC TCACCGTCTT
16621 TCATTGCCAT ACGAAATTCC GGATGAGCAT TCATCAGGCG GGCAAGAATG TGAATAAAGG
16681 CCGGATAAAA CTTGTGCTTA TTTTTCTTTA CGGTCTTTAA AAAGGCCGTA ATATCCAGCT
16741 GAACGGTCTG GTTATAGGTA CATTGAGCAA CTGACTGAAA TGCCTCAAAA TGTTCTTTAC
16801 GATGCCATTG GGATATATCA ACGGTGGTAT ATCCAGTGAT TTTTTTCTCC ATTTTAGCTT
16861 CCTTAGCTCC TGAAAATCTC GATAACTCAA AAAATACGCC CGGTAGTGAT CTTATTTCAT
16921 TATGGTGAAA GTTGGAACCT CTTACGTGCC GATCAACGTC TCATTTTCGC CAAAAGTTGG
16981 CCCAGGGCTT CCCGGTATCA ACAGGGACAC CAGGATTTAT TTATTCTGCG AAGTGATCTT
17041 CCGTCACAGG TATTTATTCG CGATAAGCTC ATGGAGCGGC GTAACCGTCG CACAGGAAGG
17101 ACAGAGAAAG CGCGGATCTG GGAAGTGACG GACAGAACGG TCAGGACCTG GATTGGGGAG
17161 GCGGTTGCCG CCGCTGCTGC TGACGGTGTG ACGTTCTCTG TTCCGGTCAC ACCACATACG
17221 TTCCGCCATT CCTATGCGAT GCACATGCTG TATGCCGGTA TACCGCTGAA AGTTCTGCAA
17281 AGCCTGATGG GACATAAGTC CATCAGTTCA ACGGAAGTCT ACACGAAGGT TTTTGCGCTG
17341 GATGTGGCTG CCCGGCACCG GGTGCAGTTT GCGATGCCGG AGTCTGATGC GGTTGCGATG
17401 CTGAAACAAT TATCCTGAGA ATAAATGCCT TGGCCTTTAT ATGGAAATGT GGAACTGAGT
17461 GGATATGCTG TTTTTGTCTG TTAAACAGAG AAGCTGGCTG TTATCCACTG AGAAGCGAAC
17521 GAAACAGTCG GGAAAATCTC CCATTATCGT AGAGATCCGC ATTATTAATC TCAGGAGCCT
17581 GTGTAGCGTT TATAGGAAGT AGTGTTCTGT CATGATGCCT GCAAGCGGTA ACGAAAACGA
17641 TTTGAATATG CCTTCAGGAA CAATAGAAAT CTTCGTGCGG TGTTACGTTG AAGTGGAGCG
17701 GATTATGTCA GCAATGGACA GAACAACCTA ATGAACACAG AACCATGATG TGGTCTGTCC
17761 TTTTACAGCC AGTAGTGCTC GCCGCAGTCG AGCGACAGGG CGAAGCCCTC GAGCTGGTTG
17821 CCCTCGCCGC TGGGCTGGCG GCCGTCTATG GCCCTGCAAA CGCGCCAGAA ACGCCGTCGA
17881 AGCCGTGTGC GAGACACCGC GGCCGGCCGC CGGCGTTGTG GATACCTCGC GGAAAACTTG
17941 GCCCTCACTG ACAGATGAGG GGCGGACGTT GACACTTGAG GGGCCGACTC ACCCGGCGCG
18001 GCGTTGACAG ATGAGGGGCA GGCTCGATTT CGGCCGGCGA CGTGGAGCTG GCCAGCCTCG
18061 CAAATCGGCG AAAACGCCTG ATTTTACGCG AGTTTCCCAC AGATGATGTG GACAAGCCTG
18121 GGGATAAGTG CCCTGCGGTA TTGACACTTG AGGGGCGCGA CTACTGACAG ATGAGGGGCG
18181 CGATCCTTGA CACTTGAGGG GCAGAGTGCT GACAGATGAG GGGCGCACCT ATTGACATTT
```

```
    18241 GAGGGGCTGT CCACAGGCAG AAAATCCAGC ATTTGCAAGG GTTTCCGCCC GTTTTTCGGC
    18301 CACCGCTAAC CTGTCTTTTA ACCTGCTTTT AAACCAATAT TTATAAACCT TGTTTTTAAC
    18361 CAGGGCTGCG CCCTGTGCGC GTGACCGCGC ACGCCGAAGG GGGGTGCCCC CCCTTCTCGA
    18421 ACCCTCCCGG TCGAGTGAGC GAGGAAGCAC CAGGGAACAG CACTTATATA TTCTGCTTAC
    18481 ACACGATGCC TGAAAAAACT TCCCTTGGGG TTATCCACTT ATCCACGGGG ATATTTTTAT
    18541 AATTATTTTT TTTATAGTTT TTAGATCTTC TTTTTTAGAG CGCCTTGTAG GCCTTTATCC
    18601 ATGCTGGTTC TAGAGAAGGT GTTGTGACAA ATTGCCCTTT CAGTGTGACA AATCACCCTC
    18661 AAATGACAGT CCTGTCTGTG ACAAATTGCC CTTAACCCTG TGACAAATTG CCCTCAGAAG
    18721 AAGCTGTTTT TTCACAAAGT TATCCCTGCT TATTGACTCT TTTTTATTTA GTGTGACAAT
    18781 CTAAAAACTT GTCACACTTC ACATGGATCT GTCATGGCGG AAACAGCGGT TATCAATCAC
    18841 AAGAAACGTA AAAATAGCCC GCGAATCGTC CAGTCAAACG ACCTCACTGA GGCGGCATAT
    18901 AGTCTCTCCC GGGATCAAAA ACGTATGCTG TATCTGTTCG TTGACCAGAT CAGAAAATCT
    18961 GATGGCACCC TACAGGAACA TGACGGTATC TGCGAGATCC ATGTTGCTAA ATATGCTGAA
    19021 ATATTCGGAT TGACCTCTGC GGAAGCCAGT AAGGATATAC GGCAGGCATT GAAGAGTTTC
    19081 GCGGGGAAGG AAGTGGTTTT TTATCGCCCT GAAGAGGATG CCGGCGATGA AAAAGGCTAT
    19141 GAATCTTTTC CTTGGTTTAT CAAACGTGCG CACAGTCCAT CCAGAGGGCT TTACAGTGTA
    19201 CATATCAACC CATATCTCAT TCCCTTCTTT ATCGGGTTAC AGAACCGGTT TACGCAGTTT
    19261 CGGCTTAGTG AAACAAAAGA AATCACCAAT CCGTATGCCA TGCGTTTATA CGAATCCCTG
    19321 TGTCAGTATC GTAAGCCGGA TGGCTCAGGC ATCGTCTCTC TGAAAATCGA CTGGATCATA
    19381 GAGCGTTACC AGCTGCCTCA AAGTTACCAG CGTATGCCTG ACTTCCGCCG CCGCTTCCTG
    19441 CAGGTCTGTG TTAATGAGAT CAACAGCAGA ACTCCAATGC GCCTCTCATA CATTGAGAAA
    19501 AAGAAAGGCC GCCAGACGAC TCATATCGTA TTTTCCTTCC GCGATATCAC TTCCATGACG
    19561 ACAGGATAGT CTGAGGGTTA TCTGTCACAG ATTTGAGGGT GGTTCGTCAC ATTTGTTCTG
    19621 ACCTACTGAG GGTAATTTGT CACAGTTTTG CTGTTTCCTT CAGCCTGCAT GGATTTTCTC
    19681 ATACTTTTTG AACTGTAATT TTTAAGGAAG CCAAATTTGA GGGCAGTTTG TCACAGTTGA
    19741 TTTCCTTCTC TTTCCCTTCG TCATGTGACC TGATATCGGG GGTTAGTTCG TCATCATTGA
    19801 TGAGGGTTGA TTATCACAGT TTATTACTCT GAATTGGCTA TCCGCGTGTG TACCTCTACC
    19861 TGGAGTTTTT CCCACGGTGG ATATTTCTTC TTGCGCTGAG CGTAAGAGCT ATCTGACAGA
    19921 ACAGTTCTTC TTTGCTTCCT CGCCAGTTCG CTCGCTATGC TCGGTTACAC GGCTGCGGCG
    19981 AGCATCACGT GCTATAAAAA TAATTATAAT TTAAATTTTT TAATATAAAT ATATAAATTA
    20041 AAAATAGAAA GTAAAAAAAG AAATTAAAGA AAAAATAGTT TTTGTTTTCC GAAGATGTAA
    20101 AAGACTCTAG GGGGATCGCC AACAAATACT ACCTTTTATC TTGCTCTTCC TGCTCTCAGG
    20161 TATTAATGCC GAATTGTTTC ATCTTGTCTG TGTAGAAGAC CACACACGAA AATCCTGTGA
    20221 TTTTACATTT TACTTATCGT TAATCGAATG TATATCTATT TAATCTGCTT TTCTTGTCTA
    20281 ATAAATATAT ATGTAAAGTA CGCTTTTTGT TGAAATTTTT TAAACCTTTG TTTATTTTTT
    20341 TTTCTTCATT CCGTAACTCT TCTACCTTCT TTATTTACTT TCTAAAATCC AAATACAAAA
    20401 CATAAAAATA AATAAACACA GAGTAAATTC CCAAATTATT CCATCATTAA AAGATACGAG
    20461 GCGCGTGTAA GTTACAGGCA AGCGATCCTA GTACACTCTA TATTTTTTTA TGCCTCGGTA
    20521 ATGATTTTCA TTTTTTTTTT TCCACCTAGC GGATGACTCT TTTTTTTTCT TAGCGATTGG
    20581 CATTATCACA TAATGAATTA TACATTATAT AAAGTAATGT GATTTCTTCG AAGAATATAC
    20641 TAAAAAATGA GCAGGCAAGA TAAACGAAGG CAAAGATGAC AGAGCAGAAA GCCCTAGTAA
    20701 AGCGTATTAC AAATGAAACC AAGATTCAGA TTGCGATCTC TTTAAAGGGT GGTCCCCTAG
    20761 CGATAGAGCA CTCGATCTTC CCAGAAAAAG AGGCAGAAGC AGTAGCAGAA CAGGCCACAC
    20821 AATCGCAAGT GATTAACGTC CACACAGGTA TAGGGTTTCT GGACCATATG ATACATGCTC
    20881 TGGCCAAGCA TTCCGGCTGG TCGCTAATCG TTGAGTGCAT TGGTGACTTA CACATAGACG
    20941 ACCATCACAC CACTGAAGAC TGCGGGATTG CTCTCGGTCA AGCTTTTAAA GAGGCCCTAC
    21001 TGGCGCGTGG AGTAAAAAGG TTTGGATCAG GATTTGCGCC TTTGGATGAG GCACTTTCCA
    21061 GAGCGGTGGT AGATCTTTCG AACAGGCCGT ACGCAGTTGT CGAACTTGGT TTGCAAAGGG
    21121 AGAAAGTAGG AGATCTCTCT TGCGAGATGA TCCCGCATTT TCTTGAAAGC TTTGCAGAGG
    21181 CTAGCAGAAT TACCCTCCAC GTTGATTGTC TGCGAGGCAA GAATGATCAT CACCGTAGTG
    21241 AGAGTGCGTT CAAGGCTCTT GCGGTTGCCA TAAGAGAAGC CACCTCGCCC AATGGTACCA
    21301 ACGATGTTCC CTCCACCAAA GGTGTTCTTA TGTAGTTTTA CACAGGAGTC TGGACTTGAC
//
```

.GB formatted Sequence for loxP-BSD-GFP-loxP fragment

```
LOCUS       loxP-BSD-GFP-loxP      3711 bp    DNA     linear   UNA 9-APR-2015
DEFINITION  Complementary copy of pUC19.
KEYWORDS    ATCC.
COMMENT     This file is created by Vector NTI
            http://www.invitrogen.com/
FEATURES             Location/Qualifiers
     misc_feature    16..41
                     /vntifkey="21"
                     /label=3frameStopFiller1
                     /label="3 frame stops"
     misc_feature    50..83
                     /feature_type="lox sites"
                     /vntifkey="21"
                     /label=loxP
     terminator      complement(117..434)
                     /vntifkey="43"
                     /label=EIF3_Term
                     /label=EIF3_T
     CDS             complement(435..833)
                     /vntifkey="4"
                     /label=BSD-ORF
                     /label=BSD
     promoter        complement(855..1856)
                     /vntifkey="21"
                     /label="TCT_Pro (Sequence corrected)"
                     /label=TCT_P
     terminator      complement(1870..2069)
                     /vntifkey="43"
                     /label=Term5
                     /label=GNPDA_T
     CDS             complement(2070..2771)
                     /vntifkey="4"
                     /label=GFP
     promoter        complement(2772..3593)
                     /vntifkey="21"
                     /label="Promoter 4AIII sequence corrected"
                     /label=4AIII_P
     misc_feature    3627..3660
                     /feature_type="lox sites"
                     /vntifkey="21"
                     /label=loxP
     misc_feature    complement(3669..3690)
                     /vntifkey="21"
                     /label=3frameStopFiller2rc
                     /label="3 frame stops"
ORIGIN
        1 TCCACAGCCC GAACCCCTTA AGCTAGACGA ACACAGTTAG CGCGGCCGCA TAACTTCGTA
       61 TAGCATACAT TATACGAAGT TATGATGCTA GCGTGTTTAA GAAGTCACTT AATTAACGTA
      121 TGGTCGACGG TTGCTCGGAT GGGGGGGGCG GGGAGCGATG GAGGGAGGAA GATCAGGTAA
      181 GGTCTCGACA GACTAGAGAA GCACGAGTGC AGGTATAAGA AACAGCAAAA AAAAGTAATG
      241 GGCCCAGGCC TGGAGAGGGT ATTTGTCTTG TTTTTCTTTG GCCAGGAACT TGTTCTCCTT
      301 TCTTCGTTTC TAGGACCCCG ATCCCCGCTC GCATTTCTCT CTTCCTCAGC CGAAGCGCAG
      361 CGGTAAAGCA TCCATTTTAT CCCACCGAAA GGGCGCTCCC AGCCTTCGTC GAGCGGAACC
      421 GGGGTTACAG TGCCTCAACC CTCCCAGACG TAGCCAGAGG GAAGCAACTC CCTGATGCCA
      481 ACCGCTGTGG GCTGCCCATC GGAATCTTTG ACAATTGCCT TGATCCCCGG GTGCAAGTCA
      541 AGCAGCACCT GCCGACATCG CCCGCACGGA GACAGAATGC CGCGGTTTTC GTTCCCGATG
      601 GCCACTATGC ACGTCAGATT TCCGGCAGCA GCCGCAGCGG CCGTTCCGAG GACCACGAGC
      661 TCCGCGCATG GCCCTCCGGT GAAATGATAT ACATTCACGC CGGTAAAGAT CCGACCGTCG
      721 GACGAGAGGG CTGCACTGGC CACCGAGTAG TCCTCGCTAA TAGGTATGCT GTTGATGGTC
      781 GCAGTTGCAC GTTCGATCAG CGTGGATTCC TCTTGGGATA AAGGCTTGGC CATCGAGCTC
```

```
      841 GGTACCCGGG GATCCATGAT TGTTGTATTA TGTACCTATG TTTGTGATGA GACAATAAAT
      901 ATGAGAAGAG AACGTTGCGG CCACTTTTTT CTCCTTCCTT CGCGTGCTCA TGTTGGTGGT
      961 TTGGGAGGCA GAAGATGCAT GGAGCGCCAC ACATTCGGTA GGACGAAACA GCCTCCCCCA
     1021 CAAAGGGACC ATGGGTAGCT AGGATGACGC ACAAGCGAGT TCCCGCTCTC GAAGGGAAAC
     1081 CCAGGCATTT CCTTCCTCTT TTCAAGCCAC TTGTTCACGT GTCAACACAA TTTTGGACTA
     1141 AAATGCCCCT CGGAACTCGG CAGGCCTCCC TCTGCTCCGT TGTCCTGGTC GCCGAGAACG
     1201 CGAGACCGTG CCGCATGCCA TCGATCTGCT CGTCTGTACT ACTAATCGTG TGCGTGTTCG
     1261 TGCTTGTTTC GCACGAAATT GTCCTCGTTC GGCCCTCACA ACGGTGGAAA TCGGTGCTAG
     1321 AATAAAGTGA GGTGGCTTAT TTCAATGGCG GCCGTCATCA TGCGGGATCA ACTGAAGTAC
     1381 GGCGGGTTCT CGAGATTTCA TCGTGCTCGT CCAGAGCAGG TGTTTTGCCT GCAGCTCTTC
     1441 ATGTTTAGGG GTCATGATTT CATCTGATAT GCCGTAAGAA AACCAATATT CACTTCTCAA
     1501 TTTTCCATGG AAAGGTGAAG GCCTAGGTTG TGTGCGAGGC AACGACTGGG GAGGGATCGC
     1561 AACATTCTTG CTAACCTCCC CTCTATCTTG GCCGCTGTGA ATCGGCATAT TTACCGGGCT
     1621 GAATTGAGAA AGTGTTTTGA GGGAATTAAA AGGTGGCTGT CTTGCAAGCT TGGCTTCAGT
     1681 GCCTGCTTAA TTCGAACCGA TCCAGCTTGT GATGAGGCCT TCCTAAGCCT GGTAGTCAGA
     1741 AGCGACATGG CGCTATAAAT TTCGTCTCAG TTGGAGAGTA GAAAAGCATG ATTCGAACAC
     1801 GGTTTTCAAC TGCCAAAGAT ATCTCCATTG TTTCCTTCAA TCTGTACACC TGCACGGGCC
     1861 AGTGAGGCCA GGAAATAAAG ATGGACAGAC GGCATGCTAG TAGACTTTGT TGAGATTAGT
     1921 GTTTGTGTTC GTCTTTATGG CTTTGAGTGG GCCCCCTTAA CCTATACACA CATGACAATC
     1981 AGGTGACGAG GAAGCTCTCG ACTCTCCAGG TCTCCAACAC ATCATGAGGA CGCCGCTCTG
     2041 CCAGGACCCT CCCCGACTCC TTCCCACCCT TATTCTTCAC CGGCATCTGC ATCCGGGGTC
     2101 TTGAAGGCGT GCTGGTACTC CACGATGCCC AGCTCGGTGT TGCTGTGATC CTCCTCCACG
     2161 CGGCGGAAGG CGAACATGGG GCCCCCGTTC TGCAGGATGC TGGGGTGGAT GGCGCTCTTG
     2221 AAGTGCATGT GGCTGTCCAC CACGGAGCTG TAGTAGCCGC CGTCGCGCAG GCTGAAGGTG
     2281 CGGGTGAAGC TGCCATCCAG ATCGTTATCG CCCATGGGGT GCAGGTGCTC CACGGTGGCG
     2341 TTGCTGCGGA TGATCTTGTC GGTGAAGATC ACGCTGTCCT CGGGGAAGCC GGTGCCCATC
     2401 ACCTTGAAGT CGCCGATCAC GCGGCCGGCC TCGTAGCGGT AGCTGAAGCT CACGTGCAGC
     2461 ACGCCGCCGT CCTCGTACTT CTCGATGCGG GTGTTGGTGT AGCCGCCGTT GTTGATGGCG
     2521 TGCAGGAAGG GGTTCTCGTA GCCGCTGGGG TAGGTGCCGA AGTGGTAGAA GCCGTAGCCC
     2581 ATCACGTGGC TCAGCAGGTA GGGGCTGAAG GTCAGGGCGC CTTTGGTGCT CTTCATCTTG
     2641 TTGGTCATGC GGCCCTGCTC GGGGGTGCCC TCTCCGCCGC CCACCAGCTC GAACTCCACG
     2701 CCGTTCAGGG TGCCGGTGAT GCGGCACTCG ATCTCCATGG CGGGCAGGCC GCTCTCGTCG
     2761 CTCTCCAACA TGTAAGCTAG GCTTTTGGTG AGAGAATGGG AAAGAAGTTA GATGTAAAAT
     2821 TGAACTTCGG TTGTCGAATT TCAGAGGTAG TGCGCGGTGC GTGCGCAACG AAGGACCGTC
     2881 TGCGACAGTC GGAGAGAATT GGGGTAGCCA CTAGAGTAGA AAACCTTCAC TTTCCCGCCT
     2941 GAGCACCGTT TCTGGAAAGG ATCTGAAGAT TGAGATATGA TTTTTCGAAC TTGCACCGAT
     3001 GTGGCCCTCG TGTAGAAGAC GAGGCAGAGT GGATATAGTG CCACTGAAGA CATGCAGCAA
     3061 GCTACCGAAC AACGCGATAA TGGAGACTAG CGCGTCTGCC ATTGGCAACC GTGCTCGCCT
     3121 TCTCGTGATC TTACGTGTCG CGTCTCTTCA TCTCCGTACA CGAAAAATAT TGGTATGCGC
     3181 GTGCATTATG CTTTCAGTAC GTGTAAATGA GAGACAGGCA ATGCCACACT ACTGGCGCAG
     3241 GACATGTTAT CCTCATCCGG GTCGCTTTTC TTGCTCTATG CAAGGAAAGG GGCGGAAATG
     3301 ATAGAGATTG ATAAATTGAT CGACGCGGAA GAGTTATTAC TCTGCATGAC AATGAAGTGT
     3361 GCTTTTAAAG TTTTGTTTAT CGAGAGGCCT CGTGCGAGAA ATTTTTGTCG CAGCATGATT
     3421 GACTTGTAGG ATAGATACTA GCTGGACTGG TCTTCGACAT CCCTACACCT CCTGCCAAAC
     3481 GGAAAAAAAA AGCATCTGTC GGCTGCACAC AGATTGCGAC TACTTATAAC TTCAAACTAT
     3541 GCTATAAGTG TCCTTTTCTT TCTTTCTTTT CTTTCCTTGC CGTCCTTTAT GCCCCTGCAG
     3601 GGTACGTTTT AGACGGACTA GGCAGTATAA CTTCGTATAG CATACATTAT ACGAAGTTAT
     3661 GGCGCGCCAG GCTACGTTAG TTCAGCAGCT GAGAACGACC ACGAACGGGA A
//
```

.GB formatted Sequence for loxN-HYGR-GFP-loxN fragment

```
LOCUS       loxN-HYGR-GFP-loxN      4341 bp    DNA     linear   UNA 26-FEB-2015
DEFINITION  Complementary copy of pUC19.
KEYWORDS    ATCC.
COMMENT     This file is created by Vector NTI
            http://www.invitrogen.com/
FEATURES             Location/Qualifiers
     misc_feature    16..41
                     /vntifkey="21"
```

```
                     /label=3frameStopFiller1
                     /label="3 frame stops"
     misc_feature    50..83
                     /feature_type="lox sites"
                     /vntifkey="21"
                     /label=loxN
     terminator      complement(117..434)
                     /vntifkey="43"
                     /label=EIF3_Term
                     /label=EIF3_T
     CDS             complement(435..1463)
                     /vntifkey="4"
                     /label=Hygro
                     /label=HygR
     promoter        complement(1485..2486)
                     /vntifkey="21"
                     /label="TCT_Pro (Sequence corrected)"
                     /label=TCT_P
     terminator      complement(2500..2699)
                     /vntifkey="43"
                     /label=Term5
                     /label=GNPDA_T
     CDS             complement(2700..3401)
                     /vntifkey="4"
                     /label=GFP
     promoter        complement(3402..4223)
                     /vntifkey="21"
                     /label="Promoter 4AIII sequence corrected"
                     /label=4AIII_P
     misc_feature    4257..4290
                     /feature_type="lox sites"
                     /vntifkey="21"
                     /label=loxN
     misc_feature    complement(4299..4320)
                     /vntifkey="21"
                     /label=3frameStopFiller2rc
                     /label="3 frame stops"
ORIGIN
        1 TCCACAGCCC GAACCCCTTA AGCTAGACGA ACACAGTTAG CGCGGCCGCA TAACTTCGTA
       61 TAGTATACCT TATACGAAGT TATGATGCTA GCGTGTTTAA GAAGTCACTT AATTAACGTA
      121 TGGTCGACGG TTGCTCGGAT GGGGGGGGCG GGGAGCGATG GAGGGAGGAA GATCAGGTAA
      181 GGTCTCGACA GACTAGAGAA GCACGAGTGC AGGTATAAGA AACAGCAAAA AAAAGTAATG
      241 GGCCCAGGCC TGGAGAGGGT ATTTGTCTTG TTTTTCTTTG GCCAGGAACT TGTTCTCCTT
      301 TCTTCGTTTC TAGGACCCCG ATCCCCGCTC GCATTTCTCT CTTCCTCAGC CGAAGCGCAG
      361 CGGTAAAGCA TCCATTTTAT CCCACCGAAA GGGCGCTCCC AGCCTTCGTC GAGCGGAACC
      421 GGGGTTACAG TGCCTCACTC CTTTGCACGC GGTCGGGTGC TCGGCCTACG GTTGCCCGAG
      481 TCCGCAAGCA CCTCAACACA GCCGTCTGTC CACACCGCAG CCGACCGGCG TGCGATTTGG
      541 GTCCGACCCA CCGTCCCAGC CCCGCTGCGG ACTATCGCGT CGCAGCGGCC CTGCGCCCAC
      601 GCGGCGTCGT CGAAGTTGCC GTCGACGAGA GACTGGTAAA GCTGATCGAG TCCGATACGC
      661 AACATATAGG CGCGGAGTCG TGGGGAGCCG GCCAGCTCCG GGTGCCTCCG TTCAAAGTAG
      721 CGTGTCTGCT GCTCCATGCA CGCCAACCAG GGACGCCAGA AGAATATGTT CGCCACTTCG
      781 TATTGGCTAT CACCAAACAT CGCTTCGGAC CAGTCGATGA CAGCAGTAAT CCGACCATTG
      841 TCTGTAAGTA CGTTATTGCT GCCGAAATCC GCGTGCACCA GGTGCCTGAC CTCAGGGCAA
      901 TCCTCGGCCC ACAACATGAG TTCGTCCAGT GCTTGGGCCA CGGATGCAGA CACGGTGTCA
      961 TCCATGACTG TCTGCCAATG ATAGACGTGA GGATCGGCAA TGGCGCAGAT GAAGTCTCGC
     1021 CAGGTCGTGT ACTGCCCGAT GCCCTGGGGC CCAAAAGGTC CAAAGCCGGA CGTCTGAGAC
     1081 AGATCTGCGG CAGCGATCGC GTCCATGGCC TCGGCCACGG GTTGCAAAAC GGCAGGCAAT
     1141 TCAGTTTCGG GCAGATCTTG CAACGTCACT CCCTGGGCTC GGCGCGAGAT GCAGTACGTG
     1201 AGAGATTCGC TAAACTCCCC AATGTCCAGT ACCTCTGGTA TGGGGAGAGC GGCGGAGGCG
     1261 AAATGACGGT AGACATACCG ATCCTTGTAG AACCCGTCCG CACAACTATT AACCCTCAAC
```

```
     1321 ACGTATCCCC GACCCCCTAC GTCAAACGAG AACGCCCTAC TCTCCTCTCC CTCGCTCAGT
     1381 TGCATCAAGT CGGAGACAGA GTCGAACTTC TCAATAAGGA ATTTCTCCAC GGACGTAGCG
     1441 GTCAGTTCCG GTTTCTTCCC CATCGAGCTC GGTACCCGGG GATCCATGAT TGTTGTATTA
     1501 TGTACCTATG TTTGTGATGA GACAATAAAT ATGAGAAGAG AACGTTGCGG CCACTTTTTT
     1561 CTCCTTCCTT CGCGTGCTCA TGTTGGTGGT TTGGGAGGCA GAAGATGCAT GGAGCGCCAC
     1621 ACATTCGGTA GGACGAAACA GCCTCCCCCA CAAAGGGACC ATGGGTAGCT AGGATGACGC
     1681 ACAAGCGAGT TCCCGCTCTC GAAGGGAAAC CCAGGCATTT CCTTCCTCTT TTCAAGCCAC
     1741 TTGTTCACGT GTCAACACAA TTTTGGACTA AAATGCCCCT CGGAACTCGG CAGGCCTCCC
     1801 TCTGCTCCGT TGTCCTGGTC GCCGAGAACG CGAGACCGTG CCGCATGCCA TCGATCTGCT
     1861 CGTCTGTACT ACTAATCGTG TGCGTGTTCG TGCTTGTTTC GCACGAAATT GTCCTCGTTC
     1921 GGCCCTCACA ACGGTGGAAA TCGGTGCTAG AATAAAGTGA GGTGGCTTAT TTCAATGGCG
     1981 GCCGTCATCA TGCGGGATCA ACTGAAGTAC GGCGGGTTCT CGAGATTTCA TCGTGCTCGT
     2041 CCAGAGCAGG TGTTTTGCCT GCAGCTCTTC ATGTTTAGGG GTCATGATTT CATCTGATAT
     2101 GCCGTAAGAA AACCAATATT CACTTCTCAA TTTTCCATGG AAAGGTGAAG GCCTAGGTTG
     2161 TGTGCGAGGC AACGACTGGG GAGGGATCGC AACATTCTTG CTAACCTCCC CTCTATCTTG
     2221 GCCGCTGTGA ATCGGCATAT TTACCGGGCT GAATTGAGAA AGTGTTTTGA GGGAATTAAA
     2281 AGGTGGCTGT CTTGCAAGCT TGGCTTCAGT GCCTGCTTAA TTCGAACCGA TCCAGCTTGT
     2341 GATGAGGCCT TCCTAAGCCT GGTAGTCAGA AGCGACATGG CGCTATAAAT TTCGTCTCAG
     2401 TTGGAGAGTA GAAAAGCATG ATTCGAACAC GGTTTTCAAC TGCCAAAGAT ATCTCCATTG
     2461 TTTCCTTCAA TCTGTACACC TGCACGGGCC AGTGAGGCCA GGAAATAAAG ATGGACAGAC
     2521 GGCATGCTAG TAGACTTTGT TGAGATTAGT GTTTGTGTTC GTCTTTATGG CTTTGAGTGG
     2581 GCCCCCTTAA CCTATACACA CATGACAATC AGGTGACGAG GAAGCTCTCG ACTCTCCAGG
     2641 TCTCCAACAC ATCATGAGGA CGCCGCTCTG CCAGGACCCT CCCCGACTCC TTCCCACCCT
     2701 TATTCTTCAC CGGCATCTGC ATCCGGGGTC TTGAAGGCGT GCTGGTACTC CACGATGCCC
     2761 AGCTCGGTGT TGCTGTGATC CTCCTCCACG CGGCGGAAGG CGAACATGGG GCCCCCGTTC
     2821 TGCAGGATGC TGGGGTGGAT GGCGCTCTTG AAGTGCATGT GGCTGTCCAC CACGGAGCTG
     2881 TAGTAGCCGC CGTCGCGCAG GCTGAAGGTG CGGGTGAAGC TGCCATCCAG ATCGTTATCG
     2941 CCCATGGGGT GCAGGTGCTC CACGGTGGCG TTGCTGCGGA TGATCTTGTC GGTGAAGATC
     3001 ACGCTGTCCT CGGGGAAGCC GGTGCCCATC ACCTTGAAGT CGCCGATCAC GCGGCCGGCC
     3061 TCGTAGCGGT AGCTGAAGCT CACGTGCAGC ACGCCGCCGT CCTCGTACTT CTCGATGCGG
     3121 GTGTTGGTGT AGCCGCCGTT GTTGATGGCG TGCAGGAAGG GGTTCTCGTA GCCGCTGGGG
     3181 TAGGTGCCGA AGTGGTAGAA GCCGTAGCCC ATCACGTGGC TCAGCAGGTA GGGGCTGAAG
     3241 GTCAGGGCGC CTTTGGTGCT CTTCATCTTG TTGGTCATGC GGCCCTGCTC GGGGGTGCCC
     3301 TCTCCGCCGC CCACCAGCTC GAACTCCACG CCGTTCAGGG TGCCGGTGAT GCGGCACTCG
     3361 ATCTCCATGG CGGGCAGGCC GCTCTCGTCG CTCTCCAACA TGTAAGCTAG GCTTTTGGTG
     3421 AGAGAATGGG AAAGAAGTTA GATGTAAAAT TGAACTTCGG TTGTCGAATT TCAGAGGTAG
     3481 TGCGCGGTGC GTGCGCAACG AAGGACCGTC TGCGACAGTC GGAGAGAATT GGGGTAGCCA
     3541 CTAGAGTAGA AAACCTTCAC TTTCCCGCCT GAGCACCGTT TCTGGAAAGG ATCTGAAGAT
     3601 TGAGATATGA TTTTTCGAAC TTGCACCGAT GTGGCCCTCG TGTAGAAGAC GAGGCAGAGT
     3661 GGATATAGTG CCACTGAAGA CATGCAGCAA GCTACCGAAC AACGCGATAA TGGAGACTAG
     3721 CGCGTCTGCC ATTGGCAACC GTGCTCGCCT TCTCGTGATC TTACGTGTCG CGTCTCTTCA
     3781 TCTCCGTACA CGAAAAATAT TGGTATGCGC GTGCATTATG CTTTCAGTAC GTGTAAATGA
     3841 GAGACAGGCA ATGCCACACT ACTGGCGCAG GACATGTTAT CCTCATCCGG GTCGCTTTTC
     3901 TTGCTCTATG CAAGGAAAGG GGCGGAAATG ATAGAGATTG ATAAATTGAT CGACGCGGAA
     3961 GAGTTATTAC TCTGCATGAC AATGAAGTGT GCTTTTAAAG TTTTGTTTAT CGAGAGGCCT
     4021 CGTGCGAGAA ATTTTTGTCG CAGCATGATT GACTTGTAGG ATAGATACTA GCTGGACTGG
     4081 TCTTCGACAT CCCTACACCT CCTGCCAAAC GGAAAAAAAA AGCATCTGTC GGCTGCACAC
     4141 AGATTGCGAC TACTTATAAC TTCAAACTAT GCTATAAGTG TCCTTTTCTT TCTTTCTTTT
     4201 CTTTCCTTGC CGTCCTTTAT GCCCCTGCAG GGTACGTTTT AGACGGACTA GGCAGTATAA
     4261 CTTCGTATAG TATACCTTAT ACGAAGTTAT GGCGCGCCAG GCTACGTTAG TTCAGCAGCT
     4321 GAGAACGACC ACGAACGGGA A
//
```

.GB formatted Sequence for lox2272-BLE-GFP-lox2272 fragment

```
LOCUS       lox2272-BLE-GFP-lox2272      4029 bp    DNA     linear   UNA 24-JUL-2017
DEFINITION  Complementary copy of pUC19.
```

```
KEYWORDS    ATCC.
COMMENT     This file is created by Vector NTI
            http://www.invitrogen.com/
FEATURES             Location/Qualifiers
     misc_feature    253..286
                     /feature_type="lox sites"
                     /vntifkey="21"
                     /label=lox2272
     terminator      complement(287..604)
                     /vntifkey="43"
                     /label=EIF3_Term
                     /label=EIF3_T
     CDS             complement(605..979)
                     /vntifkey="4"
                     /label=bleR
                     /label=BLE
     promoter        complement(1001..2002)
                     /vntifkey="21"
                     /label="TCT_Pro (Sequence corrected)"
                     /label=TCT_P
     terminator      complement(2016..2215)
                     /vntifkey="43"
                     /label=Term5
                     /label=GNPDA_T
     CDS             complement(2216..2917)
                     /vntifkey="4"
                     /label=GFP
     promoter        complement(2918..3739)
                     /vntifkey="21"
                     /label="Promoter 4AIII sequence corrected"
                     /label=4AIII_P
     misc_feature    3740..3773
                     /feature_type="lox sites"
                     /vntifkey="21"
                     /label=lox2272
ORIGIN
        1 GCCACCTCTG ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGGCGG AGCCTATGGA
       61 AAAACGCCAG CAACGCGGCC TTTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA
      121 TGTTCTTTCC TGCGTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG
      181 CTGATACCGC TCGCCGCAGC CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG
      241 AAGAGCGGCC GCATAACTTC GTATAGGATA CTTTATACGA AGTTATCGTA TGGTCGACGG
      301 TTGCTCGGAT GGGGGGGGCG GGGAGCGATG GAGGGAGGAA GATCAGGTAA GGTCTCGACA
      361 GACTAGAGAA GCACGAGTGC AGGTATAAGA AACAGCAAAA AAAAGTAATG GGCCCAGGCC
      421 TGGAGAGGGT ATTTGTCTTG TTTTTCTTTG GCCAGGAACT TGTTCTCCTT TCTTCGTTTC
      481 TAGGACCCCG ATCCCCGCTC GCATTTCTCT CTTCCTCAGC CGAAGCGCAG CGGTAAAGCA
      541 TCCATTTTAT CCCACCGAAA GGGCGCTCCC AGCCTTCGTC GAGCGGAACC GGGGTTACAG
      601 TGCCTTAGTC CTGCTCCTCG GCCACGAAGT GCACGCAGTT GCCGGCCGGG TCGCGCAGGG
      661 CGAACTCCCG CCCCCACGGC TGCTCGCCGA TCTCGGTCAT GGCCGGCCCG GAGGCGTCCC
      721 GGAAGTTCGT GGACACGACC TCCGACCACT CGGCGTACAG CTCGTCCAGG CCGCGCACCC
      781 ACACCCAGGC CAGGGTGTTG TCCGGCACCA CCTGGTCCTG GACCGCGCTG ATGAACAGGG
      841 TCACGTCGTC CCGGACCACA CCGGCGAAGT CGTCCTCCAC GAAGTCCCGG GAGAACCCGA
      901 GCCGGTCGGT CCAGAACTCG ACCGCTCCGG CGACGTCGCG CGCGGTGAGC ACCGGAACGG
      961 CGCTGGTCAG CTTGGCCATC GAGCTCGGTA CCCGGGGATC CATGATTGTT GTATTATGTA
     1021 CCTATGTTTG TGATGAGACA ATAAATATGA GAAGAGAACG TTGCGGCCAC TTTTTTCTCC
     1081 TTCCTTCGCG TGCTCATGTT GGTGGTTTGG GAGGCAGAAG ATGCATGGAG CGCCACACAT
     1141 TCGGTAGGAC GAAACAGCCT CCCCCACAAA GGGACCATGG GTAGCTAGGA TGACGCACAA
     1201 GCGAGTTCCC GCTCTCGAAG GGAAACCCAG GCATTTCCTT CCTCTTTTCA AGCCACTTGT
     1261 TCACGTGTCA ACACAATTTT GGACTAAAAT GCCCCTCGGA ACTCGGCAGG CCTCCCTCTG
     1321 CTCCGTTGTC CTGGTCGCCG AGAACGCGAG ACCGTGCCGC ATGCCATCGA TCTGCTCGTC
     1381 TGTACTACTA ATCGTGTGCG TGTTCGTGCT TGTTTCGCAC GAAATTGTCC TCGTTCGGCC
```

```
1441 CTCACAACGG TGGAAATCGG TGCTAGAATA AAGTGAGGTG GCTTATTTCA ATGGCGGCCG
1501 TCATCATGCG GGATCAACTG AAGTACGGCG GGTTCTCGAG ATTTCATCGT GCTCGTCCAG
1561 AGCAGGTGTT TTGCCTGCAG CTCTTCATGT TTAGGGGTCA TGATTTCATC TGATATGCCG
1621 TAAGAAAACC AATATTCACT TCTCAATTTT CCATGGAAAG GTGAAGGCCT AGGTTGTGTG
1681 CGAGGCAACG ACTGGGGAGG GATCGCAACA TTCTTGCTAA CCTCCCCTCT ATCTTGGCCG
1741 CTGTGAATCG GCATATTTAC CGGGCTGAAT TGAGAAAGTG TTTTGAGGGA ATTAAAAGGT
1801 GGCTGTCTTG CAAGCTTGGC TTCAGTGCCT GCTTAATTCG AACCGATCCA GCTTGTGATG
1861 AGGCCTTCCT AAGCCTGGTA GTCAGAAGCG ACATGGCGCT ATAAATTTCG TCTCAGTTGG
1921 AGAGTAGAAA AGCATGATTC GAACACGGTT TTCAACTGCC AAAGATATCT CCATTGTTTC
1981 CTTCAATCTG TACACCTGCA CGGGCCAGTG AGGCCAGGAA ATAAAGATGG ACAGACGGCA
2041 TGCTAGTAGA CTTTGTTGAG ATTAGTGTTT GTGTTCGTCT TTATGGCTTT GAGTGGGCCC
2101 CCTTAACCTA TACACACATG ACAATCAGGT GACGAGGAAG CTCTCGACTC TCCAGGTCTC
2161 CAACACATCA TGAGGACGCC GCTCTGCCAG GACCCTCCCC GACTCCTTCC CACCCTTATT
2221 CTTCACCGGC ATCTGCATCC GGGGTCTTGA AGGCGTGCTG GTACTCCACG ATGCCCAGCT
2281 CGGTGTTGCT GTGATCCTCC TCCACGCGGC GGAAGGCGAA CATGGGGCCC CCGTTCTGCA
2341 GGATGCTGGG GTGGATGGCG CTCTTGAAGT GCATGTGGCT GTCCACCACG GAGCTGTAGT
2401 AGCCGCCGTC GCGCAGGCTG AAGGTGCGGG TGAAGCTGCC ATCCAGATCG TTATCGCCCA
2461 TGGGGTGCAG GTGCTCCACG GTGGCGTTGC TGCGGATGAT CTTGTCGGTG AAGATCACGC
2521 TGTCCTCGGG GAAGCCGGTG CCCATCACCT TGAAGTCGCC GATCACGCGG CCGGCCTCGT
2581 AGCGGTAGCT GAAGCTCACG TGCAGCACGC CGCCGTCCTC GTACTTCTCG ATGCGGGTGT
2641 TGGTGTAGCC GCCGTTGTTG ATGGCGTGCA GGAAGGGGTT CTCGTAGCCG CTGGGGTAGG
2701 TGCCGAAGTG GTAGAAGCCG TAGCCCATCA CGTGGCTCAG CAGGTAGGGG CTGAAGGTCA
2761 GGGCGCCTTT GGTGCTCTTC ATCTTGTTGG TCATGCGGCC CTGCTCGGGG GTGCCCTCTC
2821 CGCCGCCCAC CAGCTCGAAC TCCACGCCGT TCAGGGTGCC GGTGATGCGG CACTCGATCT
2881 CCATGGCGGG CAGGCCGCTC TCGTCGCTCT CCAACATGTA AGCTAGGCTT TTGGTGAGAG
2941 AATGGGAAAG AAGTTAGATG TAAAATTGAA CTTCGGTTGT CGAATTTCAG AGGTAGTGCG
3001 CGGTGCGTGC GCAACGAAGG ACCGTCTGCG ACAGTCGGAG AGAATTGGGG TAGCCACTAG
3061 AGTAGAAAAC CTTCACTTTC CCGCCTGAGC ACCGTTTCTG GAAAGGATCT GAAGATTGAG
3121 ATATGATTTT TCGAACTTGC ACCGATGTGG CCCTCGTGTA GAAGACGAGG CAGAGTGGAT
3181 ATAGTGCCAC TGAAGACATG CAGCAAGCTA CCGAACAACG CGATAATGGA GACTAGCGCG
3241 TCTGCCATTG GCAACCGTGC TCGCCTTCTC GTGATCTTAC GTGTCGCGTC TCTTCATCTC
3301 CGTACACGAA AAATATTGGT ATGCGCGTGC ATTATGCTTT CAGTACGTGT AAATGAGAGA
3361 CAGGCAATGC CACACTACTG GCGCAGGACA TGTTATCCTC ATCCGGGTCG CTTTTCTTGC
3421 TCTATGCAAG GAAAGGGGCG GAAATGATAG AGATTGATAA ATTGATCGAC GCGGAAGAGT
3481 TATTACTCTG CATGACAATG AAGTGTGCTT TTAAAGTTTT GTTTATCGAG AGGCCTCGTG
3541 CGAGAAATTT TTGTCGCAGC ATGATTGACT TGTAGGATAG ATACTAGCTG GACTGGTCTT
3601 CGACATCCCT ACACCTCCTG CCAAACGGAA AAAAAAAGCA TCTGTCGGCT GCACACAGAT
3661 TGCGACTACT TATAACTTCA AACTATGCTA TAAGTGTCCT TTTCTTTCTT TCTTTTCTTT
3721 CCTTGCCGTC CTTTATGCCA TAACTTCGTA TAGGATACTT TATACGAAGT TATCCTGCAG
3781 GCAGTTGGTA CGGCATATTA TGGTTTAAAC ATCTATCCTC CAGATCACCA GGGCGCGCCT
3841 ATGCTTGTAA ACCGTTTTGT GAAAAAATTT TTAAAATAAA AAAGGGGACC TCTAGGGTCC
3901 CCAATTAATT AGTAATATAA TCTATTAAAG GTCATTCAAA AGGTCATCCA GACGAAAGGG
3961 CCTCGTGATA CGCCTATTTT TATAGGTTAA TGTCATGATA ATAATGGTTT CTTAGACGTC
4021 AGGTGGCAC
//
```