Non-transgenic marker-free gene disruption by an episomal CRISPR system in the oleaginous microalga, *Nannochloropsis oceanica* CCMP1779

Eric Poliner[1,2], Tomomi Takeuchi[2,3], Zhi-Yan Du[2,3], Christoph Benning[2,3,4], Eva M. Farré[4*]

1. Cell and Molecular Biology Program, Michigan State University, East Lansing, Michigan
2. MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, Michigan
3. Biochemistry and Molecular Department, Michigan State University, East Lansing, Michigan
4. Plant Biology Department, Michigan State University, East Lansing, Michigan

*Corresponding author

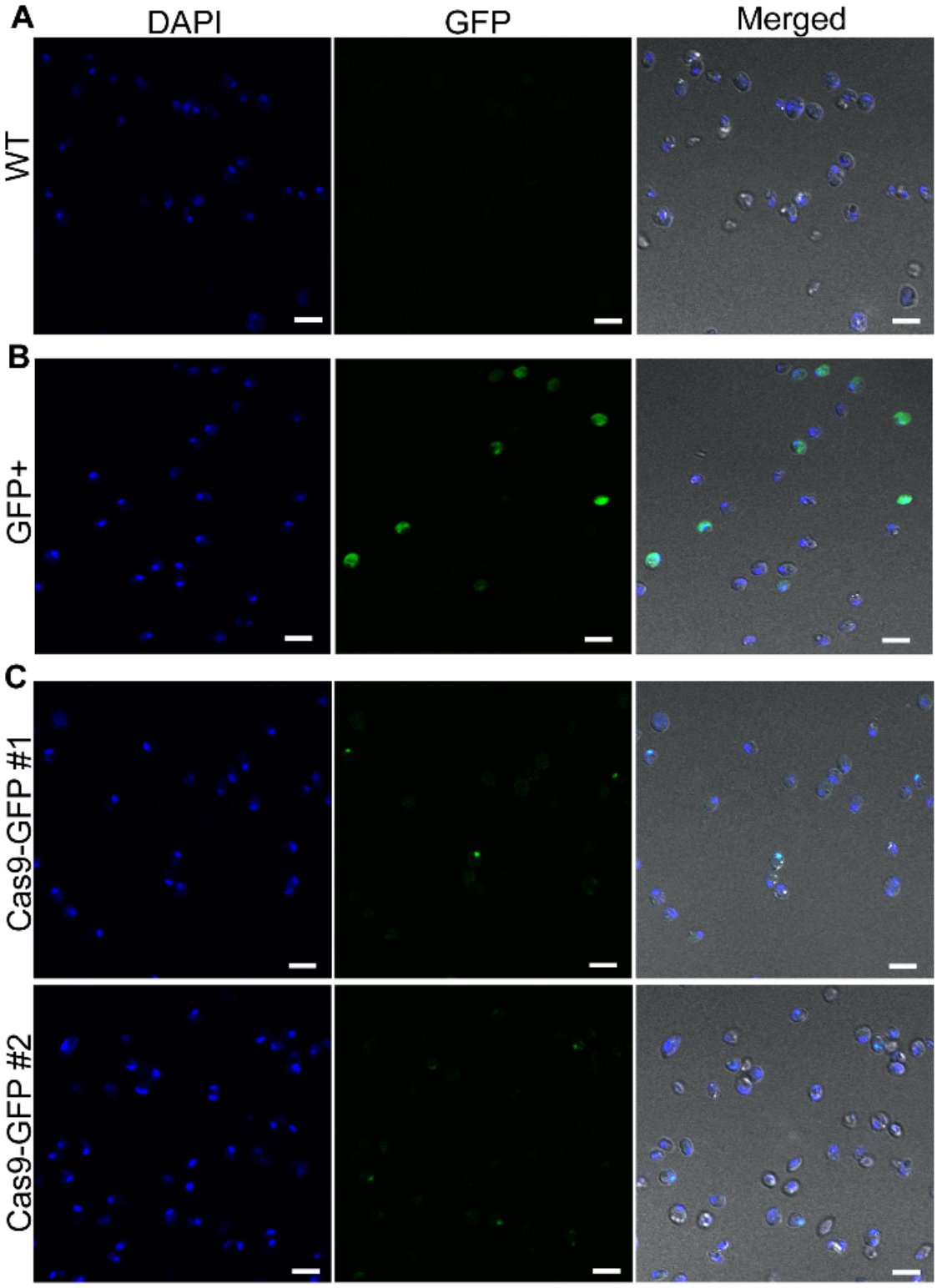


Figure S1. Confocal microscopy of Cas9-GFP expressing *N. oceanica* with the nucleus stained by DAPI. Scale bars of 5 µm. (A) Wild-type (WT) cultures with DAPI nuclear stain, GFP signal, and merged view of DAPI, GFP and brightfield signals. (B) Transformants producing GFP without a targeting sequence DAPI nuclear stain, GFP signal, and merged view of DAPI, GFP and brightfield signal. (C) Transformants producing Cas9-GFP with DAPI nuclear stain, GFP signal, and merged view of DAPI, GFP and brightfield signal. Two examples are shown.

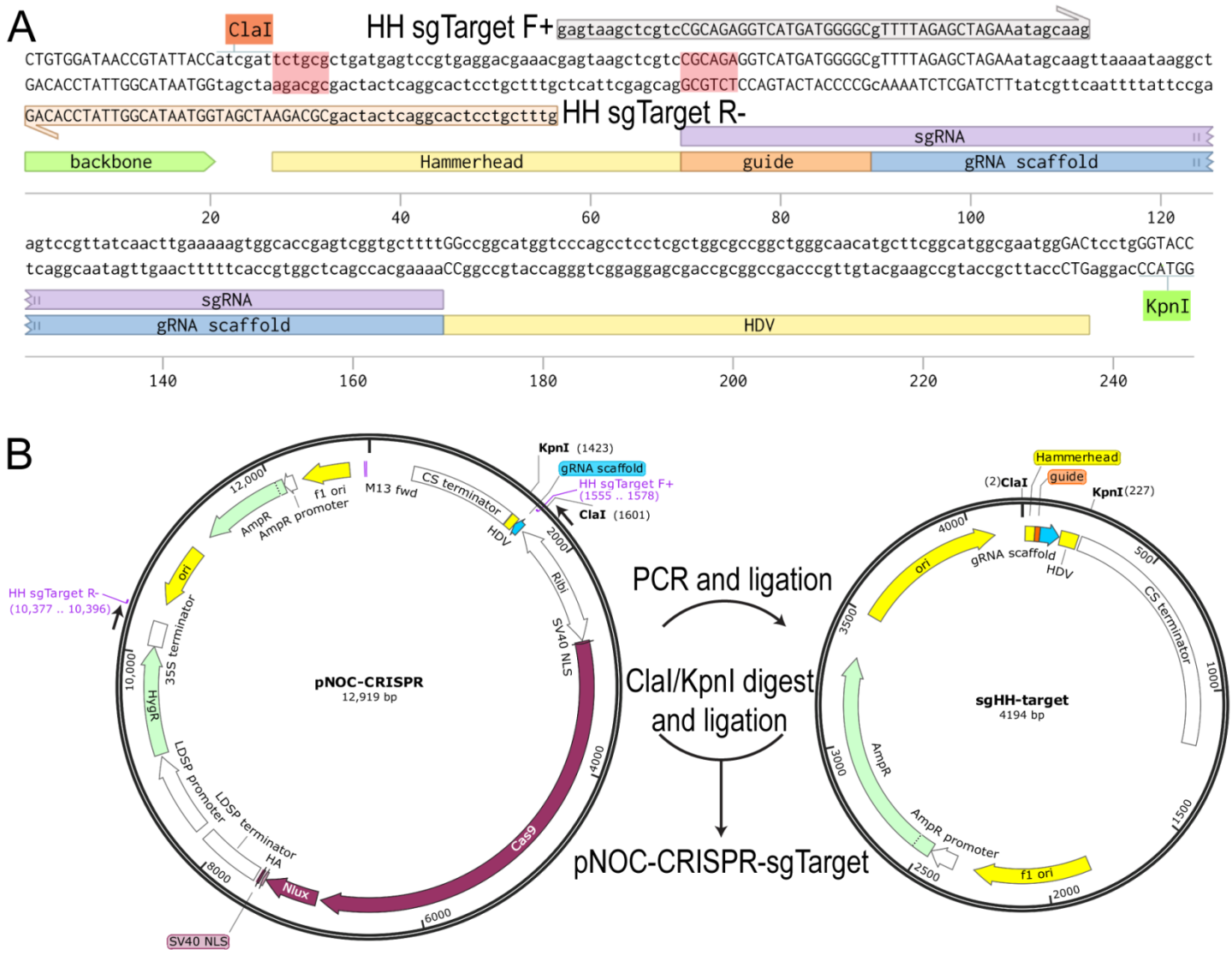


Figure S2. Cloning strategies for the generation of a ribozyme-sgRNA. (A) Example of a ribozyme (highlighted in yellow) flanked sgRNA (highlighted in purple) with unique ClaI and KpnI restriction sites indicated. The hammerhead ribozyme is customized to each guide sequence (highlighted in orange) with a reverse complement of the first six basepairs (highlighted in red). The 80 bp gRNA scaffold (highlighted in blue) has a HDV ribozyme at the 3' end. Primers used to clone the hammerhead-guide sequence, HH sgTarget F+ and HH sgTarget R-, are indicated as grey and orange arrows respectively. (B) Subcloning strategy to generate ribozyme flanked sgRNA in pNOC-CRISPR series. The HH sgTarget F+ primer, anneals at the 5' end of the gRNA scaffold (blue) on pNOC-CRISPR vectors and includes a 5' extension containing the 20 bp guide sequence and 3' end of the HH sequence. The HH sgTarget R- primer anneals to a region in the vector backbone downstream of the hygromycin resistance gene (HygR, green) and includes a 5' extension containing

the 5’ end of the HH ribozyme sequence, 6 bp specific to the guide sequence for loop formation and a ClaI site. The vector map indicates the location of primer annealing. After amplification, ligation, and transformation into *E. coli*, a subcloning vector (sgHH-target) containing the HH-guide-sgRNA-HDV sequence is recovered. The ClaI/KpnI digested fragment of the sgHH-target vector is then ligated with ClaI/KpnI digested pNOC-CRISPR vector to form the final pNOC-CRISPR-sgTarget vector. (C) Sequences of generic sgTarget F+ and HH sgTarget R- primers. The 20 bp guide sequence is placed in the guide region (shown in uppercase) of a HH sgTarget F+ primer, and the first 6 bp of the guide sequence (uppercase, highlighted in red) placed in the asterisk region (*) of the HH sgTarget R- primer.

A

	1	2	3	4	5
A	48	48	62	53	57
B	43	67	57	34	43
C	48	53	2984	*651*	43
D	* 378	62	57	91	1380
E	24	62	62	38	48
F	53	57	2591	43	62
G	43	57	738	* 1734	920
H	* 335	62	~~6993~~	57	62

B

	1	2	3	4	5	6	7	11	12
A	14	14	10	14	~~7975~~	29	24	~~8138~~	*2543*
B	10	19	10	19	29	24	* *9556*	~~8459~~	* 21105
C	19	24	115	14	24	~~7721~~	43	48	34
D	6807	14	6356	14	53	48	6850	34	38
E	**4206**	24	~~6237~~	14	6821	34	34	*12066*	38
F	14	72	105	**623**	53	~~6931~~	**7544**	34	19
G	14	48	~~4833~~	62	34	182	48	29	19
H	10	72	**1059**	14	19	29	14	34	8828

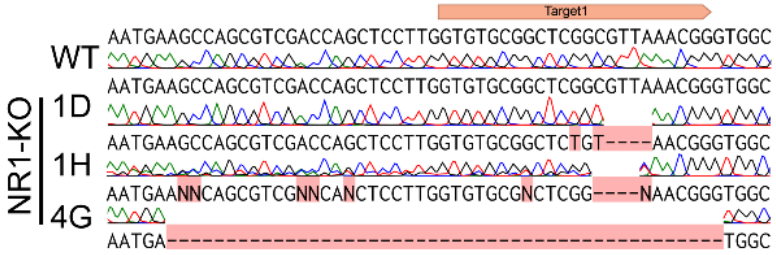
C

NR1-KO	Frame-shifted	In-frame	Low-quality	Wild-type	Total
Number	8	1	1	0	10
Percent	80%	10%	10%	0%	100%

NR2-KO	Frame-shifted	In-frame	Low-quality	Wild-type	Total
Number	6	3	7	4	20
Percent	30%	15%	35%	20%	100%

Cumulative	Frame-shifted	In-frame	Low-quality	Wild-type	Total
Number	14	4	8	4	30
Percent	47%	13%	27%	13%	100%

D



E

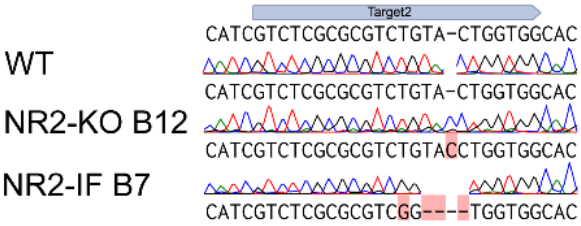


Figure S3. Identification of NR knockout mutants by CRISPR/Cas9. (A) Representative 96-well Nlux luminescence screen of NR1-KO lines. The PCR products of the *NR* genomic locus from bolded wells were submitted for Sanger sequencing. Frame-shift mutations are red, in-frame mutations italicized, colonies returning poor quality sequences are indicated by strikethroughs, wild-type sequences are unmodified (bolded), and lines selected are marked with an asterisks (*). (B) Representative 96-well Nlux luminescence screen of NR2-KO lines. Lines are highlighted as described in panel A. (C) Rates of sequencing results of NR1-KO and NR2-KO lines, and cumulative results. (D) Chromatographs of target1 region from wild-type and NR1-KO lines (identified by plate location). (E) Chromatographs of target2 region from wild-type, NR2-KO, and NR2-IF line (identified by plate location).

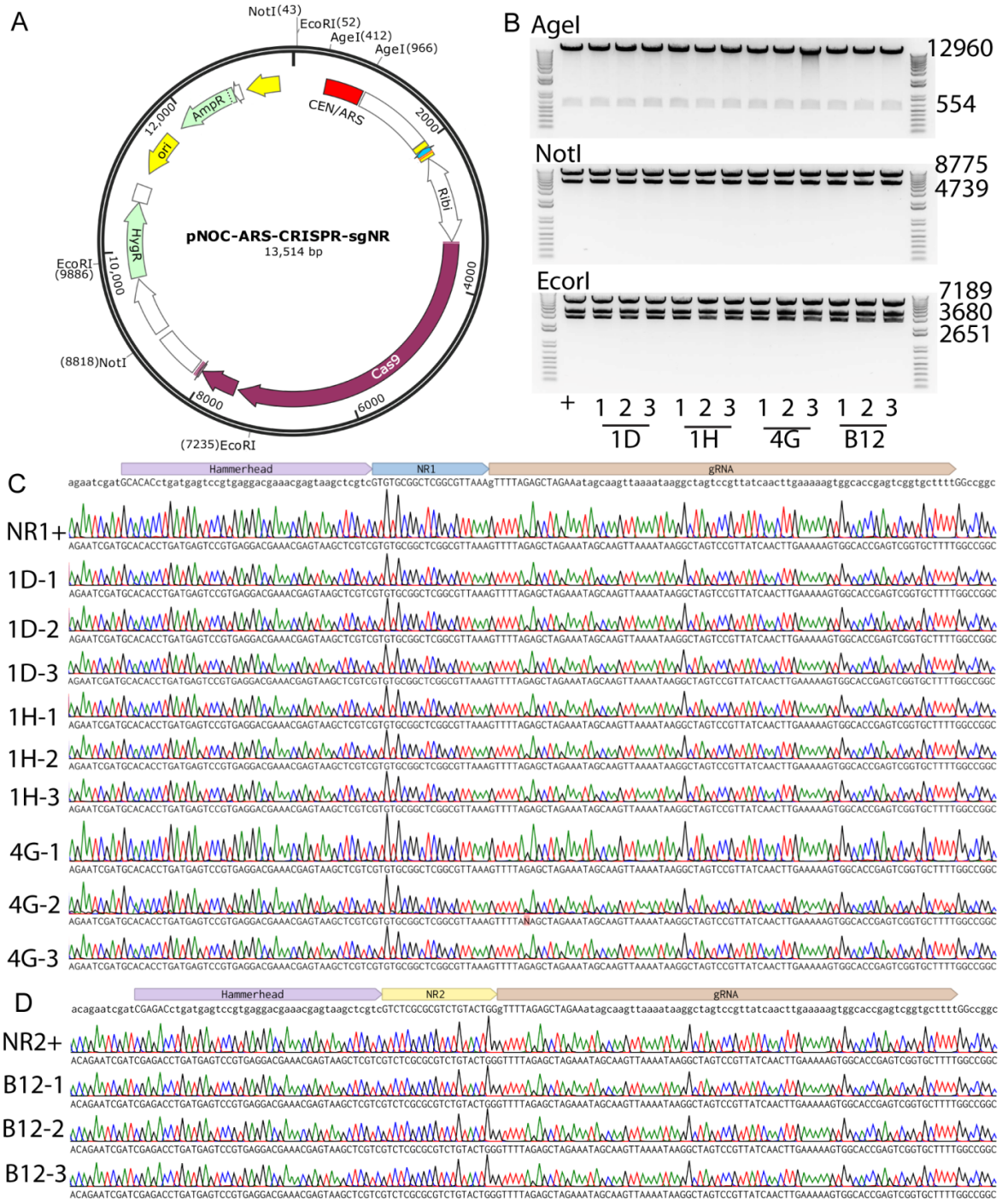
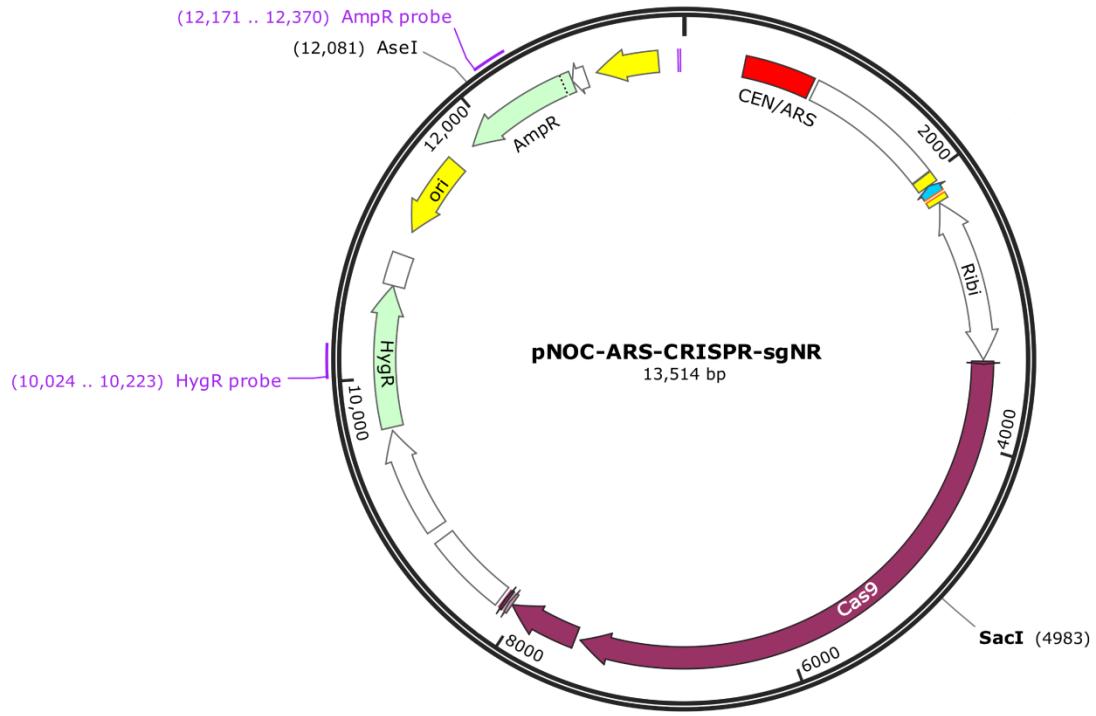
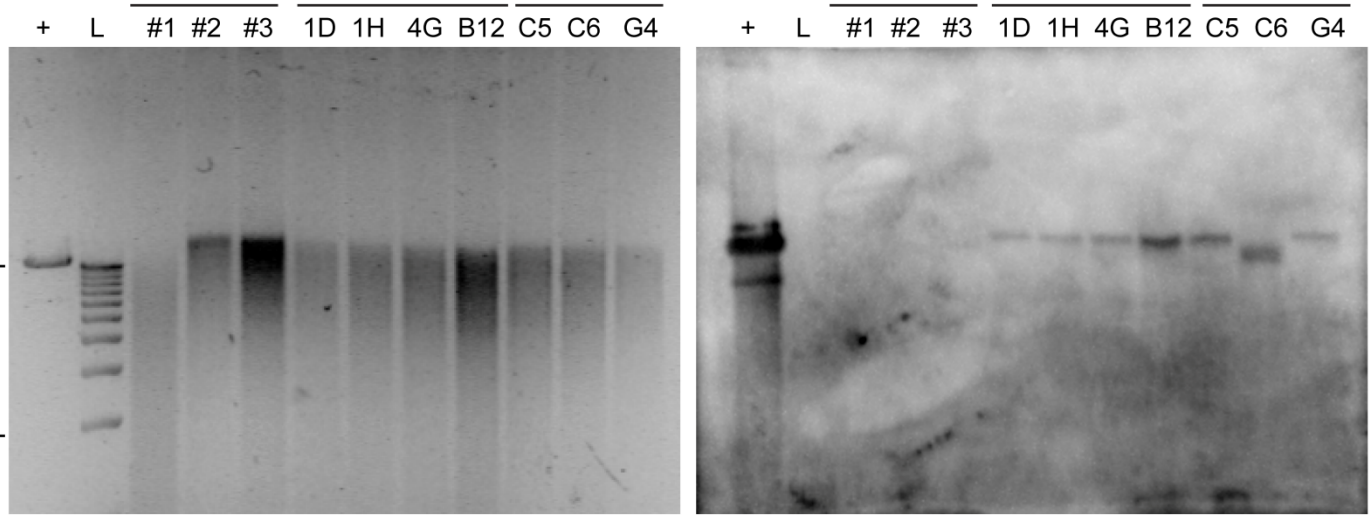


Figure S4. Verification of rescued episomes. (A) Plasmid map indicating restriction sites used for restriction fragment analysis. (B) Triplicate recovered plasmids from each NR-KO episomal line were digested with AgeI, NotI, and EcorI endonucleases. The positive control plasmid (+), pNOC-ARS-CRISPR-sgNR2, was used. Restriction fragments were visualized after separation by agarose (0.9%) gel electrophoresis using a 1 kb+ size ladder (Invitrogen). (C) Sanger sequencing of the sgRNA region of the recovered episomes from NR1-KO lines aligned to pNOC-ARS-CRISPR-sgNR1. (D) Sanger sequencing of the sgRNA region of the recovered episomes from the NR2-KO line aligned to pNOC-ARS-CRISPR-sgNR2.

A



B HygR probe



C AmpR probe

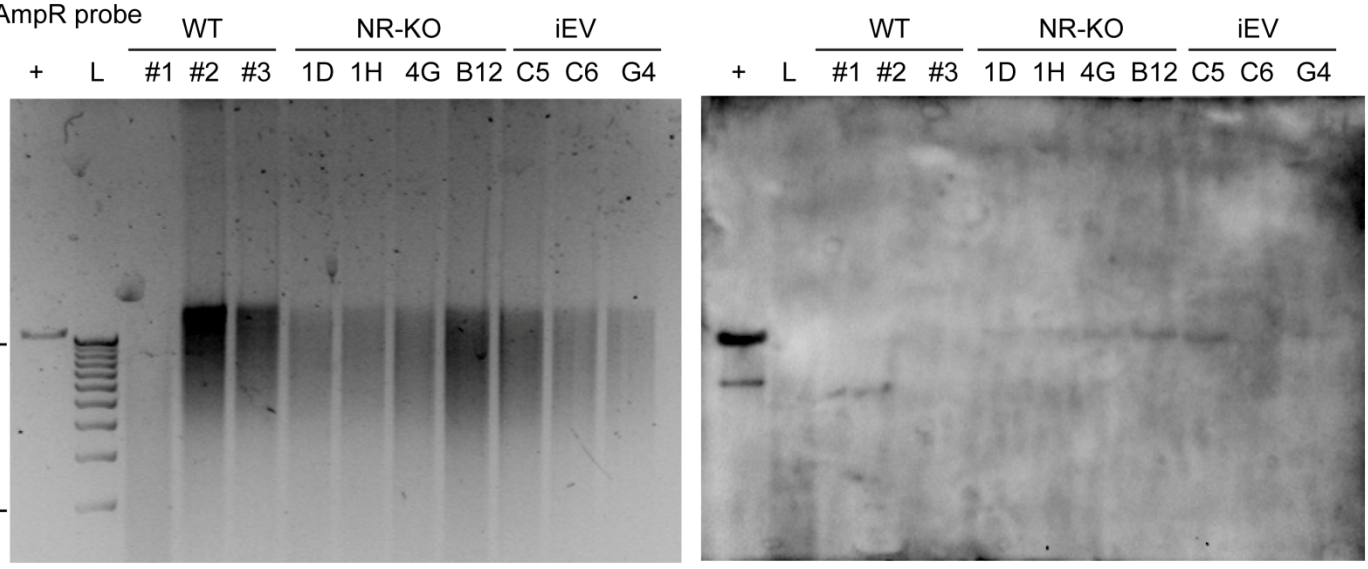
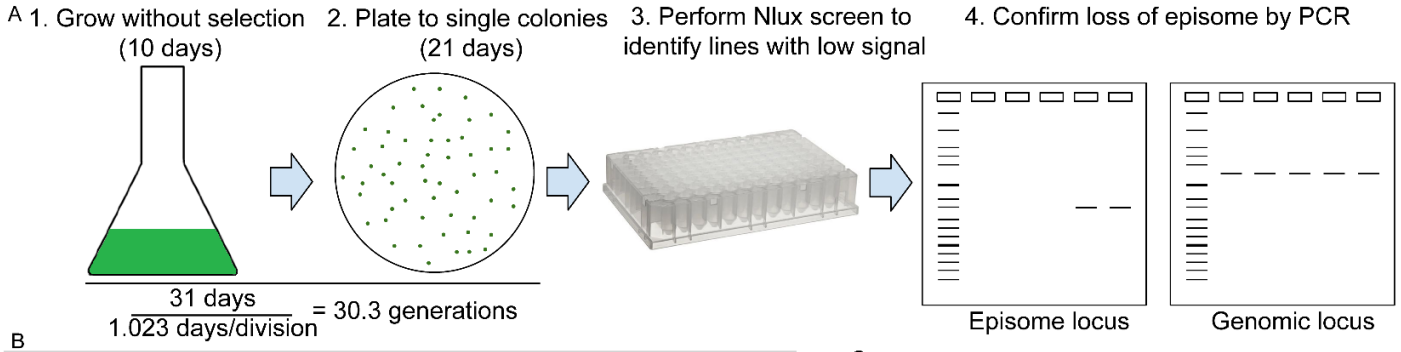


Figure S5. Southern blot analysis of the episomal and integrated empty-vector CRISPR mutants. (A) Plasmid map showing the pNOC-ARS-CRISPR-sgNR vectors used for transformation. DNA was digested SacI site prior to analysis. HygR and AmpR regions were used as probes. The AseI site was used to linearize the plasmid for integrated empty-vector controls. (B) and (C) Genomic DNA digested with SacI was separated with a 0.9% agarose gel. Left panels show the stained agarose gels and right panels show the hybridization signals. The ladder location of 12 kb and 3 kb bands are noted. The blot contains the plasmid pNOC-ARS-CRISPR-sgNR2 as a positive control (+), a 1 kb+ DNA ladder (L), three wild-type DNA (WT) samples as negative controls, DNA from episomal mutant lines (NR-KO), and DNA from integrated empty-vector lines (iEV). HygR and AmpR probes were used in (B) and (C), respectively.



B

	NR1-KO 1D						NR1-KO 1H					
	1	2	3	4	5	6	7	8	9	10	11	12
A	3195	**5044**	2802	4622	3286	1293	1485	704	53	91	19	**19**
B	1844	556	2088	2294	38	38	115	**43**	354	268	647	38
C	1274	125	1207	474	1025	129	350	125	398	34	29	**1485**
D	541	484	738	**91**	445	436	77	177	53	24	29	556
E	2496	2500	1763	2979	**1001**	1044	1356	**2678**	134	101	38	43
F	1025	647	4244	704	1284	263	680	29	747	833	43	48
G	86	**91**	326	115	153	431	675	38	848	29	139	**38**
H	220	216	163	331	29	**34**	216	273	**53**	762	**34**	522
	Low luminescence				25%		Low luminescence				58%	
	Segregation efficiency				99%		Segregation efficiency				97%	

	NR1-KO G4						NR2-KO B12					
	1	2	3	4	5	6	7	8	9	10	11	12
A	**4876**	86	53	77	**72**	77	**43**	369	53	**62**	72	364
B	**34**	273	144	704	627	81	158	**48**	77	163	211	1997
C	163	43	96	556	4182	57	38	57	125	77	239	48
D	29	153	43	48	96	1154	211	307	340	57	57	62
E	24	77	38	**34**	757	43	297	48	72	72	91	**2620**
F	53	57	38	43	105	881	216	1174	158	**623**	1073	484
G	38	**38**	57	**1466**	896	163	91	86	129	273	67	**43**
H	560	182	43	158	53	96	575	129	546	96	924	120
	Low luminescence				75%		Low luminescence				60%	
	Segregation efficiency				96%		Segregation efficiency				97%	

C

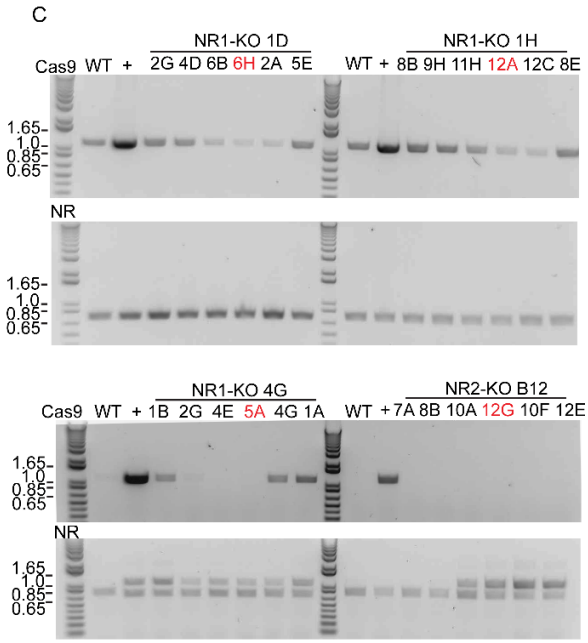


Figure S6. Curing episomes from NR-KO lines. (A) Strategy for the generation of episome-less NR-KO lines. NR-KO mutants carrying the episome were grown without selection for 10 days, and then plated on medium without selection. Single colonies were isolated in 96-well plates and screened for luminescence signal, followed by PCR to detect the *Cas9* gene. A PCR against the genomic sequence of the *NR* gene was used as genomic locus control. The number of generations used to determine segregation efficiency in panel B is shown. (B) Nlux luminescence screen of independent colonies after curing procedure of NR-KO lines. A cutoff of 300 counts/0.3 seconds was used to define low luminescent lines. The percentage of lines falling below the threshold are displayed beneath the raw measurements. The segregation efficiency based on the luminescence screen of each line is shown. Bolded wells were screened by PCR and lines in red were selected as putative "cured" lines. (C) Colony PCR for *Cas9* (episomal marker) and *NR* (genomic marker) genes separated on 0.9% agarose gels. 1 Kb+ ladder (L), wild-type (WT), episome carrying parental lines were used as positive control (+), 4 low-Nlux signal, and 2 high-Nlux recovered colonies (identified by location in the 96-well growth plate) were tested. Final cured lines are highlighted in red. The ladder location for 1.65, 1.0, 0.85, and 0.65 kb bands are noted on the left side of each gel. In some screens Cas9 PCR product was present in WT control reactions indicating cross contamination during colony PCR setup, therefore, we carried out follow-up confirmation of episome curing using multiple independent methods as shown in Figure 3.

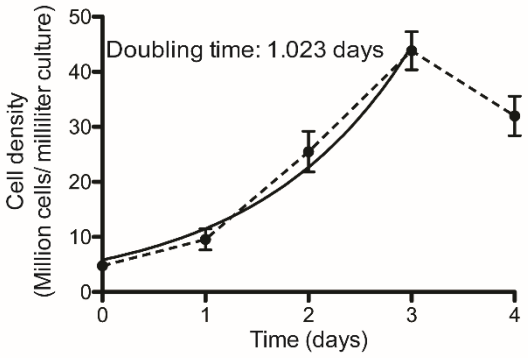


Figure S7. Growth curve of *N. oceanica* in F/2 with NH_4. Cell density was measured every 24 hours, cell density is plotted in closed circles (n = 4 biological replicates) and connected with a dashed line. The best-fit exponential curve is shown as a solid line, and the deduced doubling time noted at the top left of the graph.

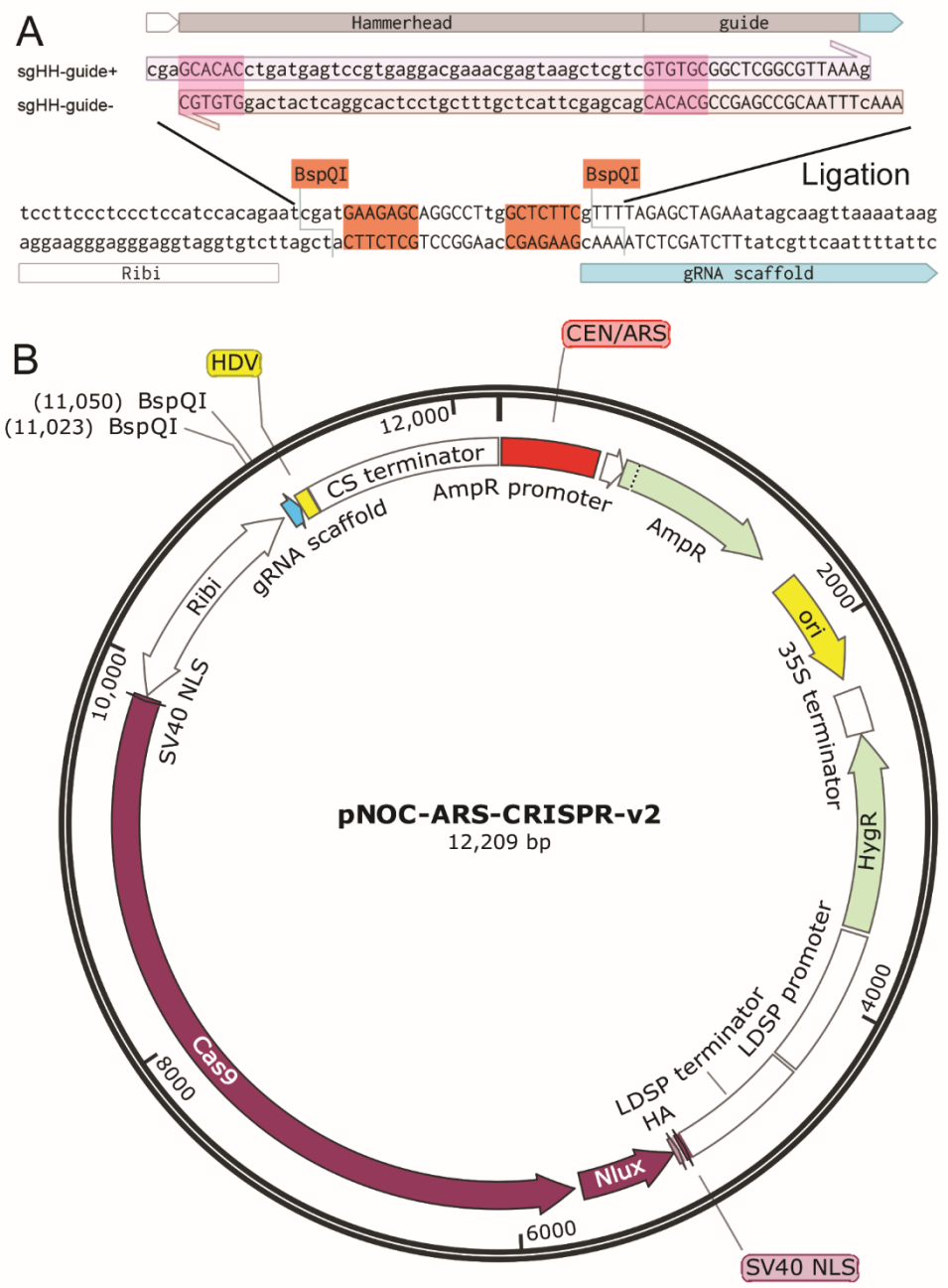


Figure S8. A one-vector CRISPR system for scarless cloning of guide sequences. (A) The multicloning site for guide sequence insertion contains a pair of BspQI sites for removal of the MCS and generation of sticky ends on the Ribi promoter and gRNA scaffold. To clone into this site, a pair of complementary oligonucleotides with overhanging ends are annealed and ligated with the digested vector. An example pair of oligonucleotides is shown. The first six bases on the Hammerhead (capitalized) correspond to the reverse complement of the first six bases of the guide sequence as indicated in red. (B) Map of pNOC-ARS-CRISPR-v2.