Cell Reports, Volume *16*

Supplemental Information

Repeat Size Determination by Two Molecular Rulers in the Type I-E CRISPR Array

Moran G. Goren, Shany Doron, Rea Globus, Gil Amitai, Rotem Sorek, and Udi Qimron

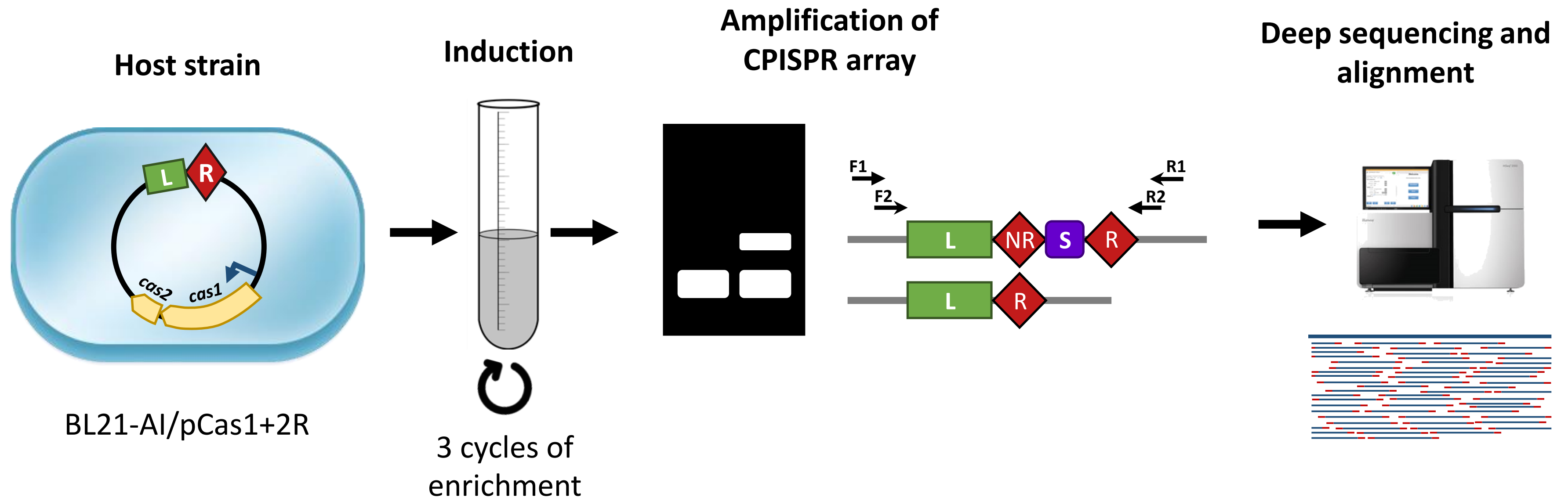

Figure S1; Related to Figures 1, 2, and 3. Assay for monitoring the effects of repeat modification on adaptation. A plasmid encoding Cas1–Cas2 and a leader–repeat sequence, occasionally with modifications to the WT sequence, was transformed into *E. coli* BL21-AI. Bacteria were induced to express Cas1–Cas2, grown, diluted, and grown again, three times in total. A sample of the culture was then taken for PCR1 using primers F1–R1. PCR2 was generated following gel extraction of the adapted band using primers F2–R2. Both products were analyzed using high-throughput DNA sequencing.

Table S1. Bacterial strains, plasmids and oligonucleotides used in this study. Related to Experimental Procedures.

Bacterial strains	Description/sequence	Source or reference
NEB5α	F^- ϕ80*lacZ*ΔM15Δ(*lacZYA-argF*) U169 *deoR recA1 endA1 hsdR17* (r_k^-, m_k^+) *gal*$^-$ *phoA supE44* λ$^-$ *thi* $^-$*1 gyrA96 relA1*	New England Biolabs
IYB5283	BL21-AI with no repeats in CRISPR I, kan^r, tet^r	(Yosef et al., 2012)[a]
Plasmids		
pCas1+2	pCDF-1b (Novagen) cloned with *cas1,2* under T7 promoter, str^r	(Yosef et al., 2012)[a]
pCas1+2R (WT)	pCDF-1b (Novagen) cloned with *cas1,2* under T7 promoter, minimal leader and single repeat of CRISPR 1 array, str^r	This study
The following plasmids are Identical to pCas1+2R except for the repeat sequence specified below:		
S1	TGTGTCCCCGCGCCAGCGGGGATAAACC	This study
S2	GTGTTCCCCGCTAACGCGGGGATAAACC	This study
S3	GTGTTCCCCGCGCCAGCGGGGAGCCCAA	This study
S4	GTGTGAAAATAGCCATATTTTCTAAACC	This study
S5	GTGTGAAAATAGCCAGCGGGGATAAACC	This study
S6	GTGTTCCCCGCGCCATATTTTCTAAACC	This study
S7	GTGTAGGGGCGGCCACGCCCCTTAAACC	This study
D1	G_GTTCCCCGCGCCAGCGGGGATAAACC	This study
D2	GTGT_CCCCGCGCCAGCGGGGATAAACC	This study
D3	GTGTTCC_CGCGCCAGCGGGGATAAACC	This study
D4	GTGTTCCCCGC_CCAGCGGGGATAAACC	This study
D5	GTGTTCCCCGCG_CAGCGGGGATAAACC	This study
D6	GTGTTCCCCGCGCC_GCGGGGATAAACC	This study
D7	GTGTTCCCCGCGCCAGCGGG_ATAAACC	This study
D8	GTGTTCCCCGCGCCAGCGGGGA_AAACC	This study
D9	GTGTTCCCCGCGCCAGCGGGGATAA_CC	This study
D10	GTGTTCCCCGCGCCAGCGGGGATAAAC_	This study
I1	GTGTTTCCCCGCGCCAGCGGGGATAAACC	This study
I2	GTGTTCCCCCCGCGCCAGCGGGGATAAACC	This study
I3	GTGTTCCCCGCGGCCAGCGGGGATAAACC	This study
I4	GTGTTCCCCGCGCCCAGCGGGGATAAACC	This study
I5	GTGTTCCCCGCGCCAAGCGGGGATAAACC	This study
I6	GTGTTCCCCGCGCCAGCGGGGGATAAACC	This study
I7	GTGTTCCCCGCGCCAGCGGGGATTAAACC	This study
I8	GTGTTCCCCGCGCCAGCGGGGATAAAACC	This study
D5S1	GTGTTCCCCGCGCAACGGGGATAAACC	This study
D5S2	GTGTTCCCCGCGCAGAGGGGATAAACC	This study
D5S3	GTGTTCCCCGCGCAGCAGGGATAAACC	This study
D5S4	GTGTTCCCCGCGCAGCGAGGATAAACC	This study
D5S5	GTGTTCCCCGCGCAGCGGAGATAAACC	This study
D5S6	GTGTTCCCCGCGCAGAAAGGATAAACC	This study
I4S1	GTGTTCCAAACGCCCAGCGGGGATAAACC	This study
Oligonucleotides	**5'→3'**	
OA2F	CCTTTGATCTTTTCTACTGA	
OA2R	ATGGGGCTGACTTCAGGTGC	
RE10RD	NNNNTGGATGTGTTGTTTGTG	
IY230R1	NNNNAAATGAGCGATGATATTTGTGCT	
MG132F	GTTATGTTTAGATGTGTCCCCGCGCCAGCGG	
MG132R	CCGCTGGCGCGGGGACACATCTAAACATAAC	
MG82F	GCGGGGATAAACCGAGCACA	
MG86R	GTTAGCGGGGAACACTCTAAACATAACCTATTAT	
MG126F	GCCCAAGAGCACAAATATCATCGCTC	
MG126R	TCCCCGCTGGCGCGGGGAACACTC	
MG85F	TATTTTCTAAACCGAGCACAAATATCA	
MG85R	TGGCTATTTTCACACTCTAAACATAACCTATTAT	
MG197F	GAAAATAGCCAGCGGGGATAAACCGAG	
MG197R	ACACTCTAAACATAACCTAT	
MG198F	TATTTTCTAAACCGAGCACAAATATCA	

MG198R	TGGCGCGGGGAACACTCTAA
MG88F	CGCCCCTTAAACCGAGCACAAA
MG196R	TGGCCGCCCCTACACTCTAAACAT
MG199F	GTTCCCCGCGCCAGCGGGGA
MG199R	CTCTAAACATAACCTATTAT
MG200F	TCCCCGCGCCAGCGGGGATA
MG200R	CACTCTAAACATAACCTATT
MG218F	TGTTTAGAGTGTTCCCGCGCCAGCGGGGAT
MG218R	ATCCCCGCTGGCGCGGGAACACTCTAAACA
MG202F	CCAGCGGGGATAAACCGAGC
MG202R	GCGGGGAACACTCTAAACAT
MG203F	CAGCGGGGATAAACCGAGCAC
MG203R	CGCGGGGAACACTCTAAAC
MG204F	GCGGGGATAAACCGAGCACA
MG204R	GGCGCGGGGAACACTCTAAA
MG205F	GGATAAACCGAGCACAAATA
MG215R	CCCGCTGGCGCGGGGAACACT
MG206F	AAACCGAGCACAAATATCAT
MG216R	AATCCCCGCTGGCGCGGGGAA
MG219F	GCGCCAGCGGGGATAACCGAGCACAAATAT
MG219R	ATATTTGTGCTCGGTTATCCCCGCTGGCGC
MG208F	CGAGCACAAATATCATCGCT
MG208R	TTTATCCCCGCTGGCGCGGG
MG200F	TCCCCGCGCCAGCGGGGATA
MG210R	AACACTCTAAACATAACCTAT
MG201F	CGCGCCAGCGGGGATAAACC
MG211R	GGGGAACACTCTAAACATAAC
MG202F	CCAGCGGGGATAAACCGAGC
MG212R	CCGCGGGGAACACTCTAAACA
MG203F	CAGCGGGGATAAACCGAGCAC
MG213R	GGCGCGGGGAACACTCTAAAC
MG204F	GCGGGGATAAACCGAGCACA
MG214R	TTGGCGCGGGGAACACTCTAA
MG205F	GGATAAACCGAGCACAAATA
MG215R	CCCGCTGGCGCGGGGAACACT
MG206F	AAACCGAGCACAAATATCAT
MG216R	AATCCCCGCTGGCGCGGGGAA
MG220F	CGCCAGCGGGGATAAAACCGAGCACAAATAT
MG220R	ATATTTGTGCTCGGTTTTATCCCCGCTGGCG
MG243F	AGTGTTCCCCGCGCAACGGGGATAAACCGAG
MG243R	CTCGGTTTATCCCCGTTGCGCGGGGAACACT
MG244F	GTGTTCCCCGCGCAGAGGGGATAAACCGAGC
MG244R	GCTCGGTTTATCCCCTCTGCGCGGGGAACAC
MG245F	TGTTCCCCGCGCAGCAGGGATAAACCGAGCA
MG245R	TGCTCGGTTTATCCCTGCTGCGCGGGGAACA
MG246F	GTTCCCCGCGCAGCGAGGATAAACCGAGCAC
MG246R	GTGCTCGGTTTATCCTCGCTGCGCGGGGAAC
MG247F	TTCCCCGCGCAGCGGAGATAAACCGAGCACA
MG247R	TGTGCTCGGTTTATCTCCGCTGCGCGGGGAA
MG248F	GTGTTCCCCGCGCAGAAAGGATAAACCGAGCAC
MG248R	GTGCTCGGTTTATCCTTTCTGCGCGGGGAACAC
MG239F	TGTTTAGAGTGTTCCAAACGCCCAGCGGGGATA
MG239R	TATCCCCGCTGGGCGTTTGGAACACTCTAAACA

[a]Yosef, I., Goren, M.G., and Qimron, U. (2012). Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli. Nucleic Acids Res *40*, 5569-5576.

Table S2. Primers and templates used for plasmid construction. Related to Experimental Procedures.

Plasmid	Primers for PCR	DNA template
S1	MG132F, MG132R	pCas1+2R
S2	MG82F, MG86R	"
S3	MG126F, MG126R	"
S4	MG85F, MG85R	"
S5	MG197F, MG197R	"
S6	MG198F, MG198R	"
S7	MG88F, MG196R	"
D1	MG199F, MG199R	"
D2	MG200F, MG200R	"
D3	MG218F, MG218R	"
D4	MG202F, MG202R	"
D5	MG203F, MG203R	"
D6	MG204F, MG204R	"
D7	MG205F, MG215R	"
D8	MG206F, MG216R	"
D9	MG219F, MG219R	"
D10	MG208F, MG208R	"
I1	MG200F, MG210R	"
I2	MG201F, MG211R	"
I3	MG202F, MG212R	"
I4	MG203F, MG213R	"
I5	MG204F, MG214R	"
I6	MG205F, MG215R	"
I7	MG206F, MG216R	"
I8	MG220F, MG220R	"
D5S1	MG243F, MG243R	"
D5S2	MG244F, MG244R	"
D5S3	MG245F, MG245R	"
D5S4	MG246F, MG246R	"
D5S5	MG247F, MG247R	"
D5S6	MG248F, MG248R	"
I4S1	MG239F, MG239R	"

Supplemental Experimental Procedures. Related to Experimental Procedures.

Reagents, strains and plasmids

Lysogeny broth (LB) medium (1% w/v tryptone, 0.5% w/v yeast extract, 0.5% w/v NaCl) and agar were from Acumedia. Antibiotics and L-arabinose were from Calbiochem. Isopropyl-β-D-thiogalactopyranoside (IPTG) was from Bio-Lab. Restriction enzymes, T4 polynucleotide kinase, Antarctic phosphatase and Quick Ligation Kit were from New England Biolabs. KAPA HiFi HotStart Ready Mix was from Kapa Biosystems. Taq DNA polymerase was from LAMDA Biotech. NucleoSpin Gel and PCR Clean-Up Kit were from Geneaid. The bacterial strains, plasmids and oligonucleotides used in this study are listed in Supplementary Table 1.

Plasmid construction

Plasmids were constructed using standard molecular biology techniques according to the manufacturers' instructions.

pCas1+2R plasmid encodes Cas1, Cas2 and a type I-E CRISPR array of a leader and a single repeat. pCas1+2R was constructed by amplifying the leader and repeat sequences of array I from the BL21-AI genome using oligonucleotides OA1F and OA1R (Supplementary Table 1). The amplified DNA was digested by *XbaI* and *SpeI* and ligated to *XbaI*-linearized pCas1+2 (Supplementary Table 1). The ligation yielded pCas1+2R plasmid that was further sequenced to exclude mutations introduced during cloning. The various mutant repeat plasmids were constructed using bidirectional PCR (repeats S2-S7, D1, D2, D4-D8, D10, and I1-I7) or site-directed mutagenesis (repeats S1, D3, D9, I8, D5S1-D5S6, and I4S1) methods. Plasmids constructed by bidirectional PCR were amplified using oligonucleotide pairs facing opposite directions followed by phosphorylation and self-ligation. Plasmids constructed by site-directed mutagenesis utilized complementary oligonucleotide pairs, each carrying the desired mutation with 15 bases of homologous sequence on both sides. Supplementary Table 2 lists the oligonucleotide combinations used to construct the various plasmids. Newly constructed plasmids were introduced into *E. coli* strain NEB5α by electroporation and sequenced to verify that the desired mutation was obtained. Once verified, the plasmids were purified from strain NEB5α and introduced into *E. coli* strain IYB5283 (Supplementary Table 1).

PCR products for deep sequencing

DNA of bacterial cultures subjected to acquisition assay was amplified in two consecutive PCRs termed PCR1 and PCR2. In PCR1, the reaction contained 25 μL Taq 2X Master Mix, 1.25 μL of 10 mM OA2F and OA2R primers (Supplementary Table 1), 5 μL bacterial culture and 16.5 μL double-distilled water. The PCR started with 3 min at 95°C followed by 35 cycles of 20 s at 95°C, 20 s at 55°C and 20 s at 72°C. The final extension step at 72°C was performed for 5 min. Part of the PCR1 content (20 μL) was purified using the DNA Clean-Up Kit and used for standard library preparation procedures followed by deep sequencing (MiSeq), while the remainder (30 μL) was loaded on a 1.8% (w/v) agarose gel and electrophoresed for 120 min at 120 V. Following electrophoresis, the expanded band was excised from the gel and purified using the DNA Clean-Up Kit. The extracted band served as the template for PCR2 aimed at amplifying the expanded CRISPR-array products. PCR2 contained 15 μL Taq 2X Master Mix, 0.5 μL of 10 mM RE10RD and IY230R1 primers (Supplementary Table 1), 2 ng of the gel-extracted DNA from PCR1 and double-distilled water to 20 μL. The PCR2 cycling program was identical to that of PCR1. The entire PCR2 content was loaded on a 1.8% agarose gel, electrophoresed, excised and purified from the gel under the same conditions as in PCR1, and used for standard library preparation procedures followed by deep sequencing (NextSeq500).