

Aberystwyth University

RapGene

Zampini, Massimiliano; Rees Stevens, Pauline; Pachebat, Justin; Kingston-Smith, Alison; Mur, Luis; Hayes, Finbarr

Published in:
Scientific Reports

DOI:
[10.1038/srep11302](https://doi.org/10.1038/srep11302)

Publication date:
2015

Citation for published version (APA):
Zampini, M., Rees Stevens, P., Pachebat, J., Kingston-Smith, A., Mur, L., & Hayes, F. (2015). RapGene: a fast and accurate strategy for synthetic gene assembly in Escherichia coli. *Scientific Reports*, 5, Article 11302.
<https://doi.org/10.1038/srep11302>

Document License CC BY

General rights

Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400
email: is@aber.ac.uk

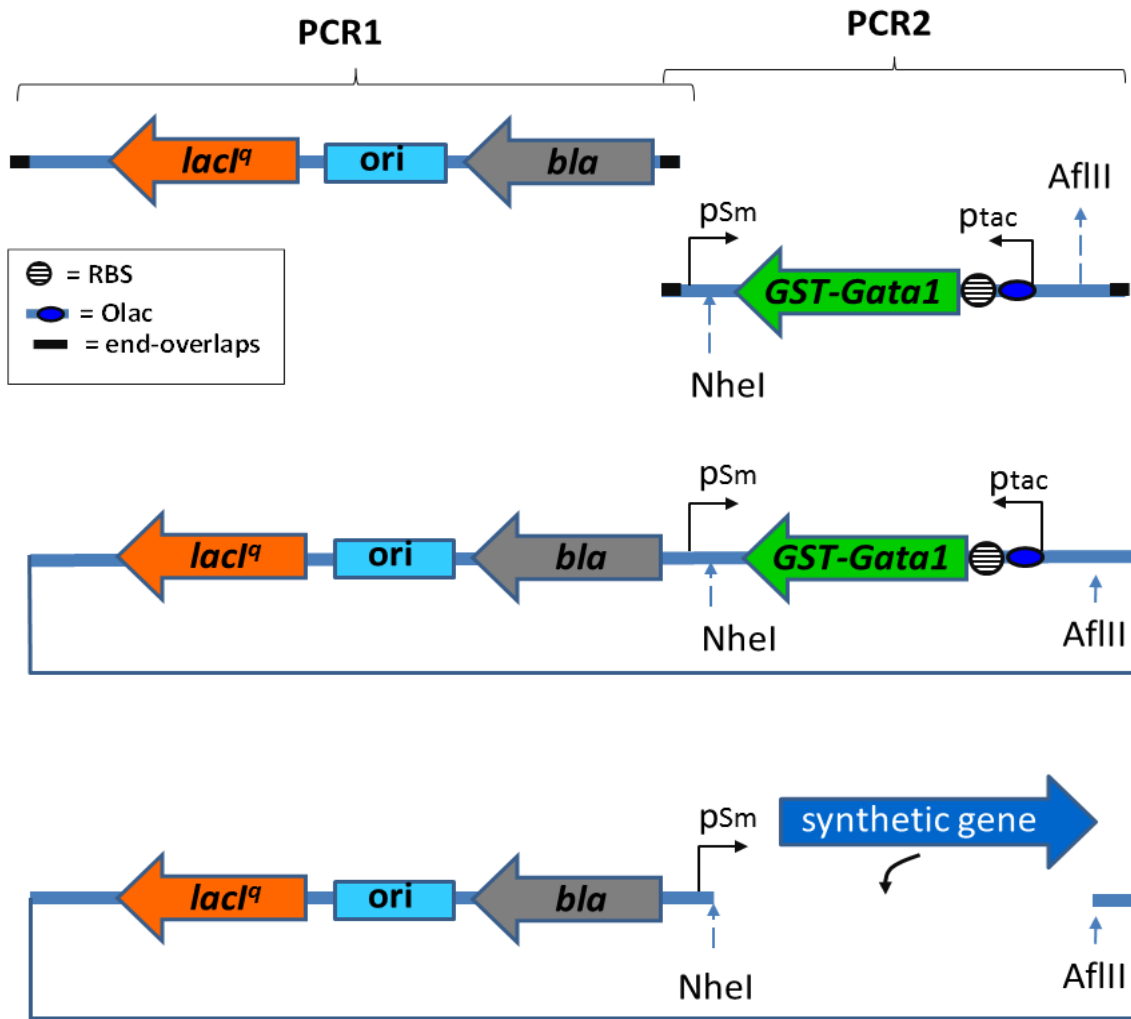
Supplementary Information

RapGene: a fast and accurate strategy for synthetic gene assembly

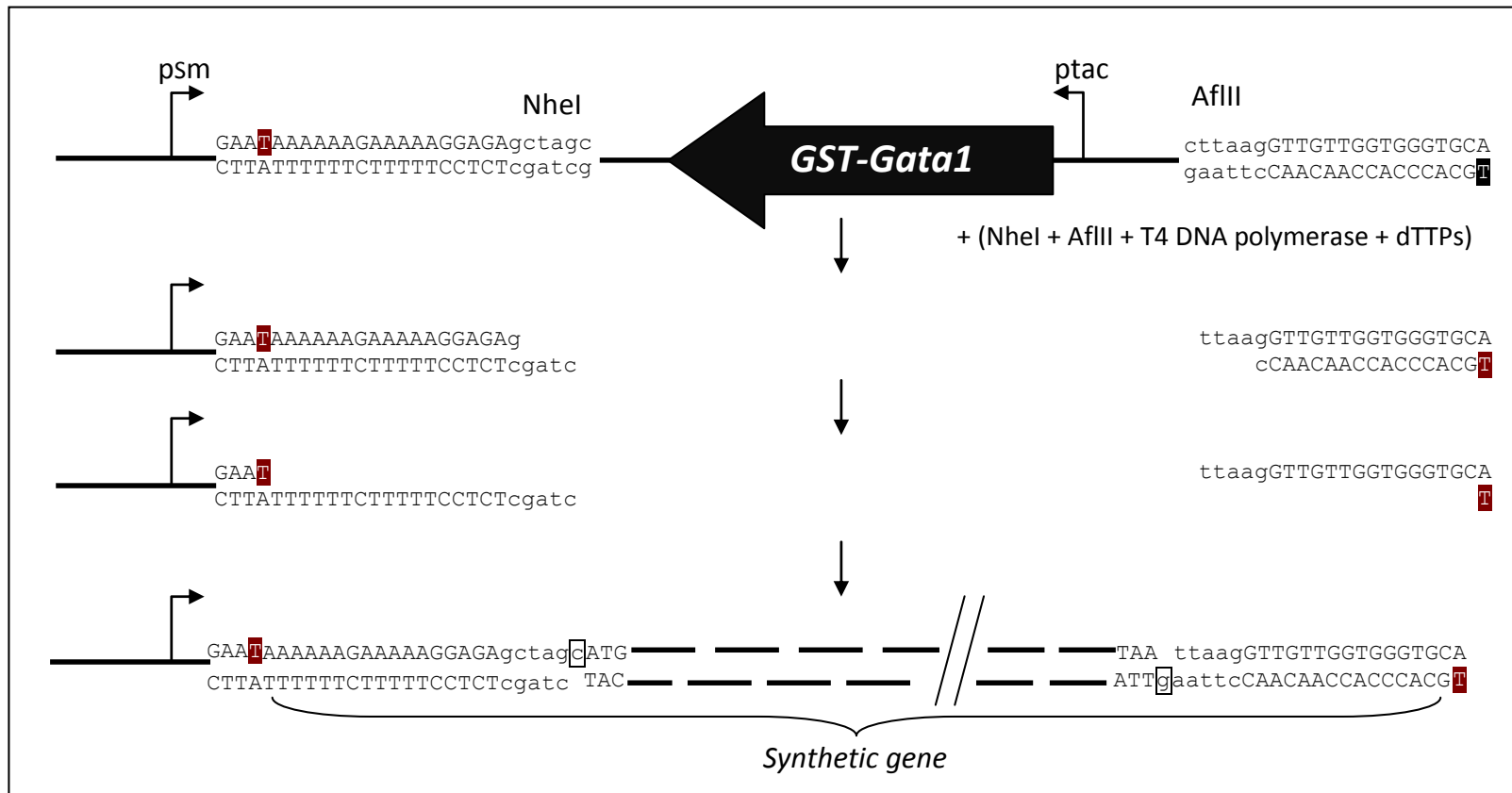
in *Escherichia coli*

Massimiliano Zampini, Pauline Rees Stevens, Justin A. Pachebat, Alison Kingston-Smith, Luis A. J.

Mur & Finbarr Hayes



Supplementary Figure S1. pRG1.0 assembly and structure (not to scale). Plasmid pRG1.0 was assembled from two PCR products derived from pGATA using the Gibson Assembly Cloning Kit. pSm, streptomycin/spectinomycin promoter from pCDFDuet-1. *ptac*, *tac* promoter. RBS, ribosome binding site. *Olac*, *lac* operator. *bla*, beta-lactamase gene. *ori*, origin of replication. *lacI^q*, *lac* repressor gene with a modified promoter. *NheI*, restriction site for *NheI*. *AflIII*, restriction site for *AflIII*. End-overlaps, homology regions required by the Gibson assembly cloning kit.



Supplementary Figure S2. pRG1.0 cloning region (not to scale). Plasmid pRG1.0 is prepared to generate long 3'-recessed cohesive ends by digesting with AflIII and NheI in the presence of T4 DNA polymerase and 1 mM dTTPs. The reaction is stopped by spin-column purification. The single-nucleotide gap (boxed on the complementary strand) between the assembled synthetic insert and the vector, produced in all inserts after annealing, is filled *in vivo* in *E. coli*. Dashed lines represent the single oligonucleotides within the synthetic gene.

Name	Sequence (5'-3')	Restriction sites	Signals	Notes
PCR1f	ACAGCTCActtaagGTTGTTGGTGGGTGCA GCCGGAAGCATAAAGTGT AAAGCC	AfIII		Template for PCR: pGATA
PCR1r	CCTAGGTTTATCAGGGTTATTGTC TCATGAGCGGATACATA TTTGAAT GTATTTAGAAAAATACCGGTACCTCTGACACATGCAGCTCCCGGAG		- promoter (-35 and -10 regions)	Template for PCR: pGATA
PCR2f	TCATGAGACAAT AACCTGATAAACCTAGG AATAAAAAAGAAAAAGG AGAgctagcTTTCACCGTCATCACCGAAACGCGCGAGGCAGATCGTCA G	NheI	- promoter (-10 region)	Template for PCR: pGATA
PCR2r	TGCACCCACCAACAACccttaagTGAGCTGT TGACAATTAATCATCGGCT CG	AfIII		Template for PCR: pGATA
Sm-1	AAAAAAGAAAAAGGAGAgctagcATGAGGGAAGCGGTGATCGCCGAAG TATCGACTCAAC	NheI		top strand
Sm-2	GACGCCAACTACCTCTGATAGTTGAGTCGATACTTCGGCGATCACCG CTTCCCTCAT			bottom strand
Sm-3	TATCAGAGGTAGTTGGCGTCATCGAGCGCCATCTCGAACCGACGTT GCTGGCCGTACA			top strand
Sm-4	CCACTGCGGAGCCGTACAAATGTACGGCCAGCAACGTCGGTTCGAG ATGGCGCTCGAT			bottom strand
Sm-5	TTTGTACGGCTCCGCAGTGGATGGCGGCCTGAAGCCACACAGTGAT ATTGATTTGCTG			top strand
Sm-6	AGCCTTACGGTCACCGTAACCAGCAAATCAATATCACTGTGTGGCTT CAGGCCGCCAT			bottom strand
Sm-7	GTTACGGTGACCGTAAGGCTTGATGAAACAACGCGCGAGCTTTGAT CAACGACCTTT			top strand
Sm-8	AGGGGAAGCCGAAGTTTCCAAAAGGTCGTTGATCAAAGCTCGCCGC GTTGTTTCATCA			bottom strand
Sm-9	TGGAAACTTCGGCTTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGT AGAAGTCACCAT			top strand
Sm-10	TGATGTCGTCGTGCACAACAATGGTGACTTCTACAGCGCGGAGAATC TCGCTCTCTCC			bottom strand
Sm-11	TGTTGTGCACGACGACATCATTCCGTGGCGTTATCCAGCTAAGCGCG AACTGCAATTT			top strand

Sm-12	TCATTGCGCTGCCATTCTCCAAATTGCAGTTCGCGCTTAGCTGGATA ACGCCACGGAA			bottom strand
Sm-13	GGAGAATGGCAGCGCAATGACATTCTTGCAGGTATCTTCGAGCCAG CCACGATCGACA			top strand
Sm-14	CAGCAAGATAGCCAGATCAATGTCGATCGTGGCTGGCTCGAAGATAC CTGCAAGAATG			bottom strand
Sm-15	TTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAACATAGCGTTGCC TTGGTAGGTCC			top strand
Sm-16	CAAAGAGTTCCTCCGCGCTGGACCTACCAAGGCAACGCTATGTTCT CTTGCTTTTGT			bottom strand
Sm-17	AGCGGCGGAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTATTTG AGGCGCTAAAT			top strand
Sm-18	TTCCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAGATCCTGTTCA GGAACCGGAT			bottom strand
Sm-19	GAAACCTTAACGCTATGGAACCTCGCCGCCGACTGGGCTGGCGATG AGCGAAATGTAG			top strand
Sm-20	AATGCGGGACAACGTAAGCACTACATTTGCTCATCGCCAGCCCAGT CGGGCGGCGAG			bottom strand
Sm-21	TGCTTACGTTGTCCCGCATTTGGTACAGCGCAGTAACCGGCAAAATC GCGCCGAAGGA			top strand
Sm-22	TTGCCAGTCGGCAGCGACATCCTTCGGCGCGATTTTGCCGGTTACT GCGCTGTACCA			bottom strand
Sm-23	TGTCGCTGCCGACTGGGCAATGGAGCGCCTGCCGGCCCAGTATCAG CCCGTCATACTT			top strand
Sm-24	AGATAAGCCTGTCTAGCTTCAAGTATGACGGGCTGATACTGGGCCG GCAGGCGCTCCA			bottom strand
Sm-25	GAAGCTAGACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTC GCGCGCAGATC			top strand
Sm-26	GTGGACAAATCTTCCAACCTGATCTGCGCGCGAGGCCAAGCGATCTT CTTCTTGTTCCA			bottom strand
Sm-27	AGTTGGAAGAATTTGTCCACTACGTGAAAGGCGAGATCACCAAGGTA GTCGGCAAATAA		- <i>aadA1</i> stop codon	top strand
Sm-28	GCACCCACCAACAACccttaagTTATTTGCCGACTACCTTGGTGTCTCG CCTTTCACGTA	AfIII	- <i>aadA1</i> stop codon	bottom strand
Sm-A1	GCACCCACCAACAACccttaagCAGCAAATCAATATCACTGTGTGGCTTC	AfIII		bottom strand

	AGGCCGCCAT			
Sm-A2	GCACCCACCAACAACcctaagAAATTGCAGTTCGCGCTTAGCTGGATAA CGCCACGGAA	AflII		bottom strand
Sm-A3	GCACCCACCAACAACcctaagATTTAGCGCCTCAAATAGATCCTGTTCA GGAACCGGAT	AflII		bottom strand
Sm-A4	GCACCCACCAACAACcctaagAAGTATGACGGGCTGATACTGGGCCGG CAGGCGCTCCA	AflII		bottom strand
GFP-1	AAAAAAGAAAAAGGAGAgctagcATGAGTAAAGGAGAAGAACTTTTCAC TGGAGTTGTCCCAATTC	NheI		top strand
GFP-2	TTTGCCCATTAACATCGCCATCTAATTCAACAAGAATTGGGACAACCTC CAGTGAAAAGTTCTTCTCCTTTACTCAT			bottom strand
GFP-3	TTGTTGAATTAGATGGCGATGTTAATGGGCAAAAATTCTCTGTCAGTG GAGAGGGTGAAGGTGATG			top strand
GFP-4	AAATAAATTTAAGGGTAAGTTTTCCGTATGTTGCATCACCTTCACCCT CTCCACTGACAGAGAATT			bottom strand
GFP-5	CAACATACGGAAAACCTTACCCTTAAATTTATTTGCACTACTGGGAAGC TACCTGTTCCATGGCCAA			top strand
GFP-6	ATTGAACACCATAAGAGAAAGTAGTGACAAGTGTGGCCATGGAACA GGTAGCTTCCCAGTAGTGC			bottom strand
GFP-7	CACTTGTCACTACTTTCTCTTATGGTGTTCATGCTTTTTCAAGATACC CAGATCATATGAAACAGC			top strand
GFP-8	AACCTTCGGGCATGGCACTCTTGAAAAGTCATGCTGTTTCATATGAT CTGGGTATCTTGAAAAGC			bottom strand
GFP-9	ATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGA ACTATATTTTACAAAGATG			top strand
GFP-10	ACTTGACTTCAGCACGTGTCTTGTAGTTCCTCGTCATCTTTGTAAAATA TAGTTCTTTCCTGTACAT			bottom strand
GFP-11	ACGGGAAC TACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACC CTTGTTAATAGAATCGAGT			top strand
GFP-12	TGTTTCCATCTTCTTTAAAATCAATACCTTTTAACTCGATTCTATTAAC AAGGGTATCACCTTCAA			bottom strand
GFP-13	TAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAAA TGGAATACAAC TATAACT			top strand
GFP-14	TTGGTTTGTCTGCCATGATGTATACATTATGTGAGTTATAGTTGTATTC CATTTTGTGTCCAAGAA			bottom strand

GFP-15	<u>CACATAATGTATACATCATGGCAGACAAACCAAAGAATGGAATCAAA</u> <u>GTAACTTCAAATTAGAC</u>			top strand
GFP-16	<u>CTGCTAATTGAACGCTTCCATCTTTAATGTTGTGTCTAATTTTGAAGTT</u> <u>AACTTTGATTCCATTCT</u>			bottom strand
GFP-17	<u>ACAACATTAAGATGGAAGCGTTCAATTAGCAGACCATTATCAACAAA</u> <u>ATACTCCAATTGGCGATG</u>			top strand
GFP-18	<u>ACAGGTAATGGTTGTCTGGTAAAAGGACAGGGCCATCGCCAATTGGA</u> <u>GTATTTTGTGATAATGGT</u>			bottom strand
GFP-19	<u>GCCCTGTCCTTTACCAGACAACCATTACCTGTCCACACAATCTGCC</u> <u>CTTTCCAAAGATCCCAACG</u>			top strand
GFP-20	<u>CAAACCTCAAGAAGGATCATGTGATCTCTCTTTTCGTTGGGATCTTTGG</u> <u>AAAGGGCAGATTGTGTGG</u>			bottom strand
GFP-21	<u>AAAAGAGAGATCACATGATCCTTCTTGAGTTTGTAACAGCTGCTGGG</u> <u>ATTACACATGGCATGGATGAACTATACAAATAA</u>		<i>gfp</i> stop codon	top strand
GFP-22	<u>GCACCCACCAACAACccttaagTTA</u> <u>TTTGTATAGTTCATCCATGCCATGTG</u> <u>TAATCCCAGCAGCTGTTA</u>	AfIII	<i>gfp</i> stop codon	bottom strand
RG1.0f	<u>CTCCGGGAGCTGCATG</u>			
RG1.0r	<u>GGCACGACAGGTTTCCCG</u>			

Supplementary Table S1. Oligonucleotides used in this study. Sequences designed to anneal to the DNA template in PCR are underlined. Boxed regions correspond to the homology sequences required by the Gibson Assembly Cloning Kit to produce pRG1.0. Sequences highlighted in black represents signals (i.e. promoters, stop codons). Restriction sites are indicated in lower case.

<i>aadA1</i>-assembly					
Type of assembly	<i>E. coli</i> strain	sample	Cycle 0	Cycle 1	Cycle 2
Oligonucleotides annealed simultaneously	NEB5 α	<i>aadA1</i> (6-oligonucleotides)	100%	100%	90%
Oligonucleotides annealed simultaneously	NEB5 α	<i>aadA1</i> (12-oligonucleotides)	100%	100%	100%
Oligonucleotides annealed simultaneously	NEB5 α	<i>aadA1</i> (18-oligonucleotides)	80%	70%	70%
Oligonucleotides annealed simultaneously	NEB5 α	<i>aadA1</i> (24-oligonucleotides)	-	20%	-
Oligonucleotides annealed simultaneously	NEB5 α	<i>aadA1</i> (28-oligonucleotides)	-	20%	10%
<i>gfp</i>-assembly					
Type of assembly	<i>E. coli</i> strain	sample	Cycle 0	Cycle 1	Cycle 2
Oligonucleotides annealed simultaneously	NEB5 α	<i>gfp</i> (22-oligonucleotides)	44%	20% 15% 15%	-
Oligonucleotides pre-annealed separately in three groups	NEB5 α	<i>gfp</i> (22-oligonucleotides)	-	20% 15% 10%	-

Supplementary Table S2. RapGene *aadA1* and *gfp* cloning efficiencies during the first two freezing-thawing cycles of the oligonucleotides used for the corresponding assemblies. Percentages of correct-size inserts as determined by colony PCR are reported. The table also reports yields for the assembly performed with three groups of consecutive oligonucleotides pre-annealed separately.