

TECHNICAL NOTES

Open Access



Lambda Red recombinase-mediated integration of the high molecular weight DNA into the *Escherichia coli* chromosome

Mario Juhas* and James W. Ajioka

Abstract

Background: *Escherichia coli* K-12 is a frequently used host for a number of synthetic biology and biotechnology applications and chassis for the development of the minimal cell factories. Novel approaches for integrating high molecular weight DNA into the *E. coli* chromosome would therefore greatly facilitate engineering efforts in this bacterium.

Results: We developed a reliable and flexible lambda Red recombinase-based system, which utilizes overlapping DNA fragments for integration of the high molecular weight DNA into the *E. coli* chromosome. Our chromosomal integration strategy can be used to integrate high molecular weight DNA of variable length into any non-essential locus in the *E. coli* chromosome. Using this approach we integrated 15 kb DNA encoding sucrose catabolism and lactose metabolism and transport operons into the *fliK* locus of the flagellar region 3b in the *E. coli* K12 MG1655 chromosome. Furthermore, with this system we integrated 50 kb of *Bacillus subtilis* 168 DNA into two target sites in the *E. coli* K12 MG1655 chromosome. The chromosomal integrations into the *fliK* locus occurred with high efficiency, inhibited motility, and did not have a negative effect on the growth of *E. coli*.

Conclusions: In addition to the rational design of synthetic biology devices, our high molecular weight DNA chromosomal integration system will facilitate metabolic and genome-scale engineering of *E. coli*.

Keywords: Synthetic biology, *Escherichia coli*, Lambda Red recombinase, Chromosomal integration, High molecular weight DNA

Background

Escherichia coli K-12 is a gram-negative model bacterium used as a host for a number of synthetic biology and biotechnology applications [1–3]. Furthermore, *E. coli* is considered to be among the most promising chassis for the development of the minimal cell factories [4–7]. Novel approaches for introducing synthetic DNA, particularly the high molecular weight DNA, into *E. coli* would therefore greatly facilitate engineering efforts in this bacterium.

Integration of synthetic DNA into the *E. coli* chromosome mitigates against many problems associated with the maintenance of DNA on plasmids or bacterial

artificial chromosomes (BACs) [8–11]. Chromosomal integration avoids complications arising from issues such as plasmid segregation or plasmid-maintenance associated metabolic burden. Furthermore, maintenance of plasmids and BACs in the cell requires constant antibiotic selection pressure [8–11]. Recent *E. coli* chromosomal integration approaches include the bacteriophage integrase-mediated recombination between phage attachment (*att*) sites dubbed clonetegration [12] and the bacteriophage lambda Red recombinase-mediated recombination employing yeast homing mitochondrial I-SceI endonuclease [13, 14], knock-in/knock-out (KIKO) vectors [15] and plasmid pSB1K3(FRTK) [16].

The chromosomal integration approaches differ in their flexibility, speed and the size of the DNA fragments they can integrate into the chromosome. While there

*Correspondence: mj417@cam.ac.uk
Department of Pathology, University of Cambridge, Tennis Court Road,
Cambridge CB2 1QP, UK

had been studies describing high molecular weight DNA integration into the chromosomes of some bacteria, such as *Bacillus subtilis* [17–19], this was not reported in *E. coli* yet. The currently used lambda Red recombination methods relying on a selection with a single antibiotic gene flanked by FLP recombinase target (FRT) sites are time-consuming due to the necessity to flip-out the antibiotic resistance gene in order to use it again. Furthermore, the traditional lambda Red recombining employing antibiotic cassette flanked by FRT sites leaves a single FRT site in the chromosome after flipping-out of the antibiotics gene [20]. The undesired recombination between the two FRT sites (first on the introduced DNA fragment and second in the chromosome) makes the repeated use of the cassettes flanked with FRT sites unsuitable for integration of large DNA sequences into the same target locus. Plasmid pSB1K3(FRTK), clonegration, KIKO-vectors, and yeast homing mitochondrial I-SceI endonuclease-based methods were shown to successfully integrate up to 10 kb DNA fragments into the *E. coli* chromosome [12–15].

Here, we developed a reliable and flexible lambda Red recombinase-based approach that utilizes overlapping DNA fragments for integration of the high molecular weight DNA of virtually any length into any non-essential target site in the *E. coli* chromosome. We demonstrate the utility of this method by integrating 15 kb DNA encoding sucrose and lactose catabolism pathways and 50 kb of *B. subtilis* 168 DNA into the *E. coli* K12 MG1655 chromosome.

Results and discussion

fliK of the flagellar region 3b as the chromosomal integration target site

The potential integration target sites in the *E. coli* chromosome vary in their suitability for integration of synthetic DNA [16]. Genes involved in the biogenesis and regulation of the flagellum are considered to be particularly suitable target sites for integration of genetic circuits due to their high expression, good characterization and non-essentiality for *E. coli*. Flagellar genes are situated in the genomic regions frequently occupied by DNA-binding proteins (highly expressed extended protein occupancy domains, heEPODs) [21, 22]. Furthermore, unlike the traditionally used phage attachment (*att*) sites, flagellar genes are present and well-conserved in all commonly used industrial *E. coli* strains, such as MG1655, DH10B, BL21-DE3, and W3110 [4, 21, 22, 25].

Here, we chose to integrate high molecular weight DNA into the *fliK* locus of the *E. coli* K12 MG1655 flagellar region 3b. We selected the *fliK* locus as the integration target site because our previous analyses of the *E. coli*

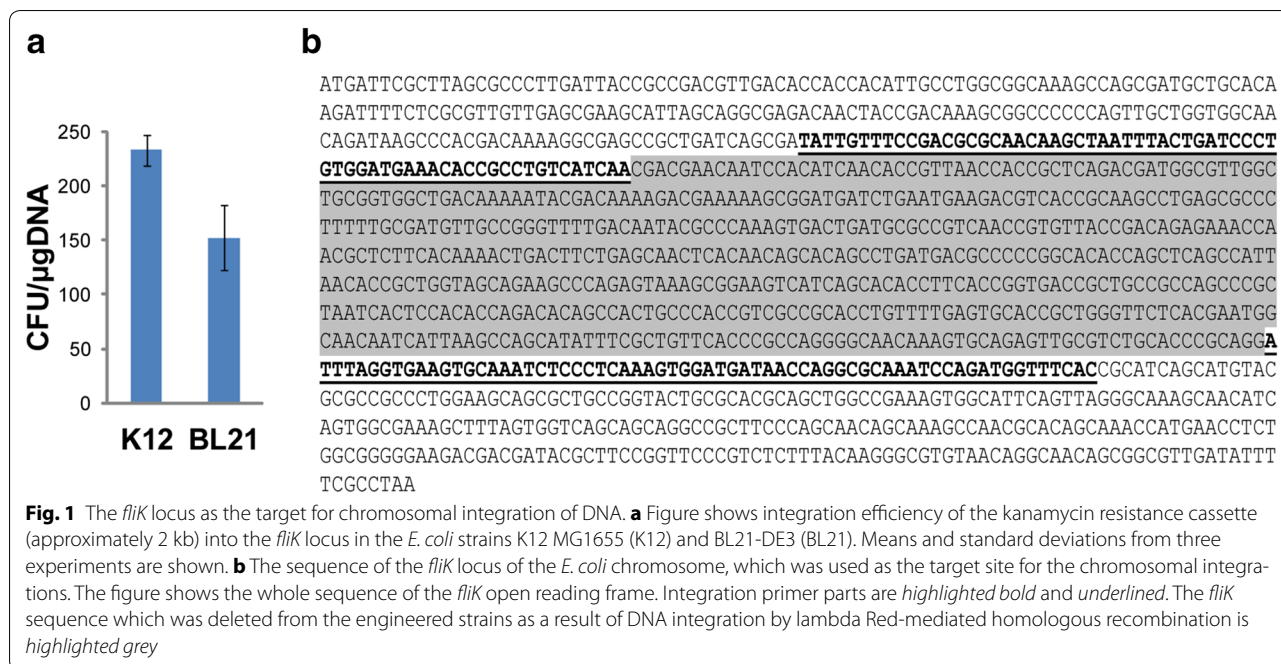
K12 MG1655 flagellar regions 1, 2, 3a, and 3b revealed that the flagellar region 3b supports the highest integration efficiency of all flagellar regions [16, 23, 24]. Furthermore, *fliK* was shown to support the highest integration efficiency from the *E. coli* K12 MG1655 flagellar region 3b.

Some commonly used industrial *E. coli* strains, such as BL21-DE3, are often causing problems by genomic engineering. This is usually due to the missing *att* sites used in many chromosomal integration approaches. Unlike *att* sites, the *fliK* locus in the strain BL21-DE3 has 100 percent homology with the strain K12 MG1655. To further verify the suitability of the *fliK* locus for integration of synthetic DNA, we compared the integration efficiency in strains BL21-DE3 and K12 MG1655. The efficiency of integration of the kanamycin resistance cassette (approximately 2 kb) into the *fliK* locus of the strain BL21-DE3 was high and comparable to that observed in the strain K12 MG1655 (Fig. 1a). This confirms the suitability of the *fliK* locus for integration of the synthetic DNA across variable *E. coli* strains. The exact position within the *fliK* open reading frame where the integrations occurred is shown in Fig. 1b.

Strategy of the high molecular weight DNA integration into the *E. coli* chromosome

Our strategy for the integration of the high molecular weight DNA molecules into the *E. coli* chromosome is outlined in Fig. 2. This approach utilizes the Red recombinase system of the bacteriophage lambda for integration of DNA fragments into the *E. coli* chromosome via recombination between homologous sequences [16]. Phage-based systems are a useful tool for genome engineering as they increase the otherwise very low rate of homologous recombination in some bacteria, such as *E. coli* [13, 15, 20, 25]. The lambda Red recombinase system employed in our approach is composed of three proteins Bet, Exo and Gam. Bet and Exo bind to the ends of the introduced DNA fragment and facilitate homologous recombination and integration into the chromosome. This is possible due to Gam-facilitated inhibition of the host RecBCD exonuclease V which, when active, degrades foreign DNA [20]. Lambda Red recombinase system in our approach is introduced into the host *E. coli* cells on plasmid pKM208. The expression of the lambda Red recombinase system from the lacZ promoter is tightly controlled by the presence or absence of Isopropyl β -D-1-thiogalactopyranoside (IPTG) in the growth medium [26].

DNA fragments to be integrated into the *E. coli* chromosome are first cloned into the plasmids p14kan and pJScav next to the kanamycin cassette (*kan*) and



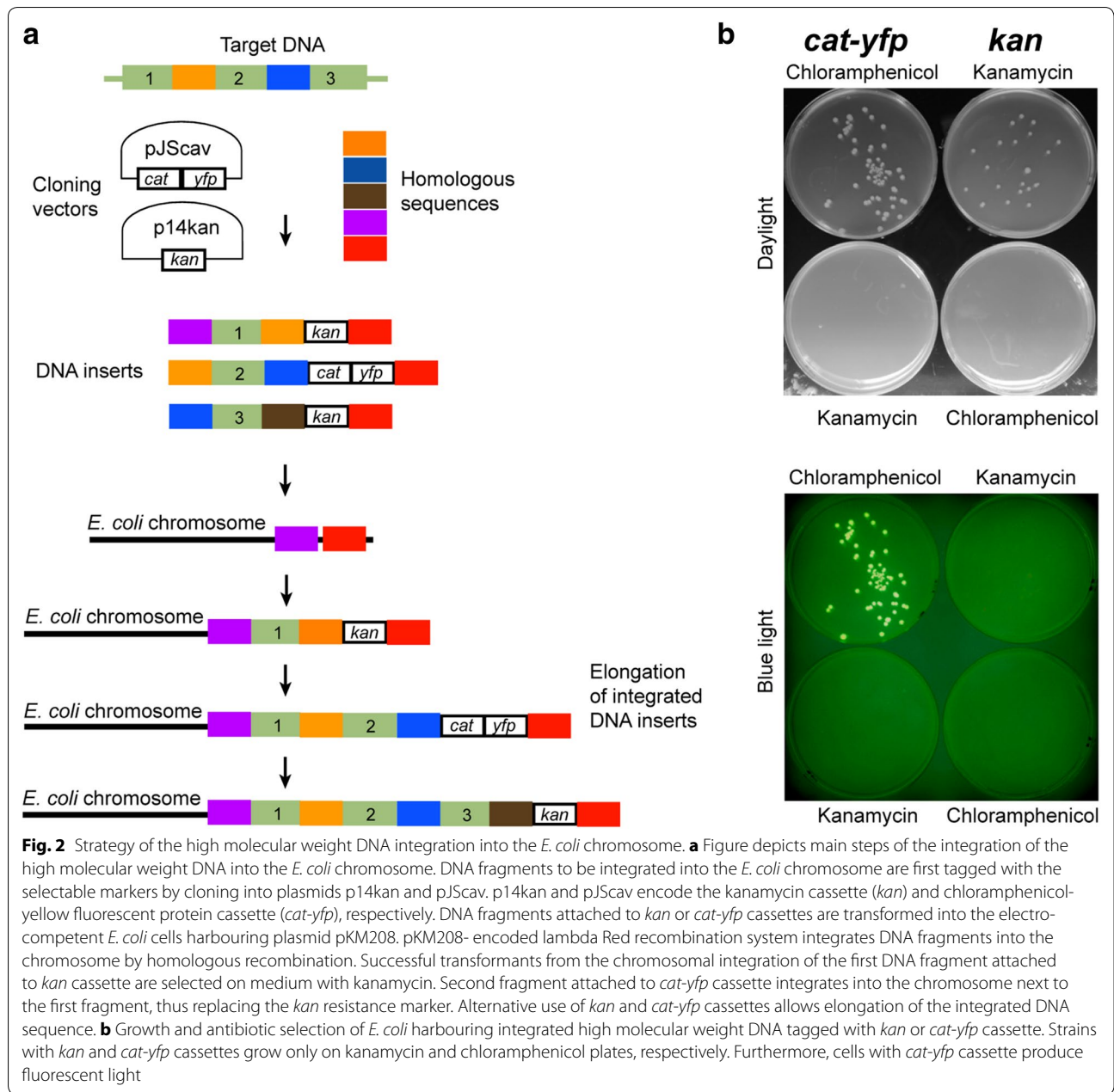
chloramphenicol-yellow fluorescent protein cassette (*cat-yfp*), respectively (Fig. 2a). Cloning of DNA fragments into p14kan and pJScav is facilitated by Gibson isothermal assembly approach [27, 28]. Next, DNA fragments are PCR amplified together with the attached *kan* or *cat-yfp* cassettes and sequences homologous to the target region in the chromosome. Sequences homologous to the target region in the *E. coli* chromosome (at least 60 bp in length) are introduced into the PCR amplified DNA fragments on primers. The length of the homology regions can be further increased by designing primers which generate longer overlapping regions between the DNA fragments for increased integration efficiency.

DNA fragments tagged with *kan* or *cat-yfp* cassettes are transformed into the electro-competent *E. coli* cells harbouring plasmid pKM208 with lambda Red recombination system. The pKM208- encoded lambda Red recombination system facilitates integration of DNA fragments into the *E. coli* chromosome (Fig. 2a). If the first DNA fragment is tagged with *kan* cassette, successful recombinants are selected by growing on kanamycin (Fig. 2b). The integration of the following DNA fragment tagged with *cat-yfp* cassette results in replacement of the *kan* resistance marker with *cat-yfp* cassette and successful recombinants are easily detectable by growth on chloramphenicol and yellow fluorescent protein (mVenus) expression (Fig. 2b). Elongation of the integrated DNA sequence is facilitated by the alternative use of *kan* and *cat-yfp* cassettes (Fig. 2a).

Integration of the sucrose and lactose catabolism pathways into the *E. coli* chromosome

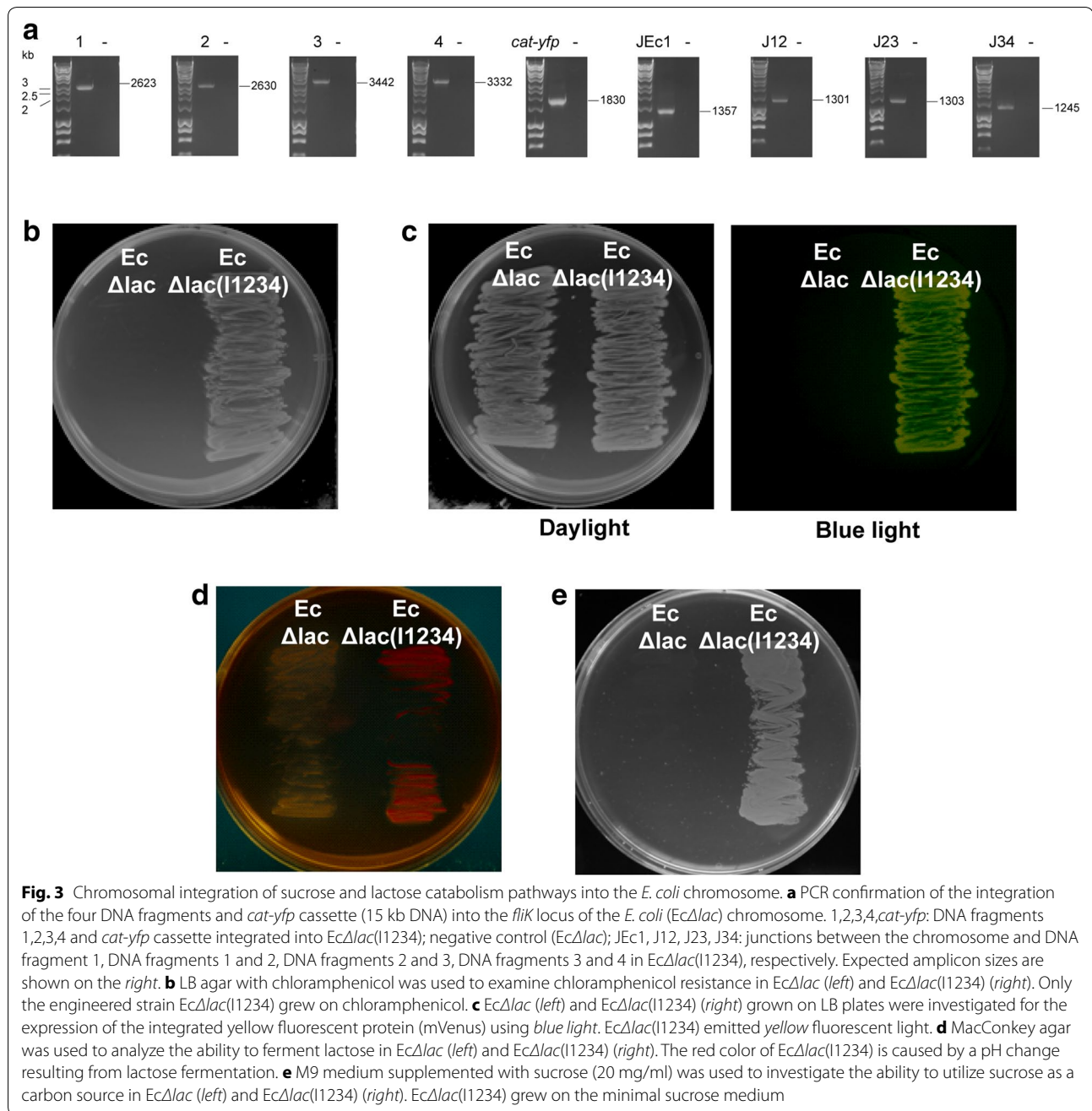
To demonstrate the utility of our system we first integrated sucrose and lactose catabolism pathways into the *E. coli* strain K12 MG1655 (*EcΔlac*) chromosome. *E. coli* K12 MG1655 can not normally utilize sucrose. We obtained DNA encoding the sucrose catabolism pathway from *E. coli* W, which is the only non-pathogenic *E. coli* strain that can utilize sucrose as a carbon source [15, 29].

Four DNA fragments (each approximately 3 kb in length) were first cloned into plasmids p14kan and pJScav for tagging with *kan* or *cat-yfp* cassettes (Fig. 2a). The first DNA fragment tagged with *kan* cassette was PCR amplified using primers harbouring sequences homologous to the *fliK* locus in the *EcΔlac* chromosome (Additional file 1) and electroporated into the host *E. coli* with pKM208-encoded lambda Red recombination system. The successful transformants were selected on plates with kanamycin (Fig. 2b). The second DNA fragment tagged with *cat-yfp* cassette was PCR amplified using forward primer with sequence homologous to the *fliK* locus and reverse primer with homology to the sequence of the first DNA fragment already integrated into the chromosome (Additional file 1). Electroporation of the second fragment into the host *E. coli* with pKM208-encoded lambda Red recombination system led to its integration into the chromosome and replacement of the *kan* resistance marker with *cat-yfp* cassette. Successful recombinants grew on plates with chloramphenicol and were



easily detectable by the yellow fluorescent protein (mVenus) expression (Fig. 2b). The alternative use of *kan* and *cat-yfp* cassettes led to the integration of all four DNA fragments into the *E. coli* chromosome. The integrated DNA has length of approximately 15 kb and encodes the whole sucrose catabolism and lactose metabolism and transport pathways, in addition to yellow fluorescent protein (mVenus) and chloramphenicol resistance-encoding open reading frames (Additional file 2). The integrations into the *E. coli* chromosome were verified by diagnostic PCR with flanking primers (Fig. 3a; Additional file 1)

and sequencing. PCR amplifications and sequencing across junctions showed that the DNA fragments were integrated into the *E. coli* chromosome in the correct order (Fig. 3a; Additional file 1). Furthermore, the chromosomal integration of the yellow fluorescent protein (mVenus), chloramphenicol resistance, sucrose and lactose catabolism pathways was confirmed phenotypically (Fig. 3b–e). The engineered strain *EcΔlac*(I1234) grew on chloramphenicol (Fig. 3b), emitted yellow fluorescent light (Fig. 3c) and fermented lactose (Fig. 3d) and sucrose (Fig. 3e).



Integration of the 50 kb *B. subtilis* DNA into the *E. coli* chromosome

We obtained high molecular weight DNA for cloning into the *E. coli* K12 MG1655 chromosome from *B. subtilis* 168. We chose *B. subtilis* due to high DNA sequence variability between *B. subtilis* and *E. coli* to avoid undesired recombination between the host and the donor DNA. Furthermore, the *B. subtilis* 168 DNA region chosen does

not harbour any known operons whose expression could be detrimental to the host *E. coli* (Additional file 3).

Seven DNA fragments (each approximately 6–7 kb in length) were first cloned into p14kan and pJScav (Fig. 2a). The first PCR amplified DNA fragment tagged with *kan* cassette and harbouring sequences homologous to the *fliK* locus (Additional file 1) was electroporated into host *E. coli* with pKM208-encoded lambda

Red recombination system. The successful transformants were selected on plates with kanamycin (Fig. 2b). The integration of the next DNA fragment tagged with *cat-yfp* cassette was detected by yellow fluorescent protein (mVenus) expression and growth on chloramphenicol (Fig. 2b). The alternative use of *kan* and *cat-yfp* cassettes led to the integration of all seven DNA fragments (with total length of approximately 50 kb) into the *E. coli* K12 MG1655 chromosome. The integrations into the *E. coli* K12 MG1655 chromosome were verified by diagnostic PCR with flanking primers (Fig. 4a; Additional file 1) and sequencing. PCR amplifications and sequencing across junctions showed that the DNA fragments were integrated into the *E. coli* K12 MG1655 chromosome in the correct order (Fig. 4a).

To assess the suitability of the *fliK* locus for integration of the high molecular weight DNA, we compared the integration efficiency of the *fliK* locus with another randomly selected integration target site in the *E. coli* chromosome, *cheW* (Fig. 4b). Integration efficiencies were determined from the number of colony forming units per microgram of DNA. This analysis showed that the *fliK* locus supports integration of the high molecular weight *B. subtilis* DNA with higher efficiency than the *cheW* locus (Fig. 4b).

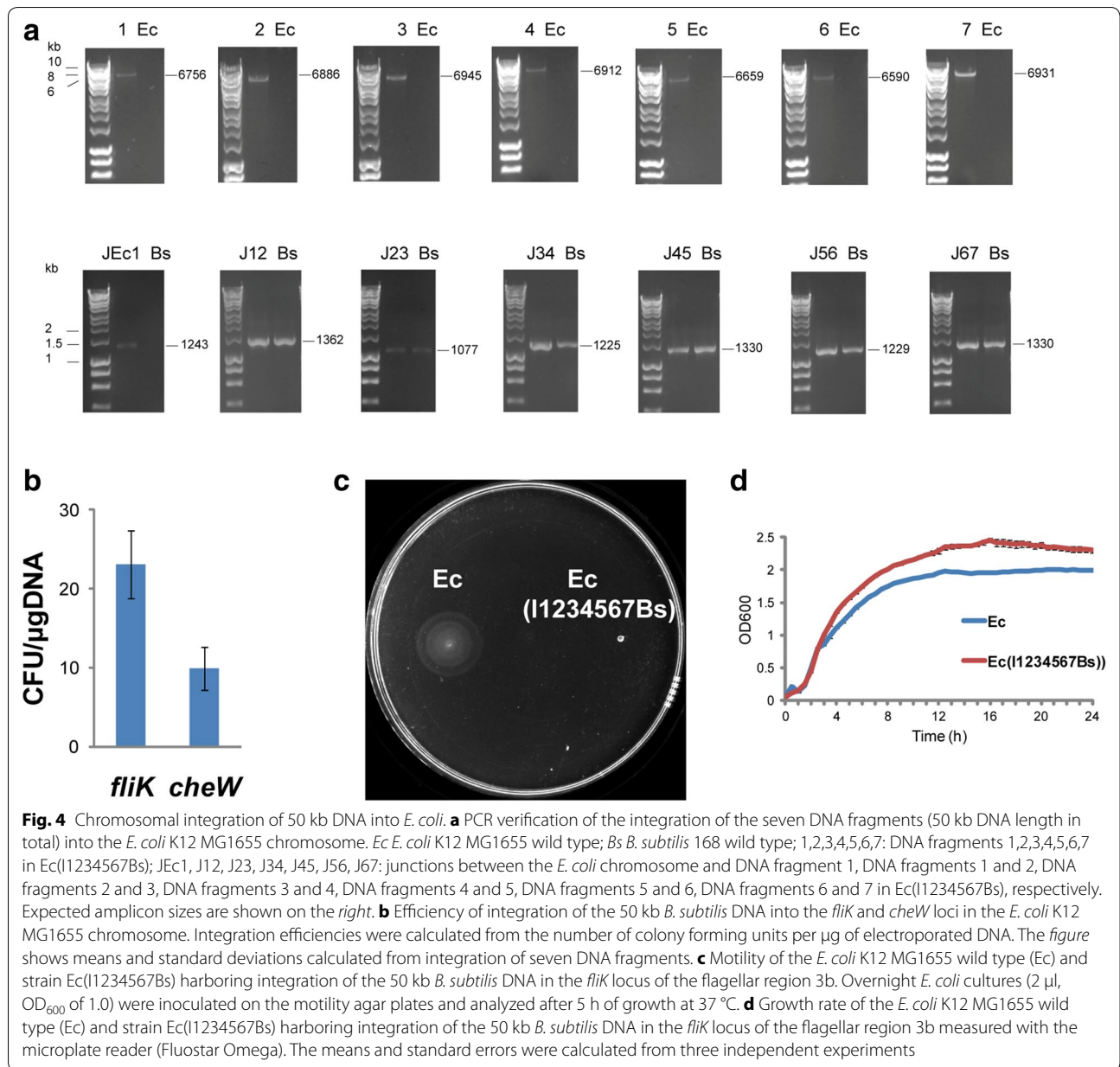
Genes encoding flagellar functions are important for the cell's motility. To verify that integration of the high molecular weight DNA into the *fliK* locus of *E. coli* K12 MG1655 impacts host cell's motility, we analyzed the engineered strain Ec(I1234567Bs) on the motility agar plate (Fig. 4c). The motility the strain Ec(I1234567Bs) was completely eliminated (Fig. 4c). Integrations of the high molecular weight DNA into the *E. coli* chromosome should not have a negative effect on the host cell's growth. To test the effect of chromosomal integration of the high molecular weight DNA on the growth rate, we used the engineered strains Ec(I1Bs), Ec(I12Bs), Ec(I123Bs), Ec(I1234Bs), Ec(I12235Bs), Ec(I123456Bs), and Ec(I1234567Bs) harbouring integrated DNA of increasing length (up to 50 kb). The cell growth was quantified by measuring absorbance of the bacterial cultures with the microplate reader (Fluostar Omega). The absorbance of the *E. coli* K12 MG1655 wild type was compared with that of the engineered strains Ec(I1Bs), Ec(I12Bs), Ec(I123Bs), Ec(I1234Bs), Ec(I12235Bs), Ec(I123456Bs), and Ec(I1234567Bs) grown at 37 °C for 24 h (Fig. 4d; Additional file 4). Integration of *B. subtilis* DNA fragments of various lengths into the *fliK* locus of the *E. coli* chromosome produced viable transformants whose growth was not impaired when compared with the wild type *E. coli* K12 MG1655. The strains Ec(I1Bs), Ec(I12Bs), Ec(I123Bs), Ec(I1234Bs), Ec(I12235Bs), Ec(I123456Bs), and Ec(I1234567Bs) harbouring integrated DNA grew

slightly better than the wild type *E. coli* K12 MG1655 (Fig. 4d; Additional file 4). Better growth of the strains with integrations into the *fliK* locus could be due to the lower metabolic burden to the cell resulting from the loss of the flagellar function. The *fliK* knockout strain grew only slightly better than the wild type *E. coli* K12 MG1655 (Additional file 4). This corroborates our previous findings, where similar growth pattern was observed for other mutants in the flagellum-encoding genes [16, 23, 24]. Good growth is a desirable feature for metabolic engineering and synthetic biology applications to generate more efficient and productive microbial cell factories.

The comparison of the integration efficiencies with another chromosomal integration target site, unimpaired growth and eliminated motility of the host cells show that the *fliK* locus of *E. coli* K12 MG1655 is suitable for chromosomal integration of the high molecular weight DNA.

Conclusions

We set out to provide the synthetic biology community with the reliable system for the integration of the high molecular weight DNA into the *E. coli* chromosome. To this end we developed a lambda Red recombinase-based method which utilizes overlapping DNA sequences for chromosomal integration of the high molecular weight DNA. Our chromosomal integration strategy is more flexible and allows integration of larger DNA molecules into the *E. coli* chromosome than the previous methods. Furthermore, the alternative use of *kan* and *cat-yfp* cassettes allows faster chromosomal integration process than the standard lambda Red recombineering methods, which rely on curing-out the lambda Red recombinase-encoding vector from the cell, before introducing second, FLP recombinase-encoding plasmid [20]. The length of the DNA fragments integrated into the *E. coli* chromosome by previous lambda Red recombinase-based approaches was about 10 kb [12–16]. The chromosomal integration strategy presented here can be used to integrate high molecular weight DNA of virtually any length into the *E. coli* chromosome. Furthermore, it does not target only a few integration sites, but allows DNA integration into any non-essential locus. Using this approach we integrated 15 kb DNA encoding sucrose and lactose catabolism pathways into the *fliK* locus of the flagellar region 3b and 50 kb *B. subtilis* DNA into two target sites in the *E. coli* K12 MG1655 chromosome. We showed that integration of 50 kb of the foreign DNA into the *fliK* locus inhibits motility of *E. coli*, but does not have a negative effect on the growth. Our chromosomal integration system will facilitate rational design of synthetic biology devices in *E. coli*. By allowing integration of the high molecular weight DNA, including whole operons and metabolic pathways into the *E. coli* chromosome, it



can be used for metabolic and genome-scale engineering of this bacterium. Furthermore, it can be easily adapted for use in other bacteria with the functional lambda Red recombination system.

Methods

Bacterial strains and growth conditions

Bacterial strains and plasmids used in this study are listed in Table 1. Luria-Bertani medium (LB) was used for the routine growth of bacteria. Bacteria cultivated in the liquid LB were grown on a rotatory shaker at 200 rpm

and 37, 30 or 42 °C, depending on the requirements of the experiments. Bacteria cultivated on LB agar (10 mg/ml) plates were grown at 37, 30 or 42 °C for 24 h. When necessary, LB medium was supplemented with kanamycin (50 $\mu\text{g}/\text{ml}$), ampicillin (100 $\mu\text{g}/\text{ml}$) or chloramphenicol (30 $\mu\text{g}/\text{ml}$) to grow *E. coli* and kanamycin (5 $\mu\text{g}/\text{ml}$) or chloramphenicol (5 $\mu\text{g}/\text{ml}$) to cultivate *B. subtilis*. Lactose fermentation experiments were performed using MacConkey agar (Sigma). The ability to utilize sucrose was investigated using M9 medium supplemented with thiamine (1 $\mu\text{g}/\text{ml}$) and sucrose (20 mg/ml) [15].

Table 1 Bacterial strains and plasmids used in this study

	Characteristics	Reference
Strains		
Ec	<i>E. coli</i> wild type strain K12 MG1655	[34]
Bs	<i>B. subtilis</i> Marburg wild type 168, <i>trpC2</i>	[35]
BL21	<i>E. coli</i> strain BL21-DE3	CGSC stock
EcΔ <i>fliK</i>	<i>E. coli</i> K12 MG1655 with deleted <i>fliK</i>	This study
EcΔ <i>lac</i>	<i>E. coli</i> K12 MG1655 with deleted <i>lac</i> operon	[14]
Ec(I1Bs)	Ec with <i>B. subtilis</i> DNA fragment 1 integrated	This study
Ec(I12Bs)	Ec with <i>B. subtilis</i> DNA fragments 1,2 integrated	This study
Ec(I123Bs)	Ec with <i>B. subtilis</i> DNA fragments 1,2,3 integrated	This study
Ec(I1234Bs)	Ec with <i>B. subtilis</i> DNA fragments 1,2,3,4 integrated	This study
Ec(I12345Bs)	Ec with <i>B. subtilis</i> DNA fragments 1,2,3,4,5 integrated	This study
Ec(I123456Bs)	Ec with <i>B. subtilis</i> DNA fragments 1,2,3,4,5,6 integrated	This study
Ec(I1234567Bs)	Ec with <i>B. subtilis</i> DNA fragments 1,2,3,4,5,6,7 integrated	This study
EcΔ <i>lac</i> (I1)	EcΔ <i>lac</i> with DNA fragment 1 integrated	This study
EcΔ <i>lac</i> (I12)	EcΔ <i>lac</i> with DNA fragments 1,2 integrated	This study
EcΔ <i>lac</i> (I123)	EcΔ <i>lac</i> with DNA fragments 1,2,3 integrated	This study
EcΔ <i>lac</i> (I1234)	EcΔ <i>lac</i> with DNA fragments 1,2,3,4 integrated	This study
Plasmids		
pJScav	<i>cat-yfp</i> cassette	[36]
p14kan	<i>kan</i> cassette	[16]
pKM208	plasmid with lambda Red system, IPTG inducible	[26]
pCP20	plasmid encoding flippase (FLP) recombinase	[20]
p14kan(I1Bs)	p14kan with <i>B. subtilis</i> DNA fragment 1	This study
pJScav(I2Bs)	pJScav with <i>B. subtilis</i> DNA fragment 2	This study
p14kan(I3Bs)	p14kan with <i>B. subtilis</i> DNA fragment 3	This study
pJScav(I4Bs)	pJScav with <i>B. subtilis</i> DNA fragment 4	This study
p14kan(I5Bs)	p14kan with <i>B. subtilis</i> DNA fragment 5	This study
pJScav(I6Bs)	pJScav with <i>B. subtilis</i> DNA fragment 6	This study
p14kan(I7Bs)	p14kan with <i>B. subtilis</i> DNA fragment 7	This study
p14kan(I1)	p14kan with DNA fragment 1	This study
pJScav(I2)	pJScav with DNA fragment 2	This study
p14kan(I3)	p14kan with DNA fragment 3	This study
pJScav(I4)	pJScav with DNA fragment 4	This study

DNA integration into the *E. coli* chromosome

Integrations of DNA fragments into the *E. coli* chromosome were performed employing previously described method [16]. Briefly, DNA fragments were first PCR amplified using primers harboring flanking sequences

homologous to the target region in the *E. coli* K12 MG1655 chromosome (at least 60 bp in length) and purified by gel extraction. As longer sequences greatly improve integration efficiency, the length of the homology regions can be further increased by designing primers generating longer overlapping regions between the DNA fragments. Plasmid pKM208 encoding lambda Red recombination system was introduced into *E. coli* K12 MG1655 by electroporation and successful transformants were selected on ampicillin at 30 °C. Overnight *E. coli* K12 MG1655 culture harboring pKM208 was inoculated into a fresh liquid LB medium (1:100 dilution) and grown to OD₆₀₀ of 0.2 at 30 °C. Then, 1 mM IPTG was added and bacteria were grown to the final OD₆₀₀ of 0.4–0.6. Bacteria were washed twice with glycerol (100 µl/ml) and resuspended in glycerol (100 µl/ml) (100 µl per 100 ml culture). Bacterial cells were then transformed with the gel purified DNA fragments by electroporation. In our setting, the best transformation efficiency was achieved when using 100–150 µl of the PCR product (approximately 5 µg of DNA). This volume was electroporated into 100 µl of the freshly prepared electro-competent cells harbouring pKM208. Transformants harboring the chromosomally integrated DNA fragments were selected on kanamycin or chloramphenicol plates at 37 °C. Successful chromosomal integrations were verified by diagnostic PCR with flanking primers and sequencing. In general, approximately 50 % of the transformants harbored correctly integrated DNA fragment. The cells were transformed newly with pKM208 after each round of integration to increase chromosomal integration efficiency. After the last round of integration the temperature sensitive plasmid pKM208 can be cured out by growing transformants at 42 °C. Chemically competent bacteria were prepared by the modified Hanahan method [30]. Electro-competent *E. coli* were prepared by the modified Miller and Nickoloff method [31]. The whole process for integration of a single DNA fragment into the *E. coli* chromosome using this system takes about 4 days. This is faster than the standard lambda Red recombineering methods which necessitate curing-out the lambda Red recombinase system- encoding vector from the cell, prior to introducing another, FLP recombinase-encoding, plasmid. On the first day, DNA fragment is cloned into pJScav or p14kan by Gibson Assembly and transformed into competent *E. coli* strain. On the second day, correctly assembled plasmids are identified by colony PCR, correctly assembled colonies are inoculated for plasmid preparation and the recipient *E. coli* harbouring pKM208-encoded lambda Red recombination system is inoculated into LB with ampicillin. On the third day, plasmids pJScav/p14kan with assembled DNA are isolated and sequenced, DNA fragments tagged with *cat-yfp*

or *kan* cassettes are PCR amplified and transformed into induced competent cells with pKM208. On the fourth day, the integrated DNA fragments are identified by colony PCR and quality controlled by sequencing and diagnostic PCR. The chromosomal integration of multiple DNA fragments can be speeded up by editing several fragments simultaneously (e.g. cloning of all fragments into pJScav and p14kan can be performed the same day).

DNA amplification and modification methodology

Dream Taq master mix kit (Thermo Scientific) or Phusion DNA polymerase (Thermo Scientific) were used to amplify DNA fragments, according to the protocol provided by the supplier. PCR amplifications of DNA fragments were routinely performed in 50 μ l reaction volumes. Oligonucleotide primers for PCR amplifications were synthesized by Sigma-Aldrich and Integrated DNA Technologies (IDT). DNA fragments were routinely assembled in a 5.2 μ l reaction volume [16] employing Gibson isothermal assembly [27, 28]. Successful assemblies of DNA fragments were confirmed using diagnostic PCR with flanking primers and sequencing. Sequencing was performed by Source Bioscience (Cambridge, UK). DNA fragments were routinely purified by extraction from the agarose (10 mg/ml) gel using Qiaquick Gel Extraction kit (Qiagen), according to the manufacturer's instructions. Plasmids were isolated with Qiaprep Spin Miniprep kit (Qiagen). *B. subtilis* genomic DNA was isolated using GeneJET genomic DNA purification kit (Thermo Scientific), according to the supplier's instructions.

Microplate reader measurement of the absorbance

Growth rate of bacterial strains was determined with the microplate reader (Fluostar Omega, BMG Labtech, UK). Briefly, overnight liquid cultures of the analyzed bacterial strains were diluted in a fresh LB medium to starting OD₆₀₀ of 0.05 and 200 μ l of these cultures were transferred into the wells of the 96 well microplates (flat-bottomed, clear, Sterilin Sero-Well, UK). The microplates were cultivated in the microplate reader at 37 °C for 24 h. Absorbance was determined with the automatically repeated protocol using following parameters: double orbital shaking at 500 r.p.m, absorbance filter 600 nm, cycle time: 30 min. The mean and standard errors were determined from three independent biological replicates.

Motility assessment

The motility of the bacterial strains was evaluated using the motility agar plates. The motility agar plates were prepared by pouring 100 ml of motility agar [tryptone (100 mg/ml), NaCl (50 mg/ml), Bacto-Agar (2.5 mg/ml) (Difco)] in the 13 cm plates and let to set overnight. 2 μ l

of the overnight bacterial cultures diluted to OD₆₀₀ of 1.0 were inoculated in the middle of the motility plates pre-warmed to 37 °C. Plates were incubated for 4–6 h at 37 °C prior to taking the pictures.

Databases and sequence analyses

The *B. subtilis* genome sequence was obtained from the SubtiWiki (<http://subtiwiki.uni-goettingen.de/>) and BioCyc (<http://bsubcyc.org/>) databases. The *E. coli* K-12 MG1655 genome sequence was obtained from the *E. coli* K-12 project website (<http://www.xbase.ac.uk/genome/escherichia-coli-str-k-12-substr-mg1655>). The sequences of the standard DNA parts were obtained from the Registry of Standard Biological Parts (http://parts.igem.org/Main_Page). Sequencing was performed by Source Bioscience (Cambridge, UK). Sequences were compared employing TBLASTX, BLASTN [32], and position-specific iterated BLAST (PSI-BLAST) [33] from the National Center for Biotechnology Information (NCBI) website (<http://ncbi.nlm.nih.gov>).

Additional files

Additional file 1. Primers used in this study.

Additional file 2. Integrated sucrose and lactose catabolism pathways. Figure shows open reading frames of the 15 kb DNA encoding sucrose catabolism and lactose metabolism and transport pathways integrated into the *flhK* locus of the *E. coli* K12 MG1655 (*EcΔlac*) chromosome. Csc: sucrose catabolism genes; Lac: lactose metabolism and transport genes; mVenus: yellow fluorescent protein; Cat: chloramphenicol.

Additional file 3. Integrated *B. subtilis* DNA. Figure shows open reading frames of the high molecular weight DNA from *B. subtilis* 168 integrated into the *E. coli* K12 MG1655 chromosome. Hypotheticals with unknown function are highlighted black. Oxred: oxidoreductase; Atransf: Acetyl transferase; Cys: cystathionine metabolism; Man: mannose metabolism; Sp: sporulation; PdaC: peptidoglycan metabolism; ThiS, ThaS: thiamine metabolism.

Additional file 4. Growth rates. Figure shows growth rates of the engineered strains Ec(11Bs), Ec(112Bs), Ec(1123Bs), Ec(11234Bs), Ec(11235Bs), Ec(1123456Bs), Ec(11234567Bs) and strain *EcΔflhK* compared to *E. coli* K12 MG1655 wild type (Ec) measured with the microplate reader (Fluostar Omega). The means and standard errors were calculated from three biological replicates. The integration into the *flhK* locus did not have a negative effect on the growth of the host *E. coli*.

Abbreviations

BAC: bacterial artificial chromosome; *cat-yfp*: chloramphenicol-yellow fluorescent protein cassette; heEPODs: highly expressed extended protein occupancy domains; IPTG: isopropyl β -D-1-thiogalactopyranoside; *kan*: kanamycin cassette; KIKO vector: knock-in/knock-out vector; LB medium: Luria-Bertani medium; FLP recombinase: flippase recombinase; FRT site: FLP recombinase target site.

Authors' contributions

Conceived and designed experiments: MJ, JWA. Performed experiments and bioinformatics analyses: MJ. Drafting of the manuscript: MJ. Critical revision of the manuscript: JWA. Obtaining funding: MJ, JWA. Both authors read and approved the final manuscript.

Acknowledgements

We would like to thank members of the laboratory for the constructive discussions.

Competing interests

The authors declare that they have no competing interests.

Availability of data and material

All data generated or analysed during this study are included in this published article and its supplementary information file.

Funding

This work was supported by grants from the UK Engineering and Physical Sciences Research Council, OpenPlant Fund and SynBio Fund.

Received: 13 June 2016 Accepted: 28 September 2016

Published online: 05 October 2016

References

- Ajikumar PK, Xiao WH, Tyo KE, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B, Stephanopoulos G. Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science*. 2010;330:70–4.
- Yim H, Haselbeck R, Niu W, Pujol-Baxley C, Burgard A, Boldt J, Khandurina J, Trawick JD, Osterhout RE, Stephen R, et al. Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat Chem Biol*. 2011;7:445–52.
- Zhou L, Niu DD, Tian KM, Chen XZ, Prior BA, Shen W, Shi GY, Singh S, Wang ZX. Genetically switched D-lactate production in *Escherichia coli*. *Metab Eng*. 2012;14:560–8.
- Juhas M, Reuß DR, Zhu B, Commichau FM. *Bacillus subtilis* and *Escherichia coli* essential genes and minimal cell factories after one decade of genome engineering. *Microbiology*. 2014;160:2341–51.
- Juhas M, Ajioka JW. High molecular weight DNA assembly in vivo for synthetic biology applications. *Crit Rev Biotechnol*. 2016;4:1–10.
- Juhas M. On the road to synthetic life: the minimal cell and genome-scale engineering. *Crit Rev Biotechnol*. 2015;36:1–8.
- Juhas M, Davenport PW, Brown JR, Yarkoni O, Ajioka JW. Meeting report: the Cambridge BioDesign TechEvent—Synthetic Biology, a new “Age of Wonder”? *Biotechnol J*. 2013;87:761–3.
- Cunningham DS, Koepsel RR, Ataai MM, Domach MM. Factors affecting plasmid production in *Escherichia coli* from a resource allocation standpoint. *Microb Cell Fact*. 2009;8:27.
- Marcellin E, Chen WY, Nielsen LK. Understanding plasmid effect on hyaluronic acid molecular weight produced by *Streptococcus equi* subsp. *zooepidemicus*. *Metab Eng*. 2010;12:62–9.
- Guan L, Liu Q, Li C, Zhang Y. Development of a Fur-dependent and tightly regulated expression system in *Escherichia coli* for toxic protein synthesis. *BMC Biotechnol*. 2013;13:25.
- Yadav VG, De Mey M, Lim CG, Ajikumar PK, Stephanopoulos G. The future of metabolic engineering and synthetic biology: towards a systematic practice. *Metab Eng*. 2012;14:233–41.
- St-Pierre F, Cui L, Priest DG, Endy D, Dodd IB, Shearwin KE. One-step cloning and chromosomal integration of DNA. *ACS Synth Biol*. 2013;2:537–41.
- Ublinskaya AA, Samsonov VV, Mashko SV, Stoynova NV. A PCR-free cloning method for the targeted $\phi 80$ Int-mediated integration of any long DNA fragment, bracketed with meganuclease recognition sites, into the *Escherichia coli* chromosome. *J Microbiol Methods*. 2012;89:167–73.
- Kuhlman TE, Cox EC. Site-specific chromosomal integration of large synthetic constructs. *Nucleic Acids Res*. 2010;38:e92.
- Sabri S, Steen JA, Bongers M, Nielsen LK, Vickers CE. Knock-in/Knock-out (KIKO) vectors for rapid integration of large DNA sequences, including whole metabolic pathways, onto the *Escherichia coli* chromosome at well-characterised loci. *Microb Cell Fact*. 2013;12:60.
- Juhas M, Evans LD, Frost J, Davenport PW, Yarkoni O, Fraser GM, Ajioka JW. *Escherichia coli* flagellar genes as target sites for integration and expression of genetic circuits. *PLoS ONE*. 2014;9:e111451.
- Itaya M, Fujita K, Kuroki A, Tsuge K. Bottom-up genome assembly using the *Bacillus subtilis* genome vector. *Nat Methods*. 2008;5:41–3.
- Yonemura I, Nakada K, Sato A, Hayashi J, Fujita K, Kaneko S, Itaya M. Direct cloning of full-length mouse mitochondrial DNA using a *Bacillus subtilis* genome vector. *Gene*. 2007;391:171–7.
- Itaya M, Tsuge K, Koizumi M, Fujita K. Combining two genomes in one cell: stable cloning of the *Synechocystis* PCC6803 genome in the *Bacillus subtilis* 168 genome. *Proc Natl Acad Sci USA*. 2005;102:15971–6.
- Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA*. 2000;97:6640–5.
- Vora T, Hottes AK, Tavazoie S. Protein occupancy landscape of a bacterial genome. *Mol Cell*. 2009;35:247–53.
- Kahramanoglou C, Seshasayee AS, Prieto AI, Ibberson D, Schmidt S, Zimmermann J, Benes V, Fraser GM, Luscombe NM. Direct and indirect effects of H-NS and Fis on global gene expression control in *Escherichia coli*. *Nucleic Acids Res*. 2011;39:2073–91.
- Juhas M, Ajioka JW. Identification and validation of novel chromosomal integration and expression loci in *Escherichia coli* flagellar region 1. *PLoS ONE*. 2015;10:e0123007.
- Juhas M, Ajioka JW. Flagellar region 3b supports strong expression of integrated DNA and the highest chromosomal integration efficiency of the *Escherichia coli* flagellar regions. *Microb Biotechnol*. 2015;8:726–38.
- Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol*. 2006;2(2):0008.
- Murphy KC, Campellone KG. Lambda Red-mediated recombinogenic engineering of enterohemorrhagic and enteropathogenic *E. coli*. *BMC Mol Biol*. 2003;4:11.
- Gibson D, Young L, Chuang R, Venter J, Hutchison Cr, Smith H. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods*. 2009;6:343–5.
- Merryman C, Gibson DG. Methods and applications for assembling large DNA constructs. *Metab Eng*. 2012;14:196–204.
- Archer CT, Kim JF, Jeong H, Park JH, Vickers CE, Lee SY, Nielsen LK. The genome sequence of *E. coli* W (ATCC 9637): comparative genome analysis and an improved genome-scale reconstruction of *E. coli*. *BMC Genom*. 2011;12:9.
- Hanahan D, Jessee J, Bloom FR. Plasmid transformation of *Escherichia coli* and other bacteria. *Methods Enzymol*. 1991;204:63–113.
- Miller EM, Nickoloff JA. *Escherichia coli* electro transformation. *Methods Mol Biol*. 1995;47:105–13.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215:403–10.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25:3389–402.
- Hayashi K, Morooka N, Yamamoto Y, Fujita K, Isono K, Choi S, Ohtsubo E, Baba T, Wanner BL, Mori H, Horiuchi T. Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110. *Mol Syst Biol*. 2006;2(2):0007.
- Burkholder PR, Giles NH. Induced biochemical mutations in *Bacillus subtilis*. *Am J Bot*. 1947;34:345–8.
- Juhas M, Ajioka JW. Integrative bacterial artificial chromosomes for DNA integration into the *Bacillus subtilis* chromosome. *J Microbiol Methods*. 2016;125:1–7.