Journal of Microbiological Methods 99 (2014) 27–34

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

Versatile broad-host-range cosmids for construction of high quality metagenomic libraries

Jiujun Cheng, Lee Pinnell, Katja Engel, Josh D. Neufeld, Trevor C. Charles *

Department of Biology, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada

ARTICLE INFO

Article history:
Received 10 December 2013
Received in revised form 22 January 2014
Accepted 25 January 2014
Available online 1 February 2014

Keywords:
Soil
DNA extraction
Functional metagenomics
Broad-host-range cosmids
Metagenomic libraries
Gateway® recombination vector

ABSTRACT

We constructed IncP broad-host-range Gateway® entry cosmids pJC8 and pJC24, which replicate in diverse *Proteobacteria*. We demonstrate the functionality of these vectors by extracting, purifying, and size-selecting metagenomic DNA from agricultural corn and wheat soils, followed by cloning into pJC8. Metagenomic DNA libraries of 8×10^4 (corn soil) and 9×10^6 (wheat soil) clones were generated for functional screening. The DNA cloned in these libraries can be transferred from these recombinant cosmids to Gateway® destination vectors for specialized screening purposes. Those library clones are available from the Canadian MetaMicroBiome Library project (http://www.cm2bl.org/).

1. Introduction

Soils host enormous taxonomic diversity (Curtis et al., 2002). Genomes of metabolically versatile soil microorganisms are potential sources of biocatalysts for various industrial applications. Metagenomics helps explore the collective genetic constituency of environmental microorganisms, even those that resist cultivation though conventional microbiological techniques. Sequence-based metagenomics identifies genes homologous to known database sequences but it is difficult, if not impossible, to reliably predict the function of truly novel genes through homology-based screens. Functional screening strategies are based on phenotypic detection of the desired activity, heterologous complementation of host strains, and, in some cases, induced gene expression (Taupp et al., 2011; Simon and Daniel, 2011).

Metagenomic libraries have been constructed using plasmids (<20 kb insert size), cosmids and fosmids (<40 kb insert size), and bacterial artificial chromosomes (>40 kb insert size). As genes involved in related metabolic pathways are often clustered in microbial genomes, metagenomic DNA of high molecular weight (HMW) is ideally cloned into cosmids or fosmids for functional screening. Most fosmid libraries have been constructed using the narrow-host-range CopyControl pCC1FOS vector (Epicentre Biotechnologies), and most functional screens have been performed in *Escherichia coli* (γ-*Proteobacteria*). Because gene expression is largely host-dependent (Gabor et al., 2004), broad-host-range systems increase opportunities for successful expression of genes of interest (Martinez et al., 2004; Li et al., 2005; Wang et al., 2006; Aakvik et al., 2009; Craig et al., 2010; Taupp et al., 2011). In addition, shuttle vectors have also been engineered for specific screening (Ly et al., 2011; Biver et al., 2013).

Gateway® LR cloning (Invitrogen) bypasses the conventional digestion and ligation strategies, and employs the λ site-specific recombinase-mediated transfer of a DNA fragment from an entry plasmid into a destination vector (Katzen, 2007). The entry vector (*attL1*–DNA–*attL2*) carries the DNA fragment of interest to be transferred, and a destination plasmid contains an *attR1*–*ccdB*–*attR2* segment. The product of CcdB protein is lethal to standard *E. coli* strains except those such as *E. coli* DB3.1 that harbor a DNA gyrase mutation. The λ integrase (Int), excisionase (Xis) and integration host factor (IHF) proteins mediate the recombination between the *attL1* × *attR1* and *attL2* × *attR2*, resulting in shuffle of the cloned DNA fragment from the entry vector into the destination plasmid. The recombinant destination transformants (in vitro recombination) or transconjugants (in vivo reaction), in which the *ccdB* gene is replaced with the cloned DNA flanked by *attB* sites, can be obtained by selection in a bacterial host such as *E. coli* DH5α. The growth of *E. coli* DH5α strains carrying the non-recombinant entry and destination vectors, and recombinant entry plasmid is inhibited due to the antibiotic selection and toxic effect of CcdB protein. In this work, we inserted the *attL* sites in the new cosmid vectors for future transfer of the cloned metagenomic DNA to Gateway® destination plasmids, in order to further expand the range of surrogate hosts for functional screening.

Despite these advances, improved screening vectors are required for additional flexibility with respect to replication and screening in multiple bacterial and eukaryal hosts. To address this limitation, we aimed to

* Corresponding author. Tel.: +1 519 888 4567x35606.
E-mail address: trevor.charles@uwaterloo.ca (T.C. Charles).

0167-7012/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.mimet.2014.01.015

construct new broad-host-range cosmids based on pRK7813 (Jones and Gutterson, 1987). Cosmid pRK7813, a derivative of broad-host-range cosmid RK2, contains the *cos* site for in vitro packaging of the recombinant cosmids, and the origin of transfer (*oriT*) for transfer of cosmids from *E. coli* to other Gram$^-$ bacterial hosts. Metagenomic libraries have been constructed using pRK7813 for functional screening of genes required for polyhydroxyalkanoate metabolism in *α-Proteobacteria Sinorhizobium meliloti* (Wang et al., 2006; Schallmey et al., 2011). We present two new broad-host-range cosmids pJC8 and pJC24 derived from pRK7813, developed to facilitate functional screening in multiple hosts. We also demonstrate the utility of pJC8 for hosting metagenomic DNA. For this, we isolated HMW DNA from agricultural soils that were cultivated with corn and wheat, using that extracted DNA for constructing metagenomic libraries. The other vector pJC24 has been used to construct a metagenomic library of human gut (Lam et al., in revision). These libraries are publicly available as part of the Canadian MetaMicroBiome Library project (http://cm2bl.org).

2. Materials and methods

2.1. Bacterial strains, plasmids and growth conditions

Multiple bacterial strains and plasmids were used in this study (Table 1). *E. coli* strains were grown at 37 °C in LB (1% tryptone, 0.5% yeast extract, 0.5% NaCl, pH 7) or TB (1.2% tryptone, 2.4% yeast extract, 72 mM K_2HPO_4, 17 mM KH_2PO_4, and 0.4% glycerol) medium. Enzymes were obtained from Thermo Scientific. Antibiotics were used at the following final concentrations: streptomycin, 100 μg/ml; rifampicin, 100 μg/ml; kanamycin, 50 μg/ml; tetracycline, 20 μg/ml; chloramphenicol, 20 μg/ml; gentamicin, 10 μg/ml.

2.2. Construction of broad-host-range cosmids

Multiple PCR primers were used in this work (Table 2). Steps involved in cosmid construction were illustrated in Fig. 1A. The Gateway *attL2* fragment was PCR amplified using primers JC5 and JC6, and ORFeome–SMc00604 (Schroeder et al., 2005) as a template, digested with *Hin*dIII/*Bam*HI, and then inserted into the same sites in pRK7813 (Fig. 1B; Jones and Gutterson, 1987) to obtain pJC5. The *attL1* site was PCR amplified using the same template and primers JC3 and JC4, restricted with *Eco*RI/*Bam*HI, then cloned into the same sites in pRK7813, yielding pJC6. Following verification of the *attL1* and *attL2* sites by DNA sequencing, the *attL2* fragment was obtained by *Hin*dIII/*Bam*HI digestion of pJC6 and then cloned into the same sites in pJC5 to obtain pJC7. A 1.1-kb DNA fragment carrying the *aacC4* gene (Gm^R) was PCR amplified from pTH1703 (Cowie et al., 2006) using primers JC17 and JC18, digested with *Bam*HI, and then inserted into the same site in pJC7 to obtain pJC8 (Fig. 1C). A 1.8-kb region encoding the *sacB* gene was PCR amplified from pK19mobsacB (Schäfer et al., 1994) using primers JC57 and JC58, and then cloned into the *Bam*HI site in pJC7, yielding cosmid pJC24 (Fig. 1D).

Table 2
DNA oligos used in this study. Restriction sites are underlined.

Oligo ID	Sequence (5′ to 3′)
JC3	CTTGAATTCAGATCTCAAATAATGATTTTATTTTG
JC4	TTTGTACAAGAAAGCTGGGTCACGTGGATCCAGCCTGCTTTTTTGTACAAA
JC5	TTTGTACAAAAAAGCAGGCTGGATCCACGTGACCCAGCTTTCTTGTACAAA
JC6	GCTAAGCTTCAAATAATGATTTTATTTTG
JC17	CCCGGATCCTCCGGCTCACGGTAACTGATG
JC18	ATAGGATCCACGTGCGTGTTGCCCCAGCAATCAG
JC57	GCGCGGATCCCAACTTTATGCCCATGCAACAG
JC58	GCGCGGATCCCACGTGTTGCGTTTTTATTTGTTAAC
JC165	GCCGCCGGCGTTGTGGATACCT
JC166	CCGGGCTTTCGTCTCTTGCCGCTTAT
JC174	CAGGGCAACGAGCCGATCGCTGATC
JC175	GGGTCAATGATGACCTGGTGCATTGCA

2.3. Preparation of cloning ready vector

Successful construction of a metagenomic cosmid library involves multiple steps (Fig. 2). In order to obtain sufficient amounts of cosmid DNA, *E. coli* DH5α carrying pJC8 or pJC24 was streaked on LB Tc plate and incubated overnight at 37 °C. Three colonies were inoculated in 50 ml of TB medium with tetracycline (15 μg/ml), and grown overnight at 37 °C with shaking. Cosmid DNA was isolated using the GeneJET Plasmid Miniprep Kit (Thermo Scientific) and DNA was quantified using a NanoDrop 1000 spectrophotometer (Thermo Scientific). DNA (50 ng) was loaded on a 0.8% agarose gel to verify the absence of RNA, which could interfere with the efficiency of subsequent dephosphorylation of linearized pJC8. In order to digest and dephosphorylate the vector simultaneously, 10 μg of pJC8 was incubated at 37 °C for 1 h with 10 units of FastDigest *Pml*I (*Eco*72I), 10 units of FastAP alkaline phosphatase and 1 × FastDigest Green buffer in a final volume of 500 μl. Five microliters of the reaction solution was loaded onto a 0.8% TAE

Table 1
Bacterial strains, plasmids and cosmids used in this study.

Bacteria and plasmids	Characteristics	References
Escherichia coli		
DH5α	*F*$^-$ Φ80*lacZ*ΔM15 Δ(*lacZYA–argF*) U169 *recA1 endA1 hdR17 phoA supE44 thi-1 gyrA96 relA1*	Lab collection
DH5α (Rif^R)	A spontaneous Rif^R mutant of DH5α, Rif^R	This work
HB101	*F*$^-$ *supE44 lacY1 ara-14 galK2 xyl-5 mtl-1 leuB6 recA13 rpsL20 thi-1 proA2 hsdSB20*, Sm^R	Lab collection
MT616	*pro82 thi-1 endA hsdR17 supE44 recA56* (pRK600), Cm^R	Finan et al. (1984)
DB3.1	*F*$^-$ *gyrA462 endA1 glnV44* Δ(*sr1–recA*) *mcrB mrr hsdS20* (rB$^-$, mB$^-$) *ara14 galK2 lacY1 proA2 rpsL20 xyl5 Δleu mtl1*, Sm^R	Life Technologies
Plasmids and cosmids		
pRK7813	IncP *oriT cos lacZα*, Tc^R	Jones and Gutterson (1987)
pK19mobsacB	Cloning vector, *sacB*, Km^R	Schäfer et al. (1994)
pCO40	pTH1705 carrying Gateway® *ccdB*-Cm^R cassette, Gm^R	Lab collection
pXINT129	Expressing λ *int* and *xis*, Km^R	Platt et al. (2000)
pRK2013	ColE1 replicon with RK2 transfer region, Km^R	Figurski and Helinski (1979)
pRK600	Helper plasmid, Cm^R	Finan et al. (1984)
pTH1703	pBR322 derivative, Gm^R	Cowie et al. (2006)
pJC5	pRK7813 carrying *attL2*, Tc^R	This work
pJC6	pRK7813 carrying *attL1*, Tc^R	This work
pJC7	pJC6 carrying *attL2* from pJC5, Tc^R	This work
pJC8	pJC7 carrying *aacC4*, Tc^R Gm^R	This work
pJC24	pJC7 carrying *sacB*, Tc^R	This work
ORFeomeSMc00604	pMK2010 with *smc00604*, Km^R	Schroeder et al. (2005)