Construction of an Expressible BAC Library of the Unculturable Insect Microorganism, Stink Bug *Plautia stali* Symbiont, for the Search of Biologically Active and Useful Symbiont Products

Hideaki Kobayashi,* a Rika Fujii-Muramatsu, b Hiroaki Noda, c and Keichi Takeishi b

*Laboratory of Genome Pharmaceuticals, School of Pharmacy, Nihon University; 7–7–1 Narashinodai, Funabashi, Chiba 274–8555, Japan; b Laboratory of Special Project, National Institute of Agrobiological Sciences; and c Division of Insect Sciences, National Institute of Agrobiological Sciences; 1–2 Owashi, Tsukuba, Ibaraki 305–8634, Japan.

Received September 2, 2013; accepted January 17, 2014

While gene products and metabolites of insect symbiotic bacteria may act as useful resources for insect-microbe studies and medicinal use, it is usually difficult to obtain the insect symbionts to some extent in quantity because most of them are unculturable. In this study, the possibility of using bacterial artificial chromosome (BAC) libraries as a heterologous gene expression tool for the discovery of novel symbiont metabolites was evaluated. A BAC library was constructed from the symbiont purified from the posterior midgut cecum of the stink bug *Plautia stali*. The BAC library, which consisted of 513 clones with an average insert size of 41 kb, represented greater than five-fold coverage of the genome. The ability of the BAC clones to express plural genes from large-sized insert DNA in *Escherichia coli* was examined by the growth of *E. coli* transformed *leu* operon-deficient DH10B cells on M9 minimal medium supplemented with glucose. Two BAC clones complemented leucine deficiency in DH10B cells; the clones contained the *leu* operon of the symbiont chromosome. The *P. stali* symbiont genes introduced into the BAC vector are functional in *E. coli*, and these genes are expressed in an operon unit. BAC libraries can be used to generate gene product- and metabolite-libraries, facilitating to characterize potential metabolites of the *P. stali* symbiont.

Key words insect; symbiont; heterologous; expression; bacterial artificial chromosome

Insects harbor various symbiotic microorganisms, i.e., symbionts.1–3) Symbionts have various influences on insects, and symbiotic lifestyles are highly diverse among insect species.3–5) Some symbionts provide nutrition to the host, while others manipulate sex and reproduction in the host. Many symbionts rely on their host for food and metabolic functions; therefore, cultivation of these microorganisms using conventional methods and media is often difficult. Some symbionts are so completely adapted to and dependent on host metabolism that they have lost many essential genes, which is evident from the analyses of insect symbiont genomes.6,7) Burke and Moran8) reported that the small genome size observed in many bacterial symbions of insects is mainly the result of genomic decay. Nevertheless, symbionts are generally thought to retain biologically important and prerequisite functional genes for symbiosis.

Culture extracts of microorganisms have been used for screening of useful natural products or in drug lead discovery.9) However, the inability to cultivate symbionts in vitro is a major obstacle to identifying and studying their gene products and metabolites. Genomic analysis of symbionts using nucleotide sequencing technology enables the prediction of putative gene products and their functions. The genomes of many insect symbionts have been sequenced since Shigenobu et al.10) sequenced the entire genome of *Buchnera*, a symbiont of the pea aphid *Acyrthosiphon pisum*. Over 25 symbiont genome sequences are currently available, and the number continues to increase with the advancement of sequencing technologies. Genome sequence data are very useful, but their applications are limited in the study of symbiont metabolites since their metabolites are not directly analyzed. One of the alternative approaches to overcome this limitation involves expressing a series of symbiont genes in an appropriate host organism and obtaining the metabolites generated by the resultant gene products.

*Escherichia coli* is widely used for gene expression analysis, as it can be easily transformed to express gene(s) of interest using various DNA vectors. Bacterial artificial chromosome (BAC) vectors stably maintain large DNA inserts (larger than 100 kb) in *E. coli*,10) facilitating the cloning of large DNA fragments that may include a specific set of genes required for a particular metabolic process. BAC genomic DNA libraries are commonly used for sequencing and more recently for heterologous (or surrogate) expression of prokaryotic genes.11,12) Furthermore, soil metagenomic DNA BAC libraries have been constructed, and a remarkable number of diverse biocatalysts have been isolated.13,14) BAC libraries constructed from the genomes of insect symbionts are expected to be effective tools for the survey of symbiont metabolites. In order to construct the libraries of gene products and metabolites derived from the unculturable insect symbiont DNA and to develop the system to search biologically active novel substances by utilizing the libraries, it is very important to obtain basic data necessary for the construction of a BAC library from unculturable insect symbiont DNA and for the heterologous expression of insect symbiont genes in BAC clones in *E. coli*.

The brown-winged green bug, *Plautia stali* (Hemiptera: Pentatomidae), a pest of fruit trees, carries obligate symbiotic bacteria in the posterior midgut cecum.15) These symbionts consist of a single species of gamma-proteobacterium, which is closely related to *Erwinia* and *Pantoea* species16); these bacteria are speculated to supply the host insects with vitamins.15) We have determined the draft sequence of the symbiont genome and its size has been revealed to be 3.8 Mb.17) Herein, we describe the construction and expression of a BAC library of the *P. stali* symbiont DNA. This study aims at (i) establishing...
ing a general procedure for the construction of a BAC library from insect symbiont DNA; (ii) confirming heterologous expression of symbiont genes in BAC clones in *E. coli*; and (iii) evaluating the possible use of BAC libraries of insect symbionts as a tool for the discovery of novel symbiont metabolites.

**MATERIALS AND METHODS**

**Insects and Sample Collection** *P. stali* was collected in Tsukuba and maintained on dry soybeans, raw peanuts, and water supplemented with sodium L-ascorbate (0.05%) and L-cysteine (0.025%).18) *P. stali* symbionts were purified from adult insects as described previously.17) Briefly, the adult insects were anesthetized on ice and their abdomens were dissected under a binocular. The posterior midgut ceca, which are easily identified by their yellow appearance, were collected in normal saline (0.9% NaCl) and crushed on a Cell Strainer with a 40-µm nylon mesh (BD Falcon, Bedford, MA, U.S.A.). The filtrate, which contained the symbionts, was passed through the Cell Strainer twice and then centrifuged at 8000×*g* for 5 min at 4°C. The pelletted symbionts were either used immediately or stored at −80°C for analysis.

**BAC Library Construction** To obtain chromosome-sized genomic DNA with minimal shearing, the symbionts from 40 adult insects were embedded in 1% agarose, which was either CleanCut Agarose from the kit or Agarose GB (Nippon Gene, Toyko, Japan). The agarose/symbiont suspensions were solidified in TAE buffer using Mupid-2 Plus (Advance, Tokyo, Japan) at 20 mmol/L; ethylenediaminetetraacetic acid (EDTA)-2Na, 1 mmol/L; MgCl2, 10 mmol/L) per plug for 1 h at 37°C, followed by >24 U/ml protease K overnight at 50°C. Residual protease K activity was inactivated with 1 mmol/L benzylsulfonyl fluoride (PMSF; Wako Pure Chemical Industries, Ltd., Osaka, Japan). The agarose plugs were incubated with 4 U Sau3 AI (Roche Applied Science, Penzberg, Germany) in 0.4 mL of Sau3 AI buffer [NaCl, 50 mmol/L; Tris–HCl, 10 mmol/L (pH 7.5); di-thiothreitol (DTT), 1 mmol/L; MgCl2, 10 mmol/L] per plug for 30 min at 37°C to partially digest the genomic DNA. The partially digested DNA was separated by pulsed-field gel electrophoresis (PFGE) in 1% pulsed field certified agarose (Bio-Rad Laboratories) at 200 V for 3 h and then dialyzed overnight against TE buffer [10 mmol/L EDTA (pH 8) and 0.2 mg/mL proteinase K for 1 h at 37°C, and proteinase K was inactivated with 2 mmol/L PMSF. The ligation solution was dropped onto a VSWP membrane (Millipore, Billerica, MA, U.S.A.) and sequentially dialysed against distilled deionized water (DDW) for 2 h at 4°C, followed by 30% (w/v) polyethylene glycol 8000 (Fluka, Steinheim, Germany) in TE buffer for 5 h at 4°C. The ligated product recovered from the VSWP membrane was transformed into DH10B electrocompetent cells (NEB) using Gene Pulser II (Bio-Rad Laboratories) at 200 V, 25 µF, 2.0 kV using a 1 mm cuvette. Transformed DH10B cells were grown overnight on LB [Bacto tryptone 10 g/L (BD, Sparks, MD, U.S.A.); Bacto yeast extract (BD) 5 g/L; NaCl 10 g/L; pH 7] or 2YT (Bacto tryptone 16 g/L; Bacto yeast extract 10 g/L; NaCl 5 g/L; pH 7) agar plates [Bacto agar (BD) 1.5% (w/v)] containing 30 µg/mL chloramphenicol (CM) coated with isopropyl-β-D-thiogalactopyranoside (IPTG) and 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal) for blue/white colony selection. White colonies were cultured in 2YT medium containing 30 µg/mL CM. Aliquots of the cultured medium were mixed with an equal volume of 50% glycerol and stored at −80°C.

BAC plasmids were purified using the High Purity Plasmid Purification System (Marligen Biosciences, Ijamsville, MD, U.S.A.), according to the manufacturer’s instructions. Sizes of DNA inserts were estimated by PFGE after NotI digestion.

**Leucine Complementation Test** Bacteria containing BACs with genomic DNA inserts were cultured for 1 to 2 d at 37°C on M9/CM agar plates consisting of M9 minimal medium (NaHPO4, 6 g/L; KH2PO4, 3 g/L; NaCl, 0.5 g/L; NH4Cl, 1 g/L; CaCl2, 0.1 mmol/L; MgSO4, 1 mmol/L) supplemented with 0.4% (w/v) glucose, 30 µg/mL CM, and 1.5% (w/v) Bacto agar (BD). The M9 minimal medium, well used for the cultivation of *E. coli*, contained only salts including a nitrogen source, and in this experiment, a carbon source but not amino acids was supplemented. In the case of a positive control for the growth of BAC clones, 40 µg/mL leucine was included in the M9/CM agar plates (Fig. 3).

**DNA Sequence Analysis** Nucleotide sequences of both ends of the DNA inserted into the BAC clones were determined using T7 (5′-TAA TAC GAC TCA CTG AGG-3′) and SP6 (5′-ATT TAG GTG ACA CTA TAG-3′) primers, ABI Prism BigDye Terminator v3.1 Cycle Sequencing Kits (Applied Biosystems), Foster City, CA, U.S.A.), and an ABI 3700 genetic analyzer (Applied Biosystems).

Three shotgun libraries were constructed from the BAC081209-62 clone in order to determine its sequence by either single digestion by *Sau3* AI (or *MspI*) or with dual digestion by *SphI* and *HindIII*. The DNA fragments were introduced into the pUC18 vector (Toyobo, Osaka, Japan), and the products were transformed into *E. coli* DH5α cells (Toyobo). Insert sequences of the plasmids purified from these shotgun libraries were determined from both sides using M13-M3 (5′-GTA AAA CAC GAG CCC AG-3′) and M13-RV (5′-CAG GAA ACA GCT ATG AC-3′) primers. The shotgun sequences were assembled and corresponding contigs of the genome sequence [DNA Data Bank of Japan (DDBJ) accession numbers: BABY01000001–BABY01000144]17) were assigned using the SeqMan program in the Lasergene software package (DNASTar, Madison, WI, U.S.A.). Sequencing gaps between the assembled sequences were filled either by sequencing the PCR products of the corresponding region amplified by ap-
propriate primers or by sequencing the corresponding clones from a P. stali symbiont DNA shotgun library, which was constructed from mechanically sheared symbiont DNA and the pUC118 vector (TaKaRa Bio Inc., Shiga, Japan) for whole genome sequence assembly of the symbiont (unpublished data). The nucleotide sequences of leu operon genes in E. coli K12 (NCBI accession number: NC_000913) and Salmonella enterica serovar. Typhimurium LT2 (GenBank accession number: AE006468) were obtained from the National Center for Biotechnology Information (NCBI) nucleotide database. The P. stali symbiont leu operon from the draft genome was analyzed using the Microbial Genome Annotation Pipeline [MiGAP, Database Center for Life Science (DBCLS), Japan][22] and the Basic Local Alignment Search Tool (BLAST; stand-alone) search (NCBI).[23] The translated amino acid sequences of the leu operon genes of the symbiont were compared with those of E. coli and S. enterica by ClustalW using the MegAlign program in the Lasergene software package (DNASTAR), and homology percentages were obtained.

Reverse Transcription Polymerase Chain Reaction (RT-PCR) Analysis  Expression of leu operon genes in BAC clones was analyzed by RT-PCR. After BAC clones were cultured in M9/CM medium consisting of M9 minimal medium supplemented with 0.4% (w/v) glucose and 30 µg/mL CM at 37°C overnight, they were collected using RNAprotect bacteria reagent (QIAGEN, Hilden, Germany). Total RNA was isolated using RNasy mini kit (QIAGEN) and RNase-Free DNase set (QIAGEN), according to the manufacturer’s instructions. To obtain cDNA, reverse transcription was performed using ReverTra Ace qPCR RT master mix with gDNA remover (Toyobo), according to the manufacturer’s instructions. LeuA cDNA was amplified by PCR using Ex taq DNA polymerase (TaKaRa Bio Inc.) and oligo DNA primers (LeuA–F: 5’-CAC CGT ATG GAA GAT ATG GTT TAC-3’ and LeuA–R: 5’-CAG ATG TTG TTC AGG GCA TTG ACC-3’), and the products were separated by agarose gel electrophoresis.

RESULTS

Construction of a P. stali Symbiont BAC Library A BAC library, consisting of 775 clones, was constructed from genomic DNA of the P. stali symbiont purified from adult midgut ceca (Fig. 1). The size of the majority of the inserts was 40–50 kb (Fig. 2), with the largest size observed to be 120 kb (data not shown). Approximately two-thirds of the total number of clones possessed DNA inserts sized larger than 10 kb.

To examine the insertion rate of symbiont DNA in the BAC library, end sequences of the insert were analyzed in 513 clones that possessed >10-kb insert DNA. The end sequences of most BAC plasmids could be determined using at least one of the primers set, either T7 or SP6 primer. The sequences from 485 clones matched the draft genome sequence of the symbiont (DDBJ accession numbers: BABY01000001–BABY01000144), indicating that at least 95% of the BAC clones contained genomic DNA from the P. stali symbiont. The average insert length of these 513 clones was 41 kb, which was sufficient to contain approximately 30 genes. Some clones would be expected to possess the entire leu operon. Some clones would be expected to possess the entire leu operon. Of the 513 BAC clones, clones BAC070802-41-2 and BAC081209-62 supported the growth of DH10B cells on M9/CM agar plates (Fig. 3, b-2, b-3), suggesting that the corresponding clones complemented leucine deficiency in DH10B cells. To confirm that this complementation was due to the DNA that was inserted into the BAC plasmid, the plasmid was extracted from the two bacterial colonies (BAC070802-41-2 and BAC081209-62) and retransformed into fresh DH10B cells. Twenty new colonies grown on LB/CM agar plates were selected from each previous colony and tested for growth on
M9/CM agar plates. All of the selected colonies grew on M9/CM agar plates, indicating that leucine deficiency was complemented by the BAC plasmids.

Characterization of DNA Inserts from Two Clones That Were Viable on M9/CM Agar Plates

The DNA end sequences of the BAC070802-41-2 and BAC081209-62 clones were assigned to relevant contigs in the draft genome sequences (DDBJ accession numbers: BABY01000001–BABY01000144) using BLAST. BAC070802-41-2 was found to bridge two contigs, 33076 and 32699, which did not overlap, whereas a single end sequence of BAC081209-62 was found in contig 33076. The nucleotide sequence of BAC081209-62, whose insert size (47 kb) was larger than that of BAC070802-41-2 (39 kb), was determined in order to examine the relationship between the clones and contigs and to determine whether the leu operon was included in this sequence. The results of an alignment of the contigs with BAC081209-62 and BAC070802-41-2 are shown in Fig. 4. The 39-kb sequence (BAC070802-41-2) was included in BAC081209-62 and possessed the entire sequence of contigs 00075, 32655, 00011, and 34050 as well as a partial sequence of contigs 32699 and 33076. Contig 00011 was confirmed to have the leu operon by BLAST; the putative leu operon protein sequences were similar to those of E. coli and S. enterica (55.6% and 51.9% in LeuL and 77.6–86.0% and 80.1–85.8% in LeuA–D, respectively; Table 1).

To confirm that the putative leu operon genes were expressed in these BAC clones, RT-PCR was performed. BAC clones were cultured in M9/CM medium and total RNA was isolated. Reverse transcription was performed with or without reverse transcriptase and a portion of leuA gene (485 kb) was amplified by PCR. As shown in Fig. 5, reverse transcription-dependent DNA band derived from leuA mRNA was observed in these BAC clones. This DNA band was not observed in the BAC clone carrying no insert cultured in M9/CM+leucine medium (data not shown). These results indicated that these BAC clones viable on M9/CM agar plates expressed leu operon genes.
DISCUSSION

We constructed a BAC library of *P. stali* symbiont DNA as a first step toward the generation of gene product- and metabolite-libraries to explore useful substances from insect symbionts. The average size of the DNA inserts of the constructed library was comparable to that of DNA inserts of the previously constructed BAC libraries from cultured bacteria or a soil metagenome, e.g., 98 kb from *Bacillus cereus* or 44.5 kb from a soil metagenome library.13) The ability of the BAC clones to express plural genes from large-sized insert DNA in *E. coli* was confirmed by the growth of BAC-transformed *leu* operon-deficient DH10B cells on M9/CM agar plates; the clones contained the *leu* operon of the symbiont chromosome.

Isolation of symbionts from insects is critical for BAC library construction. The *P. stali* gut symbiont used herein had two important advantages that aided in the construction of the library. First, many adult *P. stali* can easily be obtained by breeding in the laboratory using simple food. Second, *P. stali* symbionts can be isolated from the host tissue with less contamination than more commonly studied insect symbionts due to their aggregation in the posterior midgut ceca. A similar method can be applied to insect symbionts residing in specific host tissues. An aphid symbiont *Buchnera*, whose whole genome sequence was described for the first time using insect symbionts, was collected by a similar procedure.9) However, this procedure is not always applicable for many insect symbionts, e.g., Wolbachia, Rickettsia, and Cardinium, which are scattered throughout the insect body, and some secondary symbionts that exist only in small numbers. Although some insect symbionts such as Wolbachia can be grown in insect cell lines for sample DNA preparation, most symbionts are difficult to culture in *in vitro* cell lines. Amplification of small quantities of DNA may be an alternative method for obtaining sufficient DNA for BAC library construction, as shown in a previous study regarding genome sequencing in the protist symbiont of termites.26) However, the quality and length of DNA typically obtained using this method is generally insufficient for BAC library construction.

Recent studies have shown that symbiotic bacteria of invertebrates can produce many interesting natural products.27) For example, putative genes responsible for the production of pedierin-type antitumor polyketides were found from the symbiont of the beetle *Paederus fuscipes* by diagnostic PCR analysis of a metagenomic cosmid library constructed from the total DNA of the beetle.28) However, PCR-based gene identification can only be applied to known characterized genes. In contrast, symbiont genes cloned in a vector of an *E. coli* system could be expressed in *E. coli* cells; the resultant libraries constructed using the vectors would be useful for exploring novel natural products from insect symbionts. We used a BAC vector, as the BAC vector is able to clone large-sized DNA that can include operon units. A complementation assay was used to demonstrate heterologous expression of symbiont genes in a host *E. coli* DH10B, which lacks the *leu* operon. The results of the current study demonstrated that two BAC clones grown on minimal medium possessed a shared genome fragment of the *P. stali* symbiont, wherein the *leu* operon was located. This heterologous expression revealed that promoters and genes of the *P. stali* symbiont were active in the host *E. coli*. Using this BAC library as a model, methods for improving library quality can be studied. For example, if biased and/or insufficient expression of some genes were observed, the use of modified version of BAC vectors and another host strain would be considered. If some genes are not incorporated in the BAC library by expressing toxic substances to host *E. coli*, host strains that can express these genes would be surveyed.

The present results suggest that the gene product- and metabolite-libraries generated from BAC libraries constructed from genomic DNA of *P. stali* symbiont are useful resources for finding products of the symbiont. If appropriate assay methods become available, the medium in which BAC-transformed bacterial cells are cultured will be utilized as a screening source for identifying target metabolites or novel substances such as lead compounds for the medicinal industry. In future studies, it will be important to examine and evaluate heterologous expression of the symbionts that are distantly related with *E. coli*, particularly when using *E. coli* for discovering natural products from the symbionts with BAC libraries.

Acknowledgments We are grateful to Dr. Shunji Natori for helpful advice. We thank Noriko Higashida for her technical assistance. This work was supported in part by the Integrated Research Project for Insects Using Genome Technology of the Ministry of Agriculture, Forestry, and Fisheries of Japan (2007–2009); a Nihon University Individual Research Grant for 08–133 (2008); the “Academic Frontier” Project for Private Universities: matching fund subsidy from MEXT (Ministry of Education, Culture, Sports, Science and Technology) of Japan 2007–2009; Special Subsidies in the Ordinary Expense Subsidies for Private Universities from MEXT of Japan; and a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (JSPS).

REFERENCES


2) Oliver KM, Degnan PH, Burke GR, Moran NA. Facultative symbionts in aphids and the horizontal transfer of ecologically important traits. *Annu. Rev. Entomol.*, **55**, 247–266 (2010).


