An alternative genome-integrated method for undomesticated *Bacillus subtilis* and related species

(Received April 2, 2018; Accepted June 8, 2018; J-STAGE Advance publication date: November 27, 2018)

Gumpanat Mahipant,1,2 Junichi Kato,3 Naoya Kataoka,4,5 and Alisa S. Vangnai2,6,*

1 Biological Sciences Program, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand
2 Biocatalyst and Environmental Biotechnology Research Unit, Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand
3 Department of Molecular Biotechnology, Graduate School of Advanced Sciences of Matter, Hiroshima University, Hiroshima 739-8530, Japan
4 Division of Agricultural Sciences, Graduate School of Sciences and Technology for Innovation, Yamaguchi University, Yamaguchi 753-8515, Japan
5 Research Center for Thermotolerant Microbial Resources, Yamaguchi University, Yamaguchi 753-8515, Japan
6 Center of Excellence on Hazardous Substance Management (HSM), Chulalongkorn University, Bangkok 10330, Thailand

Full Paper

Given their applicability in genetic engineering, undomesticated *Bacillus* strains are extensively used as non-natural hosts for chemical production due to their high tolerance of toxic substrates or products. However, they are difficult to genomically modify due to their low transformation efficiencies. In this study, the *Bacillus*-*E. coli* shuttle vector pHY300PLK, which is widely used in gram-positive bacteria, was adopted for genome integration in organic solvent-tolerant *Bacillus* isolates. The *Bacillus*-replicative vector was used to deliver homologous recombinant DNA and propagate itself inside the host cell, increasing the likelihood of genome integration of the recombinant DNA. Then, the unintegrated vectors were cured by cell cultivation in antibiotic-free medium with facilitation of nickel ions. The developed protocol was successfully demonstrated and validated by the disruption of *amyE* gene in *B. subtilis* 168. With an improved clonal selection protocol, the probability of clonal selection of the *amyE::cat* genome-integrated mutants was increased up to 42.0 ± 10.2%. Genome integration in undomesticated, organic solvent tolerant *Bacillus* strains was also successfully demonstrated with *amyE* as well as *proB* gene creating the gene-disrupted mutants with the corresponding phenotype and genotype. Not only was this technique effectively applied to several strains of undomesticated *B. subtilis*, but it was also successfully applied to *B. cereus*. This study validates the possibility of the application of *Bacillus*-replicative vector as well as the developed protocol in a variety of genome modification of undomesticated *Bacillus* species.

Key Words: *Bacillus cereus, Bacillus subtilis*, genome modification, organic solvent-tolerant bacteria, undomesticated bacteria

Abbreviations: *B. subtilis*, Bacillus subtilis; *B. cereus*, Bacillus cereus; LB, Luria-Bertani; SMM, Spizizen minimal medium; CRISPR, clustered regularly interspaced short palindromic repeat; TcR, Tetracycline resistant; Tc S, Tetracycline sensitive; Cm R, Chloramphenicol resistant; Cm S, Chloramphenicol sensitive; Ts, Temperature-sensitive

Introduction

Domesticated *Bacillus* strains well-adapted to laboratory conditions are widely used for many biotechnological applications, especially natural product production, due to their ease of handling and “generally regarded as safe (GRAS)” property (Schallmey et al., 2004). However, since the demands for the microbial production of commodity, specialty, and non-natural chemicals are growing,
Bacillus strains with desired metabolic activities or unique properties are required, and thus, new strains are isolated, and investigated for properties such as tolerance of harmful concentrations of chemical products, or the ability to thrive in unconventional fermentation conditions, such as an aqueous-organic-solvent two-phase system (Schmid et al., 2001). Then, these natural-isolated or undomesticated microbes are generally subjected to genetic or metabolic engineering for strain improvement and to enhance their production efficiency, generating efficient microbial cell factories (Prather and Martin, 2008). In many cases, genome modification is necessary and has more advantages over extrachromosomal plasmid handling because of the higher stability of introducing gene(s) and the manipulability of large or multiple gene fragments (Tas et al., 2015). To date, various types of bacterial-genome editing tools have been developed, including homologous recombination (Sun et al., 2015; Wang, Y. et al., 2012), and nuclelease-assisted genome excision or recombineering, such as the CRISPR-Cas9 based system (Westbrook et al., 2016). Although the CRISPR-Cas9 system is considered a powerful genome engineering technique, its efficiency depends on several factors, including Cas9 activity, delivery methods, off-target effects, target site selection, and the known genome sequence required for the effective short-guide RNA design (Peng et al., 2016); therefore, this technique may not be fully applicable to undomesticated strains with unknown genetic backgrounds.

Homologous recombination methods are universal tools for bacterial genome modification. However, the recombination efficiency in undomesticated Bacillus strains is limited not only by low transformation efficiency (Romero et al., 2006; Wang, H. et al., 2016), but also by the substrate DNA availability (Wang, Y. et al., 2012). Unlike the natural competence of some reference Bacillus strains in which foreign DNA can be naturally taken up from the surrounding environment by cells, the transformation of undomesticated strains requires exhaustive optimization, and has been mostly successful using electroporation (Dong and Zhang, 2014; Liu et al., 2014). In terms of DNA substrates, linear double-stranded DNA molecules introduced into host cells are generally prone to degradation. Consequently, temperature-sensitive (Ts) shuttle vectors have been developed and are used to deliver recombinant DNA into thermophilic or thermostolerant Bacillus cereus, Bacillus amyloboulefaciens, and Bacillus thuringiensis (Liang et al., 2007; Wang, Q. et al., 2012; Zakataeva et al., 2010).

As an alternative, this work has developed a simple genome modification method that can be applied to a wide range of undomesticated Bacillus strains. The technique relies on a Bacillus-replicative vector that can be cured under non-selective conditions while delivering the homologous recombinant DNA for genome integration. This developed vector was constructed from a commercially available vector, pHY300PLK, which is replicable in Bacillus and other gram-positive bacteria (Ishiwa and Shibahara, 1985); therefore, the amount of DNA substrate is sufficient for homologous recombination. In addition, while the selection process of other temperature-sensitive methods requires an increase in temperature, the selection process in this study relies on the segregational instability of the plasmid in antibiotic-free conditions. The developed genome integration protocol was validated with undomesticated organic-solvent-tolerant B. subtilis and B. cereus.

Materials and Methods

Bacterial strains and plasmids. The bacterial strains and plasmids used in this study are listed in Table S1. Bacillus subtilis strain 168 was used as a wild-type model for B. subtilis. The organic-solvent-tolerant B. subtilis strains were obtained from previous studies (Kataoka et al., 2011; Navacharoen and Vangnai, 2011). The bacterial strains were routinely cultured at 37°C in Luria-Bertani (LB) medium. The antibiotics used in the selective medium were 20 mg/L tetracycline (Tc), 5 mg/L chloramphenicol (Cm) and 100 mg/L spectinomycin (Sp).

Vector construction. To conduct genome integration in this study, Bacillus-replicative vectors were constructed. The vectors included two important parts, i.e. the Bacillus-E. coli shuttle vector part and the homologous recombinant part. For the Bacillus-E. coli shuttle vector part, the pHYA vector was used. As shown in the construction scheme, the pHYA vector was constructed using pHY300PLK (TAKARA, Japan) as a PCR template (Fig. S1). The primer pair of F-phya and R-phya with additional restriction sites was used for PCR amplification (Table S2). Then, the PCR product was digested by SpII and self-ligated, resulting in a 4.9-kb pHYA vector (Fig. 1A).

The homologous recombinant part consisted of a genome-integration target region flanked by an antibiotic resistance gene. The following vectors were constructed with different homologous recombinant parts. The homologous recombinant part was constructed in an E. coli vector (i.e. pBlueScriptSK(II)-, a non-replicative vector in Bacillus) (Table S1) generating pBS-AMC (Fig. S2), then subcloned into the pHYA vector. The amylE gene flanked by the chloramphenicol resistance gene (cat) in pBS-AMC was used to construct pHYAMC (Fig. 1). In brief, the pair of F-amylE-HindIII and R-amylE-EagI primers was used to amplify the wild-type amylE using the genomic DNA of B. subtilis 168 as a template. The 2.0-kb PCR product was further cloned to pBlueScriptSK(II)- with KpnI-EagI sites, resulting in pBS-amyE. The cat gene encoding chloramphenicol acetyltransferase from pH101 (MoBiTec, Germany) was amplified with F-cat-MCS and R-cat-SacI primers and further cloned into pBS-amyE with PstI-SalI sites, resulting in pBS-AMC. The cat-flanked-amylE fragment was cloned into pHYA with HindIII-EagI sites, resulting in the 7.5-kb pHYAMC vector (Fig. 1B).

The proB-genome-integration cassette obtained from a previous study (Mahipant et al., 2017) was also used for proB disruption by the Bacillus-replicative vector. Briefly, the proB gene encoding glutamate 5-kinase was amplified using F-53proB and R-53proB-SphI primers, while the sgc gene was amplified using F-Spg-SacII and R-Sgc-SacII primers. The proB PCR product was cloned into pUC19 at the SphI-EcoRI sites, and then the proB gene was flanked in the middle by the sgc gene at the SacII site, resulting in pUC-KOB. The sgc-flanked-proB frag.
ment was cloned into pHYA at the SphI-EcoRI sites, resulting in pHYKOB. To disrupt proB gene in B. cereus, the recombinant plasmid pHYKOBc was constructed. The UP-proBc (829 bp; SphI-XhoI) and DO-proBc (892 bp; XhoI-SalI) were amplified from B. cereus ATCC14579, cloned into pHYA according to their additional restriction enzyme sites, then the spc gene was flanked at XhoI site resulting pHYKOBc.

The standard genetic manipulation carried out as described by Mahipant et al. (2017). The KOD plus NEO (Toyobo, Japan) was used for the polymerase chain reaction (PCR) experiment. All plasmids were constructed and maintained in E. coli strain DH5α. The primers used in this study were listed in Table S2.

Electroporation. The plasmid was transformed into the Bacillus competent cell host by electroporation as described by Kataoka et al. (2011), with some modifications. Briefly, an inoculum (2% v/v) of an overnight, LB-grown culture was transferred into 50 mL of fresh LB medium and incubated with shaking at 37°C for 3 h (OD600 = 0.6–0.8). The culture was then chilled and washed 4 times with 50 mL HSMG buffer (1 mM HEPES, 0.5 M sorbitol, 1 mM MgCl2, 10% glycerol) and resuspended in 1 mL of HSMG buffer. The competent cells (200 μL) were mixed with the plasmid DNA (5–20 μL) and incubated for at least 10 min on ice. The chilled mixture was pulsed at 2.5 kV, 25 μF, 200 Ω in a 2-mm-gap electroporimeter (BTX ECM630, Harvard Apparatus, USA). The pulsed cell was immediately chilled, recovered with 2 mL of Super Optimal broth with Catabolite repression (SOC) medium, incubated with shaking at 37°C for 3 h, and spread on LB agar supplemented with tetracycline.

Genome integration protocol. The genome integration protocol consisted of 3 main steps: (i) plasmid delivery, (ii) homologous recombination and plasmid segregation, and (iii) clonal selection, as illustrated in Fig. 2, in which amyE-cat genome integration using pHYAMC was shown as an example. First, the recombinant plasmid was transformed into the Bacillus cell host and the clones were selected as described in the electroporation section. The recombinant plasmid was allowed to homologously recombine with the target gene and segregate the unintegrated vector when cells were cultured under non-selective conditions in an antibiotic-free medium. The plasmid-harboring clone (TcR, CmR: tetracycline resistant, chloramphenicol resistant) was subcultured from Tc-containing LB into 5 mL of fresh antibiotic-free LB broth and incubated with shaking at 37°C for 1–9 cultivation cycles. Each cultivation cycle took approximately 10–14 h, and then the culture (2%, v/v) was further transferred to a new cultivation cycle. To estimate the plasmid segregation efficiency, the cell suspension was serially diluted and spread on LB agar both with and without tetracycline. The plasmid segregation percentage was calculated by: 100 – (number of cells on Tc plate/number of cell on non-Tc plate ¥ 100). A high plasmid-segregation percentage represented a high number of cured cells (i.e. cells without plasmid).

The clonal selection step was then performed to select the genome-integrated clones. During several rounds of cultivation in an antibiotic-free medium, the cell suspension still consisted of the mixed populations of wild-type cells (TcS-CmS), cells harboring an unintegrated plasmid (TcR-CmR) and genome-integrated cells (TcS-CmR). To select the genome-integrated clones (Fig. 2), the cell suspension with mixed populations was subcultured (2%, v/v) into 5 mL of Cm-containing LB medium, incubated with shaking at 37°C for 6 h, serial diluted, spread on Cm-containing LB agar, and incubated overnight. Each CmR clone was picked and resuspended in 200 μL of Cm-containing...

Fig. 1. The plasmid maps of the Bacillus-E. coli shuttle vector pHYA and the pHYAMC. Both pHYA (A) and pHYAMC (B) plasmids contain the origin of replication (ori-pA1) that drives the self-replication in Bacillus, a tetracycline resistance gene (TcR), an ampicillin resistance gene (ApR), and the ori-177, which is replicable in E. coli (not shown). The pHYAMC contains the cat-flanked-amyE fragment for selection and genome-integration purposes.
Genome integration method for *Bacillus* 99

LB medium. Half of the cell suspension (100 μL) was kept in a 96-well plate as a master plate, while the other half (i.e. the replica 96-well plate) was challenged in a Tc-containing LB medium with shaking incubation overnight. The positive genome-integrated clones were selected based upon the TcS-CmR phenotype. As the demonstrative case, disruption of *proB* with *spc* gene using pHYKOB was performed, and the clonal selection conditions were adopted from genome-integration by pHY AMC in the presence of proper metal ion (NiCl₂) as described below, with slight modification where chloramphenicol was replaced by spectinomycin.

In addition, the effect of metal ion, NiCl₂ (Ni) or CdSO₄ (Cd) and chlorotetracycline (CT) on clonal selection efficiency was investigated. Initially, the effect of the substances at various concentrations on growth *B. subtilis* WT and *B. cereus* WT were determined by cultivating cells for 12 h in 5 mL LB medium with and without NiCl₂ or CdSO₄ at 0.5–10 mM, or chlorotetracycline at 10–30 μM. The cell density (OD₆₀₀) was measured as OD<sub>control</sub> and OD<sub>treated</sub>, respectively. The percentage of growth inhibition was then calculated by 100 × (OD<sub>control</sub>–OD<sub>treated</sub>/OD<sub>control</sub>).

The proper concentration(s) of metal ion and chlorotetracycline was then incorporated into the clonal selection step to reduce the numbers of wild-type cells, and unintegrated-plasmid-harboring cells (and, hence, increasing the chance of obtaining genome-integrated clones). Cells from each cultivation cycle were treated with NiCl₂ at a selected concentration with and without chlorotetracycline for 6 h prior to the selection of Cm R clones by plating on Cm-containing LB agar followed by the Tc challenge as described above. The clonal selection efficiency was calculated by (100 × number of TcS-CmR clone)/(number of Cm R clone). The higher percentage of TcS-CmR clones indicated a higher clonal selection efficiency.

**Amylase activity plate assay.** An amylase activity assay on a starch-containing plate was used to determine the phenotype of *amyE* gene product in the *B. subtilis* WT and mutant cells. Ten microliters of cell suspension were

---

*Fig. 2.* The protocol for genome integration in *B. subtilis* using pHYAMC. The protocol includes 3 main steps: (i) plasmid delivery by electroporation, (ii) homologous recombination and plasmid segregation in antibiotic-free medium for several cultivation rounds, and (iii) clonal selection in LB medium supplemented with chloramphenicol (Cm) to prevent growth of wild-type cells (WT; Tc<sup>S</sup>-Cm<sup>R</sup>), and nickel ion (Ni) to reduce the number of cells harboring an unintegrated plasmid (Tc<sup>R</sup>-Cm<sup>R</sup>). Then, after selection in tetracycline (Tc)-containing LB medium, the positive genome-integrated clones (*Bacillus*-AMC) are selected based upon the Tc<sup>S</sup>-Cm<sup>R</sup> phenotype.
Proline auxotroph assay. To test the proline auxotroph activity, a clear halo zone on the plate indicated amylase was supplemented. A 1% (v/v) iodine solution was overlaid. The phenotype of the α-amylase was used as a gene model for genome integration. The common distal phenotype was constructed as a homologous recombinant cassette. The target site of the genome was constructed to facilitate genome modification. In this study, the proB genome integration and selection protocol was developed based on a Bacillus-replicative vector, in order to facilitate genome modification in undomesticated Bacillus strains. The plasmid contains two parts including (i) a Bacillus-replicative vector backbone, and (ii) homologous regions of the target gene flanked with an antibiotic resistance gene. The Bacillus-replicative vector was used to deliver the modified target gene, and was able to multiply in Bacillus cells thereby increasing the chance for homologous gene recombination to occur. The pHYA was constructed based on the Bacillus-E. coli shuttle vector pHY300PLK (Fig. 1A). This vector was capable of genetic engineering in E. coli and self-propagation in Bacillus. The pHYA-harboring Bacillus clones were selected by the tetracycline-resistant (TcR) phenotype. The target site of the genome was constructed as a homologous recombinant cassette. The common distal genome integration site in the B. subtilis amyE (encoded alpha-amylase) was used as a gene model for genome integration. The phenotype of the amyE mutant strain was readily detected by starch degradation (Guerout-Fleury et al., 1996; Trauth and Bischofs, 2014). The recombinant plasmid for amyE disruption was constructed as pHYAMC (Fig. 1B). The amyE-integration cassette contained 0.8-kb front and 1.0-kb back homologous regions of the amyE gene flanked by the cat gene (Fig. S2).

Results

Vector construction and plasmid segregation efficiency

In this study, a genome integration and selection protocol was developed based on a Bacillus-replicative vector, in order to facilitate genome modification in undomesticated Bacillus strains (Fig. 2). The plasmid contains two parts including (i) a Bacillus-replicative vector backbone, and (ii) homologous regions of the target gene flanked with an antibiotic resistance gene. Bacillus-replicative vector was capable of genetic engineering in E. coli and self-propagation in Bacillus. The pHYA-harboring Bacillus clones were selected by the tetracycline-resistant (TcR) phenotype. The target site of the genome was constructed as a homologous recombinant cassette. The common distal genome integration site in the B. subtilis amyE (encoded alpha-amylase) was used as a gene model for genome integration. The phenotype of the amyE mutant strain was readily detected by starch degradation (Guérout-Fleury et al., 1996; Trauth and Bischofs, 2014). The recombinant plasmid for amyE disruption was constructed as pHYAMC (Fig. 1B). The amyE-integration cassette contained 0.8-kb front and 1.0-kb back homologous regions of the amyE gene flanked by the cat gene (Fig. S2). The recombinant plasmid for amyE disruption was constructed as pHYAMC (Fig. 1B). The amyE-integration cassette contained 0.8-kb front and 1.0-kb back homologous regions of the amyE gene flanked by the cat gene (Fig. S2).

To successfully apply the Bacillus-replicative vector for genome modification, a high rate of plasmid segregation is required. To estimate the plasmid segregation efficiency, strains of B. subtilis and B. cereus harboring pHYA were used as model strains. When cultured in an antibiotic-free medium, B. cereus rapidly segregated pHYA up to 94% after one cultivation cycle (approximately 12 h), while B. subtilis required at least seven cultivation cycles to reach the same level (Fig. 3). After three cultivation cycles, B. cereus segregated the plasmid more than 99%, while B. subtilis segregated the plasmid by 35%.

In addition, different target gene could also be applied as a homologous recombinant cassette. In this study, the proB-genome-integration cassette was constructed as pHYKOB (Table S1), and, later on, was effectively integrated into B. subtilis genome (as described below).

Genome integration and clonal selection protocol development

The genome integration of the cat-flanked-amyE gene (i.e. amyE mutant) was performed in a type-strain B. subtilis 168 using pHYAMC as a demonstrative guideline for the genome integration protocol by a Bacillus-replicative vector (Fig. 2). After plasmid delivery, cells were grown in several cultivation cycles (up to 9 cycles) in the antibiotic-free medium to facilitate homologous recombination and plasmid segregation prior to clonal selection on Cm-containing LB agar. The amyE-cat genome-integrated clones with a TcR-CmR phenotype were obtained at the 7th cultivation cycle with a clonal selection efficiency of 4.8 ± 0.4% (Fig. 4, -Ni -CT condition). The amyE mutants were then verified by the genotype by PCR amplification of amyE gene from the genomic DNA using F-AMYE-XbaI and R-AMYE-KpnI primers. The PCR product size of ca. 2.6 kb (amyE::cat) from amyE mutant clones confirmed the success of amyE-genome integration, when...
compared with that from the WT at ca. 2 kb (amyE) (Fig. 5A). In addition, the mutants were also tested for their phenotype by an amylase plate assay. All tested clones lost the amylase activity, while the wild-type clones retained the activity (Fig. 5B). To test the genetic stability, the mutants were cultured in the antibiotic-free medium for up to 9 cultivation cycles, and all tested clones were remained resistant to chloramphenicol. These results showed the desirable characteristics of both the genotype and the phenotype in the mutants.

Although genome integration using the developed vector and protocol was successful, the probability of clonal selection was still low. Less than 10% of the selected clones were obtained from the 7th cultivation cycles of cells in the antibiotic-free medium. Therefore, the clonal selection method was improved based on the leakiness of the tetracycline resistance gene (TcR or tetA) in the pHYA vector. The TcR gene encoded the tetracycline multiple-drug efflux pump, which exports toxic tetracycline out of cells; however, it also takes up toxic heavy metal ions from the medium (Bochner et al., 1980). In this case, to reduce the number of unintegrated plasmid-containing cells, nickel and cadmium ions were tested as metal ions that potentially inhibit the growth of cells harboring a pHYA-based vector with a TcR gene, while they should not affect the growth of B. subtilis WT. The concentration of NiCl2 that did not affect the growth of B. subtilis WT was at 0.5 mM (Fig. S3), while it caused 66% growth inhibition in pHYA-harboring cells (Fig. 6A). Thus, Ni at 0.5 mM was included in the clonal selection step. On the other hand, cadmium at a low concentration completely inhibited cell growth, thus it was not used (Fig. S3). In the presence of 0.5 mM NiCl2, the TcR-CmR clone selection efficiency was significantly increased by 9 fold up to 42.0 ± 10.2% at the 7th cultivation cycle (two-tailed unpaired t-test p = 0.0006) (Fig. 4, +Ni -CT condition).

In addition, to further increase the probability of clonal selection, the number of clones harboring an unintegrated vector could be reduced with an increase of nickel sensitivity. Thus, the nickel sensitivity was increased by induction of TcR expression. Various TcR gene inducers that do not inhibit the growth of the wild type were tested.
Generally, tetracycline inhibits the growth of the wild type cells at selective concentrations (10–20 mg/L). The effect of tetracycline at a lower concentration, i.e. 5 mg/L, and the autoclaved tetracycline (20–50 mg/L) were tested, but they still inhibited the growth of wild type (data not shown). As an alternative, chlorotetracycline, a low-toxicity tetracycline-derivative, was then tested. The result showed that a high concentration of chlorotetracycline, up to 20 μM, could be used without inhibition of the growth of the wild type cells (Fig. S4), and thus, it was tested in the clonal selection step. However, when chlorotetracycline (20 μM) was incorporated into the clonal selection step in addition to NiCl₂ (0.5 mM) (Fig. 4, +Ni +CT condition), it did not increase the chance of obtaining TcS-CmR genome-integrated clones (Fig. 4). As a result, the optimal clonal selection condition for *B. subtilis* was to cultivate cells harboring plasmid in the antibiotic-free medium for 7 cultivation cycles, and then subculture into LB supplemented with chloramphenicol (5 mg/L) and NiCl₂ (0.5 mM). The result described in this section thoroughly validated the effectiveness of the protocol for genome integration in the type-strain *Bacillus subtilis* 168.

**Genome integration in undomesticated strains of *B. subtilis* and *B. cereus***

After the validation of the genome-integration protocol in the type-strain of *Bacillus subtilis*, the developed vector and protocol were further applied for genetic manipulation to undomesticated strains of *B. subtilis* and *B. cereus*. Three *B. subtilis* strains with an organic-solvent-tolerant property were used. These included the sea-water-isolated *B. subtilis* strain GRSW1-B1 and strain GRSW2-B1 previously reported as butanol-tolerant and salt-tolerant strains (Kataoka et al., 2011), and the soil-isolated *B. subtilis* strain 3C previously reported to be high tolerant to toluene and diethyl phthalate (Navacharoen and Vangnai, 2011). As with most undomesticated bacteria, the genetic manipulation technique of these organic-solvent tolerant bacteria have not yet been proven. In this study, the transformation of pHYA into undomesticated *Bacillus* strains by electroporation was successful with significantly different transformation efficiencies (Table S3). This result clearly showed that not only low efficiencies of chromosome integration were a problem, but low efficiencies of introducing foreign DNA were also a problem in undomesticated strains as well. This result supported the concept of this work that, namely, once the transformation was achieved, multiplication of the introduced plasmid was required in order to increase the chance of genome integration in such a bacterial host. In this study, genome integration of these undomesticated strains by either natural transformation or electroporation using non-replicative plasmids, including pDG364, pDG1730 (Guérout-Fleury et al., 1996), and pUC-KOB, were failed, while it was accomplished in the domesticated *B. subtilis* strain 168 (Guérout-Fleury et al., 1996; Mahipant et al., 2017). On the other hand, with the application of a *Bacillus*-replicative vector, and the optimal clonal selection conditions reported in this study, pHYAMC was successfully integrated into the genomic DNA of undomesticated *B. subtilis* strains GRSW1-B1, GRSW2-B1 and 3C creating TcS-CmR clones with an efficiency of 5.0 ± 1.0, 13.3 ± 6.2, and 7.3 ± 1.0% (+Ni, -CT condition), respectively. This result strengthened the importance of plasmid replication ability increasing the chance of the homologous gene recombination in undomesticated *B. subtilis*.

Beside *amyE* gene, the developed vector and protocol were successfully applied for gene integration at another site of the genome as well. As a demonstration, the gene encoding glutamate 5-kinase (*proB*) that is involved in the proline biosynthesis pathway in *B. subtilis* 168 (Mahipant et al., 2017), was a target for gene disruption using a *Bacillus*-replicative vector. In this case, the pHYKOB vector (i.e. the vector harboring *spc*-flanked-*proB* fragment for genome integration at the *proB* gene) was transformed into *B. subtilis* GRSW2-B1 and GRSW1-B1. The result showed that the *proB* gene of both strains was successfully disrupted with an efficiency of 5% (data not shown). The phenotype of *proB* mutants was also confirmed. With-
out proline, both mutants were unable to grow in the minimal medium (Fig. S5). These results confirmed the possibility of the application for other target genes.

Moreover, the developed method could be applied with different Bacillus sp. In this case, B. cereus strain ATCC14579 was used. As an important food pathogen, this strain was reported to be tolerant to acid, osmotic and antibiotic stress (Mols et al., 2007). To perform genome integration in B. cereus ATCC14579, the pHYKOBc was used. The optimal condition was tested for application in B. cereus in terms of the plasmid segregation percentage and the optimum nickel concentration as described earlier. The cells cultured in an antibiotic-free medium for one cycle were selected for clonal selection in LB supplemented with 1.0 mM NiCl2 (Fig. 6B) and 5 mg/L chloramphenicol. The proB genome integration using pHYKOBc was accomplished in B. cereus ATCC14579 with ca. 3% efficiency, and the genotype was confirmed by PCR. This result endorsed the possibility of using the Bacillus-replicative vector in other related species.

Discussion

To accomplish genome modification in undomesticated Bacillus, previous studies have used temperature-sensitive (Ts) shuttle vectors to deliver the recombinant DNA into cells. In general, the self-replication ability of Ts plasmids in the host cells at a low temperature (~30–35°C) compensates for the low transformation efficiency (Le Breton et al., 2006; Leenhouts et al., 1991; Wilson et al., 2007). However, with this Ts vector, not only are several rounds of screening and selection of the transformants needed, but also the last step, incubation at a high temperature (e.g. 42–55°C) is strictly required to eliminate unintegrated Ts plasmids (Qiu et al., 2014, 2016; Wang, P. et al., 2016; Wang, Q. et al., 2012; Zakataeva et al., 2010). As a consequence, the application of a Ts vector is perhaps suitable for thermophilic or thermotolerant bacteria. This point is a limitation of applying a Ts vector to newly isolated strains and undomesticated strains of which the genetic and physiological backgrounds are unknown. In this study, a similar concept regarding the self-replication ability of the plasmid was adopted as a Bacillus-replicative vector to increase the possibility of obtaining the transformants. Nevertheless, even though this technique required several rounds of cultivation for homologous recombination to occur, plasmid segregation did not rely on the temperature sensitivity. In addition, an effective protocol was developed for clonal selection. As a consequence, the genome integration technique reported here is an effective system to apply to undomesticated Bacillus.

In this study, pHY300PLK was selected as the backbone for vector construction. The pHY300PLK is a widely used E. coli-Bacillus vector with a thermostatle property and has a high copy number. It harbors the pAMZl origin of replication from Streptococcus faecalis (Ishiwata and Shibahara, 1985). The estimated copy number of pHY300PLK in B. subtilis is 50 copies per cell (Morimoto et al., 2008). pHY300PLK has been widely used in many Bacillus species e.g. B. subtilis, B. cereus, B. licheniformis, B. brevis, B. nematocida, B. thuringiensis, B. antracis, and B. pumilus (Cai et al., 2016; Chen et al., 2003; Ishiwata and Shibahara, 1985; Lertcanawanichakul et al., 2004; Makino et al., 2002; Niu et al., 2015; Oda et al., 2012; Shaw et al., 2000; Toth et al., 2016). In addition, this vector has been successfully used in other related Bacillus such as Lactobacillus, Streptococcus, Paenibacillus, Staphylococcus, and Listeria (Fukuda et al., 2009; Hanawa et al., 2002; Kimoto and Taketo, 2003; Shiratsuchi et al., 2010; Watanabe et al., 1994). It has broad range thermo-stability as it was reported in a psychrophilic Listeria monocytogenes (Okada et al., 2006) as well as a thermophilic B. licheniformis under the selective conditions (Niu et al., 2009; Qiu et al., 2016). Accordingly, the vector potentially has a broad application in other related bacteria as well.

The plasmid segregation efficiency is an important factor for the application. To maintain bacterial fitness under non-selective conditions, the bacteria normally lose plasmids that are no longer necessary (Smith and Bidochka, 1998). Nonetheless, different species may show large differences on the plasmid segregation. The results of this study agree with those of a previous study in which a pHY300PLK-based vector was more stable in B. subtilis than B. cereus. Within 2 cultivation cycles in non-selective conditions, the pHY300PLK-derived plasmid was reported to be stable in isolated B. subtilis (Katoaka et al., 2011). On the other hand, the pHY300PLK-derived plasmid was segregated promptly at the first cultivation cycle in B. cereus, and probably B. thuringiensis, which is categorized in the same group as B. cereus (Lertcanawanichakul and Wiwat, 2000). Without either homologous recombination or replication, the foreign DNA was rapidly degraded by endonucleases as an immune-responsive mechanism in bacteria (Anders et al., 2014). These data showed that the plasmid (pHYA) is unstable under non-selective conditions and readily segregates, and, thus, has a potential application for genome integration in different groups of Bacilli.

The use of a Bacillus-replicative vector was not without flaw since the probability of clonal selection was low. Alternatively, the clonal selection can be scaled up and applied to robotic systems as was done in the selection using microtiter plates. To increase the probability of clonal selection without further genetic manipulation, the sensitivity to nickel of cells harboring the Te6 gene was used. With the improved clonal selection protocol with nickel, the likelihood of obtaining the Te6-Cm6 clones was much higher. In general, nickel is used as a co-factor for several enzymes in bacterial cells, including Bacillus sp., and, thus, it is required at a low concentration. Nickel at a high concentration may be harmful to bacterial cells, although the nickel tolerance of bacteria is varied and is strain-dependent. In this study, the effect of nickel concentration to the cell growth of B. cereus wild-type and B. cereus harboring pHYA is shown in Fig. 6B. This result is straightforward and shows that the growth of B. cereus wild-type increased with an increase of nickel concentration from 0.5–1.0 mM. As for B. cereus harboring pHYA, since pHYA contains the tetracycline-resistant gene (Te6) that codes for multi-drug efflux pump, more nickel can be
pumped into the recombinant cells resulting in higher cell growth at Ni 0.5 and 0.75 mM. Because *B. cereus* harboring pHYA could uptake nickel more than that of the wild-type cells (i.e. the recombinant cells have a higher sensitivity to nickel), exposure of cells to 1 mM nickel (as well as at 1.5 and 2 mM) resulted in higher cell growth inhibition compared to that of the wildtype.

For future developments, driving the TsR expression by inducible promoters is expected to efficiently reduce the plasmid-harboring cell population, as described in a previous study in *E. coli* (Nomura and Yokobayashi, 2007). Alternatively, the vector can be further improved by combining with complex markerless methods to eliminate the unwanted region (Dong and Zhang, 2014). For example, the combination of a Ts vector and recombinase were applied for target genome integration in *B. licheniformis* BL1 (Wang, Q. et al., 2012), while the Ts plasmid and the recombinase in Mob/oriT system were used for genome integration in *B. thuringiensis* and *B. cereus* (Wang, P. et al., 2016). Although these developed techniques reported a success, they required a tremendous time for unintegrated plasmid curing by repeated bacterial sub-culturing and clonal selection. On the other hand, this study not only reported the effectiveness of a *Bacillus*-replicative vector for genome integration in undomesticated *Bacillus* strains, but also reported the simple clonal selection to obtain the homologous recombinant strains. This alternative and simple method of genome modification can potentially be applied to a wide range of isolated *Bacillus* and other related bacteria.

Acknowledgments

A research grant to GM and ASV by the Thailand research fund and Chulalongkorn University through the Royal Golden Jubilee PhD program (RGI-PhD.) (Grant No. PHD/0078/2553) and the 90th Anniversary of Chulalongkorn University Fund (Ratchaphadisekomphot Endowment Fund) are appreciated. This work was partially supported by The Japan Society for the Promotion of Science (JSPS) and the National Research Council of Thailand (NRCT) (Bilateral Project).

Supplementary Materials

Supplementary figures and tables are available in our J-STAGE site (http://www.jstage.jst.go.jp/browse/jgam).

References


