

# SCIENTIFIC REPORTS



OPEN

## Far rapid synthesis of giant DNA in the *Bacillus subtilis* genome by a conjugation transfer system

Mitsuhiro Itaya<sup>1</sup>, Mitsuru Sato<sup>1</sup>, Miki Hasegawa<sup>1</sup>, Nobuaki Kono<sup>1</sup>, Masaru Tomita<sup>1</sup> & Shinya Kaneko<sup>2</sup>

*Bacillus subtilis* offers a platform for giant DNA synthesis, which is mediated by the connection of overlapping DNA segments called domino DNA, in the cloning locus of the host. The domino method was successfully used to produce DNA fragments as large as 3500 kbp. However, domino DNA is limited to <100 kbp because of size restrictions regarding the transformation (TF) of *B. subtilis* competent cells. A novel conjugal transfer (CT) method was designed to eliminate the TF size limit. The CT method enables rapid and efficient domino reactions in addition to the transfer of giant DNA molecules of up to 875 kbp to another *B. subtilis* genome within 4 hours. The combined use of the TF and CT should enable significantly rapid giant DNA production.

*Bacillus subtilis* Marburg 168, a Gram-positive spore-forming bacterium, is used as a platform for giant DNA synthesis<sup>1–3</sup>. The DNA synthesis is achieved by the connection of overlapping DNA molecules, called domino, at the cloning locus prepared in the *proB* gene locus of the *B. subtilis* genome, as illustrated in Fig. 1. The domino method has been used for the synthesis of various genomes, such as those of mouse mitochondria (16.7 kbp)<sup>2</sup>, lambda phage (48.5 kbp)<sup>4</sup>, rice chloroplast (135 kbp)<sup>2</sup>, and those of *Synechocystis* spp PCC6803 (3542 kbp)<sup>1</sup>. The domino method is based on the inherent natural transformation (TF), which is accurate enough to incorporate DNA<sup>3</sup>. Domino connection through TF requires liquid genomic DNA biochemically isolated from donor *B. subtilis*. As shown in Fig. 1, the liquid DNA is added to recipient competent cells. However, TF efficiency dramatically decreases with the increase in the size of the liquid DNA to >100 kbp<sup>5–8</sup>. Accordingly, our latest TF-based protocol restricts one domino size to ≤100 kb<sup>8</sup>.

We developed a conjugal transfer (CT) system for transmission of the *B. subtilis* genome to eliminate the TF-associated size restriction<sup>9</sup>. The CT enables the transfer of antibiotic resistance markers from the donor *B. subtilis* genome, distant from the origin of transfer *oriT*<sup>10</sup> at the *xkdE* locus, to the recipient *B. subtilis* genome by mating within 4 hours at 30 °C<sup>9</sup>. Large DNA inserted in the *B. subtilis* genome should be similarly transferable with the help of a Type IV secretion system called T4SS, which is encoded and supplied by the plasmid pLS20hyg<sup>9</sup>. Orientation of the *oriT*<sup>10</sup> from BEST390<sup>9</sup> and in Table 1 was used throughout the study.

### Methods and Results

As illustrated in Fig. 2, the CT was used to connect two dominos (196 kbp and 171 kbp with a 22 kbp overlap) covering the mouse *jmj* region to yield the 345 kbp product efficiently and accurately.

The original domino strains, BEST6568 (171 kbp, SpcR) and BEST6571 (196 kbp, BSR), which are described in Fig. 2 and Table 1, were converted to CT donors. Two functional elements should be added<sup>9</sup>, namely the sequence *oriT*<sup>10</sup>-*stm* unit, introduced by TF from BEST390 and selected with streptomycin, and pLS20hyg, transferred by CT from BEST23124 and selected with hygromycin<sup>9</sup>. Donors are illustrated in Fig. 2 and intermediate strains are listed in Table 1. Mating starts by simple mixing of the stationary donor culture (50 µL) and stationary recipient culture (50 µL) in 2.0 mL fresh LB in a 15 mL-culture tube in the presence of DNase I (final concentration, 3.4 µg per mL). The recipient *B. subtilis* that accepts the DNA transferred from the donor cell by CT is called the transcient. After incubation at 30 °C for 4 hours, the transcient was selected on LB plates supplemented with Blasticidin S (BS, 500 µg per mL) and Spectinomycin (Spc, 150 µg per mL). As listed in Table 1, the domino

<sup>1</sup>Institute for Advanced Biosciences, Keio University, Nipponkoku, Daihoji, Tsuruoka-shi, Yamagata, 997-0017, Japan.

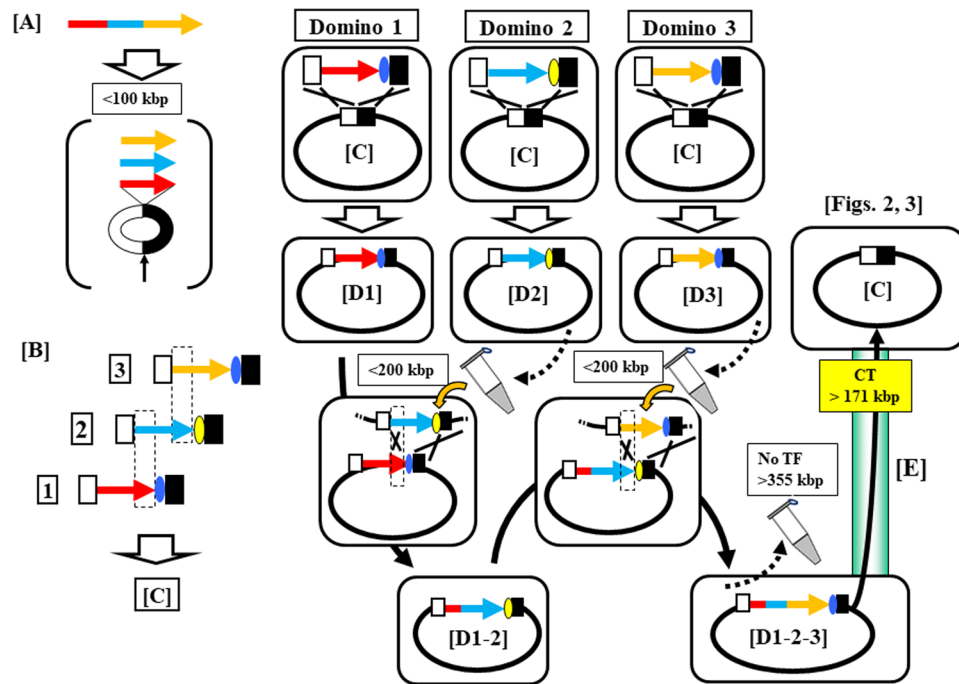
<sup>2</sup>School of Life Science and Technology, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama-shi, Kanagawa, 226-8501, Japan. Correspondence and requests for materials should be addressed to M.I. (email: [mita2001@sfc.keio.ac.jp](mailto:mita2001@sfc.keio.ac.jp))

Size <sup>a</sup>	<i>B. subtilis</i> <sup>b</sup>	Genotypes <sup>b</sup>	Antibiotic markers <sup>c</sup>	transcipts <sup>d</sup>	transformants <sup>e</sup>	Domino connection	
	RM125 <sup>9</sup>	<i>leuB8, arg-15, hsdRM</i>					
	BEST386 <sup>9</sup>	pHY300PLK	Tc				
5.7	BEST390 <sup>9</sup>	<i>xkdE::oriT<sup>110</sup>-stm</i> unit	Sm		362 (SmR)		
	BEST23124 <sup>9</sup>	pLS20hyg	Hm				
1.3	BEST4306 <sup>9</sup>	<i>xkdE::oriT<sup>110</sup>-stm</i> unit, pLS20hyg, <i>yjcl::ne-spc</i> , <i>pycA::eo-bsr</i>	Sm Hm Spc BS	318			
						recipient BEST6571	recipient BEST6568
	BEST6568 <sup>7</sup>		Spc				
171	BEST6621	<i>oriT<sup>110</sup>-stm</i>	Sm Spc		1 (SpcR)		
	BEST6627	<i>oriT<sup>110</sup>-stm</i> , pLS20hyg	Hm Sm Spc	176		23 <sup>f</sup>	
	BEST6571 <sup>7</sup>		BS				
196	BEST6623	<i>oriT<sup>110</sup>-stm</i>	Sm BS		18 (BSR)		
	BEST6629	<i>oriT<sup>110</sup>-stm</i> , pLS20hyg	Hm Sm BS	175			33 <sup>g</sup>
	BEST6586 <sup>7</sup>		BS Spc				
345	BEST6625	<i>oriT<sup>110</sup>-stm</i>	Sm BS Spc		0 (SpcR)		
	BEST6631	<i>oriT<sup>110</sup>-stm</i> , pLS20hyg	Hm Sm BS Spc	146			
	BEST7095 <sup>1</sup>		Sp Em				
509	BEST7817	<i>oriT<sup>110</sup>-stm</i>	Sm Spc Em		0 (SpcR)		
	BEST7817hyg	<i>oriT<sup>110</sup>-stm</i> , pLS20hyg	Hm Sm Spc Em	125			
	BEST7103 <sup>1</sup>		Sp Em				
586	BEST7821	<i>oriT<sup>110</sup>-stm</i>	Sm Spc Em		0 (SpcR)		
	BEST7821hyg	<i>oriT<sup>110</sup>-stm</i> , pLS20hyg	Hm Sm Spc Em	86			
	BEST7118 <sup>1</sup>		Sp Em				
683	BEST7823	<i>oriT<sup>110</sup>-stm</i>	Sm Spc Em		0 (SpcR)		
	BEST7823hyg	<i>oriT<sup>110</sup>-stm</i> , pLS20hyg	Hm Sm Spc Em	140			
	BEST7154 <sup>1</sup>		Sp Em				
875	BEST7819	<i>oriT<sup>110</sup>-stm</i>	Sm Spc Em		0 (SpcR)		
	BEST7834	<i>oriT<sup>110</sup>-stm</i> , pLS20hyg	Hm Sm Spc Em	3			

**Table 1.** *B. subtilis* strains used for TC and TF. <sup>a</sup>Size of DNA in the *proB* gene, except BEST4305 and BEST390, which carry a blasticidin resistance gene (1.3 kbp) near *proB*<sup>9</sup> and *oriT<sup>110</sup>-stm* unit (5.4 kbp) at the *xkdE*<sup>9</sup>, respectively. <sup>b</sup>Previously published strains are shown with references in brackets. Strains constructed in this study are italicized. For strains referring<sup>7</sup>, their actual insert sizes are determined this study. <sup>c</sup>Blasticidin S (Bs, 500 µg/mL), hygromycin (Hm, 150 µg/mL), spectinomycin (Spc, 100 µg/mL), streptomycin (Sm, 100 µg/mL), and tetracycline (Tc, 10 µg/mL). <sup>d</sup>BEST386 (TcR) is used as the common recipient for all donors. Number of transcipts/mL in 4 hours by standard CT protocol<sup>9</sup> selected by Spc and Tc, except for BEST6629 by BS and Tc. <sup>e</sup>RM125 is used as the common recipient for TF with all genomic DNA. Volume of the competent cells is increased eight times that using standard transformation<sup>1,5</sup>. The antibiotics used for selection of the transformants are indicated in parentheses. <sup>f</sup>A representative named CT[171/196] was analyzed by PFGE (Fig. 2) and sequenced. <sup>g</sup>A representative named CT[196/171] was analyzed by PFGE (Fig. 2) and sequenced.

connection by CT yielded 23 transcipts (171 kbp-carrier as donor) and 33 transcipts (196 kbp-carrier as donor) per mL. Representatives of the 23 and 33 transcipts were analyzed by Pulsed Field Gel Electrophoresis (PFGE) as shown in Fig. 2. The production of 345 kbp inserts of the transcipts was confirmed. Sequence comparison of the insert from the donor (171 kbp and 196 kbp) and the transcipt (345 kbp) showed no nucleotide alterations, indicating that the domino connection was achieved accurately.

Next, we investigated whether the CT allows the transmission of large DNA fragments of >200 kb, which was difficult by TF as shown in Fig. 1. We selected four strains carrying synthesized DNA of 509, 586, 683, and 875 kbp, originated from the Cyanobacterium *Synechocystis* spp. PCC6803 genome<sup>1</sup> as shown in Fig. 3. The original strains shown in Fig. 3 and listed in Table 1 were converted to CT donors by the same method described in Fig. 2. Mating with the common recipient BEST386, which was previously described<sup>9</sup> and shown in Fig. 3, and selection of transcipts on LB plates supplemented with spectinomycin (Spc, 150 µg per mL) and tetracycline (Tc, 10 µg per mL) yielded transcipts (per mL), 125, 86, 140 and 3, as summarized in Table 1. Representatives of each transcipt were analyzed by PFGE, as shown in Fig. 3, which demonstrated the transfer of undisrupted DNA of up to 875 kbp. Sequence comparison of the inserts from donors and transcipts revealed no nucleotide alterations. The efficiency of CT from donors of mouse domino strains to the recipient BEST386 was assessed, and the transcipts were scored as (per mL), 176 (196 kbp), 175 (171 kbp), and 146 (345 kbp), as summarized in Table 1. It seems likely that CT for *B. subtilis* proceeded in a largely size-independent manner, consistent



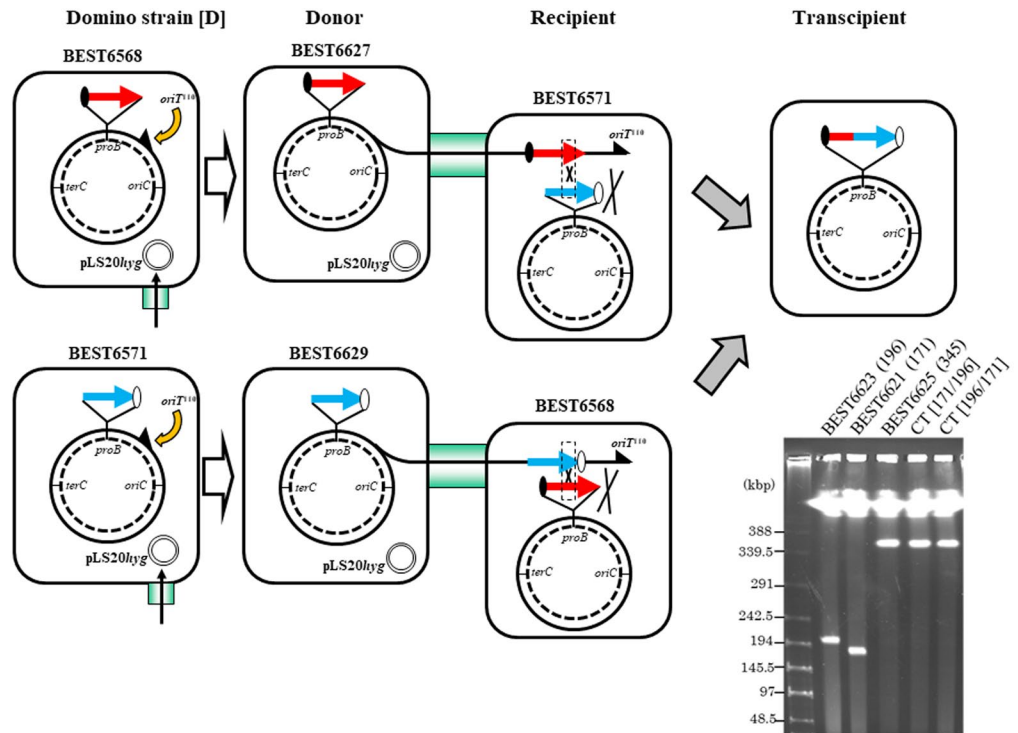
**Figure 1.** TF-mediated domino method in the *B. subtilis* genome requires liquid DNA and limits DNA size to <100 kbp per domino step. (A) The target DNA to be synthesized is divided into domino DNA. (B) Designed dominos possessing overlapping DNA sequences are prepared by the *E. coli* plasmids BAC<sup>7,8</sup> or pBR322<sup>1</sup> depending on the domino size. The size of dominos is restricted to <100 kbp even for BAC-based dominos. Dominos 1, 2, and 3 are shown in linearized form opened at the site of *E. coli* vectors, as indicated by the vertical arrow. Overlapping sequences between adjacent dominos are shown as rectangular dots. The red or blue ovals indicate antibiotic resistance genes conferring CmR or EmR, which are alternatively used for the selection of connected strains<sup>1,2,8</sup>. (C) The three domino DNA numbers are inserted in the *proB* locus by double homologous recombination between the *E. coli* vector sections shown by X, resulting in (D1) to (D3). Domino 2 elongates by double homologous recombination at the overlapping sequence, as shown by the rectangular dots between (D1) and (D2) DNA (the left X) and at the *E. coli* vector (right X). Similarly, Domino 3 integration is completed by the double homologous recombination sites. Note that selection markers are alternated at the elongation step. The small tube icon indicates genomic DNA prepared in liquid form for TF, indicating an average DNA size of <200 kbp. (E) The expected DNA transfer by the CT system of >171 kbp is indicated by solid arrows through a light green box. This is confirmed in Figs 2 and 3 and Table 1.

with the Hfr frequency for F-mediated *Escherichia coli*, which is constant throughout the movement of the whole genome<sup>10,11</sup>. The reason for the decrease in the number of transcipts for BEST7834 (875 kbp) remains unclear. A slower growth rate of the donor BEST7834 (875 kbp) (our unpublished observations) may have affected the efficiency of CT.

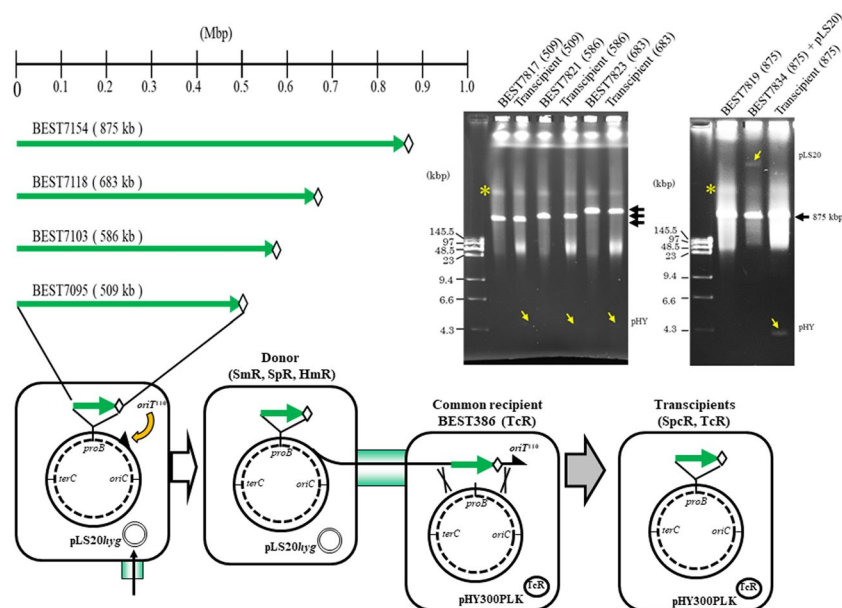
Regardless of the CT efficiency, the transformants generated by TF, a few from 171, or 196 and none from DNA 345 kb, as summarized in Table 1, are consistent with our previous conclusion<sup>6–8</sup>.

## Discussion

The TF-mediated connection of one domino DNA of 100 kbp requires at least 2 weeks, including the preparation of high quality genomic DNA for TF and selection, followed by screening by antibiotic resistance, and confirmation of insert structures by PFGE. More than 40 weeks was required for the synthesis of 875 kbp DNA by TF<sup>1</sup>. One domino connection by CT is expected to require 1 week including the introduction of two functional elements and the selection of transcipts. Considering that CT-based DNA connection as domino donor and/or recipient can be achieved for up to 683 kbp, the time required for the connection of larger DNA fragments should be dramatically reduced. It should be addressed that the lack of nucleotide alterations caused by CT could minimize the screening for correct transcipts. In addition to the currently established genome synthesis methods using yeast<sup>12,13</sup>, the CT should be a valuable additional tool to accurately edit and synthesize large DNA by *B. subtilis*. The synthesis of existing DNA, even that with a high GC content of up to 70% from *Thermus thermophilus*<sup>8</sup>, chloroplast DNA carrying a long 25-kbp inverse repeat<sup>2</sup>, and mouse genomes with abundant small repeats<sup>7,14,15</sup> has been demonstrated. In addition to existing DNA, preparation of long DNA *de novo* can be achieved by our method called OGAB<sup>16,17</sup>. Since few nucleotide changes occur during TF<sup>1,16,18</sup>, CT can be used as a daily tool to synthesize DNA fragments of >1000 kbp using long DNA fragments as initial dominos. *B. subtilis* allows the



**Figure 2.** Domino connection by CT Top: The original BEST6568 was converted to donor BEST6627 by transfer of the *oriT<sup>110</sup>* unit, as shown by a half triangle, and the helper plasmid pLS20hyg. Insertion of the *oriT<sup>110</sup>* unit was achieved by TF from BEST390, as indicated by a curved yellow arrow. Introduction of pLS20hyg was mediated by CT from BEST23124, as indicated by a vertical arrow through a light green box. The donor BEST6627 carries a spectinomycin resistance gene (solid oval) located at the left end of the 171-kbp insert (red bold arrow). Single DNA transfer by the CT system is represented by solid arrows through a light green box. In the mating with the recipient BEST6571 containing a blasticidin S resistance gene (open oval) at the right end of the 196-kbp insert (blue bold arrow), the 171-kbp of transferred DNA recombines with the homologous sequences of the recipient, as indicated by two Xs. Twenty three transipients with a 345-kbp carrier were selected by markers at both ends as described in Table 1. Bottom: The donor and recipient are reversed. Conversion from BEST6571 to BEST6629 was performed as described above. Thirty three transipients were selected by both markers as summarized in Table 1. Inserts by *I-PpoI* digestion<sup>1,2,7,8</sup> from the indicated *B. subtilis* strains were separated by PFGE. Running conditions: 1% agarose gel with 0.5 × TBE buffer; constant 5 volts/cm; 24 sec pulse time; and 24 hours running time at 14 °C.



**Figure 3.** Transfer of DNA as large as 875 kbp by CT. *B. subtilis* strains with the indicated DNA inserts from 509 kbp to 875 kbp, which were reported previously<sup>1</sup>, were converted to the donors listed in Table 1 by the

same sequential introduction of the two devices described in Fig. 2. The common recipient BEST386, carrying tetracycline resistance (TcR) by pHY300PLK, is shown and listed in Table 1. The donor marker is Spectinomycin (Spc) + erythromycin (Em), as indicated by an open diamond. Inserts of representative transcient and donor derivatives (sizes are shown within parentheses) were analyzed by PFGE after complete digestion by *I-PpoI*. Left gel photo: the positions of the *I-PpoI* segments of 509, 586, and 683 kbp are indicated by bold horizontal arrows on the right. The same sizes for each donor and transcient are identified. The size marker lambda DNA ladder (48.5 kbp x n) and *HindIII* digests are shown in the left with their sizes in kbp. The plasmid pHY300PLK (4.7 kbp), indicated by pHY on the right, is identified by yellow arrows (↘) in the photo. (\*) may be produced from *B. subtilis* genome by unidentified *I-PpoI* star activity (see also Supplementary Fig. 1). Right gel photo: the 875 kbp *I-PpoI* segment of donor and transcient is indicated by a bold arrow on the right. The donor BEST7834 carries pLS20hyg (66 kbp), as indicated on the right by a yellow arrow (↙) (see also Supplementary Fig. 2). Note that running of circular plasmids >50 kbp on PFGE is difficult to predict<sup>10</sup>. Running conditions: constant 5 volts/cm; 360 sec pulse time; and 17 hours running time at 14 °C.

stable synthesis of DNA as large as 3500 kbp<sup>1</sup> as well as the synthesis of several mouse genome segments,<sup>7,14,15</sup> therefore, it might be suitable as a workhorse for the Human Genome Wright project<sup>19</sup>.

Lastly, the accurate transfer of 875 kbp DNA inserted in the genome is, to the best of our knowledge, the first and largest exogenous DNA horizontally transferred between *Bacillus* species. The present results and experimental tools might shed light on studies about horizontal DNA transfer<sup>1,3</sup>.

## References

- 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* **102**, 15971–15976 (2005).
- Itaya, M., Fujita, K., Kuroki, A. & Tsuge, K. Bottom-up genome assembly using the *Bacillus subtilis* genome vector. *Nat Methods* **5**, 41–43 (2008).
- Itaya, M. Systems biology and Synthetic Biology (eds. Fu, P. and Panke, S.), Ch. 5, 153–190 (John Wiley & Sons, 2009).
- Itaya, M. Toward a bacterial genome technology: integration of the *Escherichia coli* prophage lambda genome into the *Bacillus subtilis* 168 chromosome. *Mol Gen Genet* **248**, 9–16 (1995).
- Itaya, M. Genetic transfer of large DNA inserts to designated loci of the *Bacillus subtilis* 168 genome. *J Bacteriol.* **181**, 1045–1048 (1999).
- Itaya, M. *et al.* Stable positional cloning of long continuous DNA in the *Bacillus subtilis* genome vector. *J Biochem.* **134**, 513–519 (2003).
- Kaneko, S., Takeuchi, T. & Itaya, M. Genetic connection of two contiguous bacterial artificial chromosomes using homologous recombination in *Bacillus subtilis* genome vector. *J Biotechnol.* **139**, 211–213 (2009).
- Ohtani, N. *et al.* Serial assembly of *Thermus* megaplasmid DNA in the genome of *Bacillus subtilis* 168: a BAC-based domino method applied to DNA with a high GC content. *Biotechnol J.* **7**, 867–876 (2012).
- Itaya, M., Hasegawa, M., Tomita, M. & Sato, M. The first high frequency of recombination-like conjugal transfer from an integrated origin of transfer sequence in *Bacillus subtilis* 168. *Biosci Biotechnol Biochem.* **82**, 356–362 (2018).
- Itaya, M. *et al.* Efficient delivery of large DNA from *Escherichia coli* to *Synechococcus elongatus* PCC7942 by broad-host-range conjugal plasmid pUB307. *J Biochem.* in press.
- Gerald, R. S. Conjugal recombination in *E. coli*: myths and mechanisms. *Cell* **64**, 19–27 (1991).
- Isaacs, F. J. *et al.* Precise manipulation of chromosomes *in vivo* enables genome-wide codon replacement. *Science* **333**, 348–353 (2011).
- Gibson, D. G. *et al.* Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **319**, 1215–1220 (2008).
- Hutchison, C. A. *et al.* Design and synthesis of a minimal bacterial genome. *Science* **351**, aad6253 (2016).
- Iwata, T. *et al.* *Bacillus subtilis* genome vector-based complete manipulation and reconstruction of genomic DNA for mouse transgenesis. *BMC Genomics* **14**, 300, <https://doi.org/10.1186/1471-2164-14-300> (2013).
- Ogawa, T. *et al.* An inducible recA expression *Bacillus subtilis* genome vector for stable manipulation of large DNA fragments. *BMC Genomics* **16**, 209, <https://doi.org/10.1186/s12864-015-1425-4> (2015).
- Itaya, M. & Tsuge, K. *Methods Enzymol.* Vol. 498 (ed. Voigt, C.), 427–447 (Academic Press, 2011).
- Tsuge, K. *et al.* Method of preparing an equimolar DNA mixture for one-step DNA assembly of over 50 fragments. *Scientific Reports.* **5**, 10655, <https://doi.org/10.1038/srep10655> (2015).
- Boeke, J. D. *et al.* The Genome Project—Write. *Science* **353**, 126–127 (2016).

## Acknowledgements

This work was supported by Research funds from the Yamagata prefectural government and the City of Tsuruoka.

## Author Contributions

M.I. and M.S. designed the experiments. M.I., M.S., M.H. and S.K. performed *in vivo* experiments. M.I., N.K. and M.T. analyzed the sequence determination. M.I. wrote the manuscript. All authors read and approved the final manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-26987-0>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018