Improved Bacterial Mutagenesis by High-Frequency Allele Exchange, Demonstrated in *Clostridium difficile* and *Streptococcus suis*

Alexandra Faulds-Pain, Brendan W. Wren

Department of Pathogen Molecular Biology, London School of Hygiene and Tropical Medicine, London, United Kingdom

Here we show that the frequency of mutant isolation by two-step allele exchange can be improved by increasing the length of homologous DNA and the opportunity for recombination, obviating the need for counterselection markers. These principles are demonstrated in *Clostridium difficile* and *Streptococcus suis* but are likely to be generally applicable.

Mutagenesis is central to the study of microorganisms, enabling the functional role of specific genes to be determined (1, 2). There are many important organisms for which reliable allele exchange methods are not well established, including several pathogens. Allele exchange, or gene replacement, is perhaps the most widely implemented strategy of directed mutagenesis in microorganisms whereby a native allele is exchanged with an introduced, alternative allele containing a mutation. The alternative allele is flanked by DNA regions identical to the DNA regions flanking the original allele (Fig. 1A). Homologous recombination events between both pairs of identical DNA regions are required to mediate the exchange. When replication-defective plasmids (suicide, conditional, or "pseudosuicide" [3–8]) are used as the delivery vehicle for the allele exchange cassette, mutagenesis is a two-step process. The first recombination event, termed a single-crossover event, causes integration of the plasmid into the genome and is simply isolated when positive, plasmid-specific antibiotic selection is applied. A second recombination event (a double-crossover event) can then occur whereby the plasmid is excised from the genome. If this second event occurs at the same region of homology as the first event, then the cell reverts to wild type; if it occurs at the other region of homology, then the alternative allele is stably localized, causing the desired mutation (Fig. 1B).

The isolation of double-crossover clones can be achieved by incorporating a counterselection marker onto the allele exchange plasmid whose presence causes cell death under appropriate conditions (3, 9–12). Identification of a suitable counterselection marker is complex and often organism specific (11); however, it is often assumed that isolating mutants using two-step allele exchange is extremely laborious without a counterselection marker due to the low frequency at which the second recombination event occurs (11). We questioned whether this assumption is justified and whether the frequency of recombination could be increased, in a generally applicable way, to a level that would make two-step allele exchange facile and obviate the requirement of a counterselection marker.

We undertook to increase the frequency of the second recombination event in two ways. The first was by using long regions of homologous DNA which we hypothesized would increase the frequency of the second recombination event (5). The second was to increase the opportunity for recombination to occur by serially subculturing single-crossover integrants. This second principle is an advantage particular to two-step allele exchange, as members of a population of single-crossover clones all have the potential to

undergo a second recombination and become double crossovers, and this potential is not lost until a recombination event occurs.

Clostridium difficile is the main cause of antibiotic-associated diarrhea in health care settings. There are at least five clonal lineages of *C. difficile*, designated clades 1 to 5 (13–17), but to date, only strains from clades 1 and 2 have been studied by a directed mutagenesis approach. The most widely implemented method of *C. difficile* mutagenesis introduces the macrolide, lincosamide, streptagramin (MLS) antibiotic resistance marker *ermB* into MLS-sensitive strains (6, 7). We are interested in studying isogenic mutants of clinical isolates from clade 4 using this approach (13, 17). However, we found that strains in this clade were resistant to the MLS antibiotic lincomycin (Lm) (see Table S1 and Materials and Methods in the supplemental material) and that this was likely to be due to the presence of a chromosomal *ermB* gene (14, 18). Therefore, our first target for our alternative approach to allele exchange was the *ermB* gene in the clade 4 sequenced strain M68.

The allele exchange cassette for the deletion of *ermB* from M68 was constructed by SOE-PCR (19) with 1,200-bp regions of DNA homologous to the DNA sequences flanking the *ermB* gene (Fig. 1A) using a high-fidelity polymerase (Phusion; NEB). This was cloned into the replication-defective *C. difficile* vector pMTL82151 (20), suitable for pseudosuicide in *C. difficile* (4, 5), using standard cloning techniques, resulting in the plasmid pAFP1200 (see Table S2 in the supplemental material for primers and restriction enzymes). pAFP1200 was transferred into *C. difficile* M68 by conjugation (4, 7, 21), and single-crossover clones were obtained by passage three times on selective media as described previously (3, 4) (Fig. 1B). *C. difficile* was cultured routinely at 37°C in an anaerobic work station (Don Whitley) on BHIS medium (brain heart infusion [BHI] medium supplemented with 5 mg · ml^{-1} yeast extract and 0.1% [wt/vol] L-cysteine) supplemented with D-cycloserine (250 μg ·
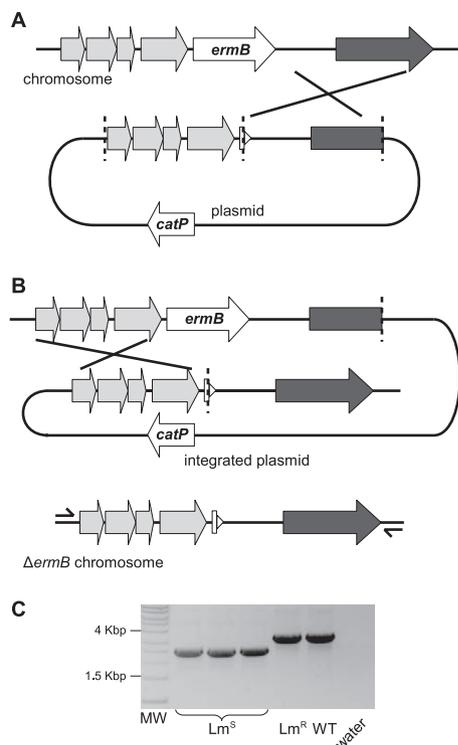
FIG 1 Diagrammatic representation of the precise deletion of *ermB* in *C. difficile* strain M68 by two-step allele exchange. (A) The first recombination event between the pAFP1200 plasmid and the *ermB* locus. In this example, integration at the right region of homology (dark gray) is represented. (B) The second recombination event causes a deletion of the *ermB* gene; if the recombination event occurs at the other region of homology to the first recombination event (light gray), the gene is deleted from the chromosome. (C) PCR screening confirms the *ermB* deletion mutants initially identified by replica plating on thiamphenicol (Tm), lincomycin (Lm), and nonselective media. PCR was performed using primers M68ΔermB screen F and M68ΔermB screen R (see Table S2 in the supplemental material). MW, Hyperladder I molecular weight marker; Lm^S, DNA from Lm-sensitive clones; Lm^R, DNA from Lm-resistant but Tm-sensitive clone; WT, wild-type DNA; water, water-only control.
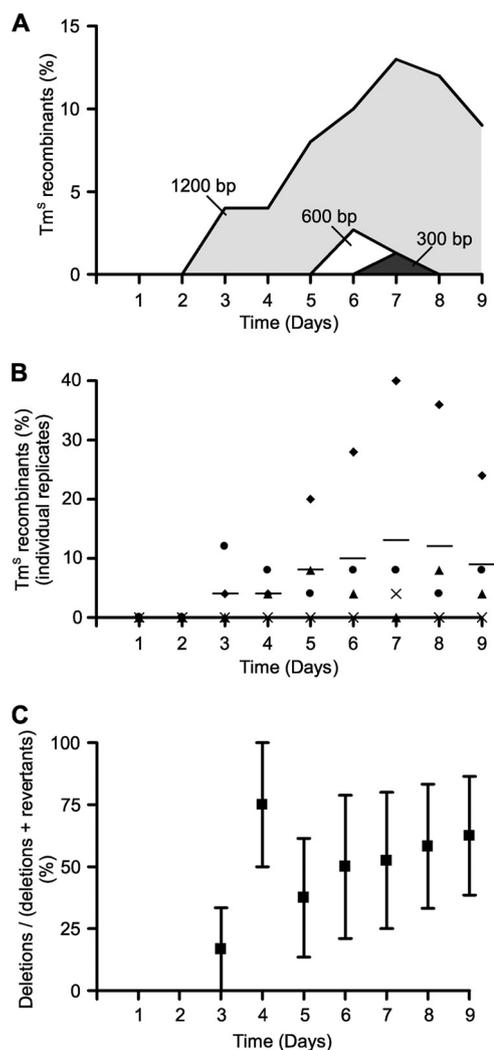

FIG 2 Recombination frequency and the importance of the lengths of the homology regions. (A) The graph represents the fraction of double-crossover recombinants as a percentage of the total clones screened. Single crossovers of the three *ermB* deletion plasmids with differing homology lengths (pAFP300, pAFP600, and pAFP1200) were passaged without selection and screened each day for loss of Tm resistance; the data plotted represent the average results of either three or four replicates, each of which screened 25 colonies a day. Light gray, 1,200-bp, double-crossover recombinants from pAFP1200 single-crossover recombinants; white, 600-bp, double-crossover recombinants from pAFP600 single-crossover recombinants; dark gray, 300-bp, double-crossover recombinants from pAFP300 single-crossover recombinants. (B) The graph represents the fraction of pAFP1200 single-crossover cells that had undergone a second recombination event and become Tm^s when passaged without selection every day for 9 days. Each of the symbols ▲, ●, ×, and ◆ represents a different replicate; the horizontal line represents the mean of the four repeats. (C) The graph represents the fraction of the pMT-LAFP1200 double-recombinant clones (deletions plus revertants) that are deletions, as a percentage, over time. The large filled squares represent the average results of four repeat experiments where 25 clones were screened per repeat; error bars represent the standard errors of the means.

ml^{-1}), cefoxatin (8 μg · ml^{-1}), and thiamphenicol (Tm) (15 μg · ml^{-1}) to maintain plasmid selection.

In the second part of the experiment, thiamphenicol selection was removed to allow growth of double-crossover clones. To allow more opportunity for the second recombination event to occur, the single-crossover clones were serially subcultured without selection for up to 9 consecutive days. Double-crossover recombinants were identified by replica plating colonies onto nonselective BHIS and BHIS media with thiamphenicol to screen for the loss of the plasmid-specific antibiotic resistance marker, *catP*. As the target gene, *ermB*, is likely to confer resistance to lincomycin, deletion mutants could also be distinguished from revertants by replica plating on lincomycin (40 μg · ml^{-1}). We found that following subculture, thiamphenicol-sensitive clones (Tm^S; recombinants) that were either lincomycin sensitive (Lm^S; deletion mutants) or lincomycin resistant (Lm^r; revertants) were isolated. By the fifth nonselective subculture, 1 clone of the 50 replica-plated clones was Tm^S Lm^S and one clone Tm^S Lm^r, indicating that one mutant and one wild-type revertant had been isolated; by passage six, 8 of the 75 replica-plated clones were Tm^S Lm^S and two

were Tm^s Lm^r. The Δ*ermB* genotype was confirmed by PCR (Fig. 1C), and the products were sequenced (Source Bioscience). One clone (designated M68Δ*ermB*) was checked and found to be sensitive to lincomycin (MIC 32 μg · ml^{-1}, while M68 was resistant at 128 μg · ml^{-1}) and erythromycin (MIC 1 μg · ml^{-1}, while M68 was resistant at 128 μg · ml^{-1}).
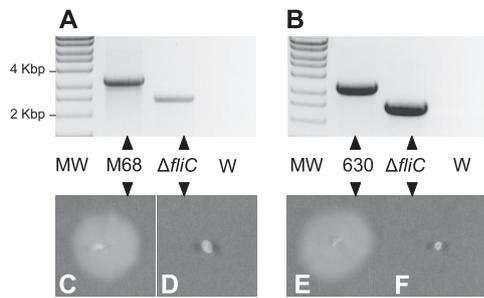
FIG 3 Deletion of the *fliC* gene encoding flagellin from strains M68 and 630 of *C. difficile*. (A) PCR screening spanning the *fliC* gene to confirm its deletion from M68 using primers M68 Δ*fliC* screenF and M68 Δ*fliC* screenR. (B) PCR screening spanning the *fliC* gene in the strain 630 and 630Δ*fliC* to confirm deletions. MW, Hyperladder I molecular weight marker; M68, M68 parental DNA; 630, 630 parental DNA; Δ*fliC*, *fliC* mutant DNA; W, water control. Motility assays in 0.3% agar of the *C. difficile* M68 wild type (C) and the Δ*fliC* mutant (D) showed that M68 is motile but that Δ*fliC* is nonmotile. Motility assays in 0.3% agar confirmed that the 630 strain is motile (E) and the Δ*fliC* mutant is nonmotile (F). Filled triangles indicate the gel lane to which the label relates, and the inverted filled triangles indicate the phenotypic verification by motility assays for these strains.

The ease with which mutant clones were obtained in this experiment suggested that, as predicted, the frequency of the second recombination event is a function of the length of the homology region. To further determine the effect of the length of the homology region on the frequency of the second recombination event, two additional allele exchange plasmids were constructed to make the same deletion in M68 *ermB* with 300-bp and 600-bp regions of homology (pAFP300 and pAFP600, respectively) and conjugated into *C. difficile* M68.

Single-crossover clones derived from pAFP300, pAFP600, and pAFP1200 were confirmed by PCR and serially subcultured without selection daily for 9 days. The empty vector pMTL82151 was used as a control. Each day, 25 colonies were replica plated to determine whether a second recombination event had occurred and, if so, whether this had resulted in reversion to the wild type or deletion of the *ermB* gene. Figure 2A shows the average frequency of Tm^s recombinants over time. We found that double-crossover clones occurred at a considerably higher frequency with 1,200-bp regions of homology than with either 300 bp or 600 bp. Although it was possible to isolate double-crossover clones using 300-bp or 600-bp regions of homology (Fig. 2A), neither was consistent. The second recombination event was not detected in every replicate for either 300-bp or 600-bp regions, and the frequency was altogether too low to make it a reliable approach to mutagenesis (one double crossover with 300 bp and three double crossovers with 600 bp of the 25 clones per subculture for nine subcultures in triplicate). Following a single subculture without selection, the empty plasmid was lost from the control strain.

Using pAFP1200, double-crossover clones tended to be isolated from the third or fourth subculture onward (Fig. 2A and B). The frequencies of the second recombination event differed between the four repeats (Fig. 2B), but on average, the frequency between days 3 and 9 remained between 4% and 13% of the total clones screened; therefore, double-crossover clones were readily isolated by replica plating 25 colonies. Deletion mutants and revertants were isolated at similar frequencies between days 3 and 9 (Fig. 2C).

These results indicate that, using 1,200-bp regions of homology, the frequency of recombination is sufficient to make
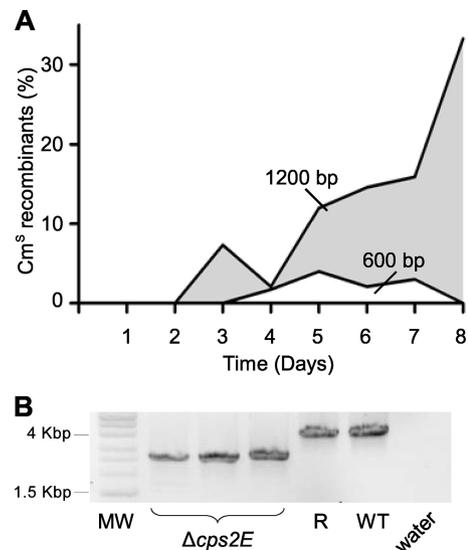

FIG 4 Deletion of the *cps2E* gene from *S. suis* P1/7 and the role of the length of the homology region in the frequency of recombination. (A) The graph represents the fraction of double-crossover recombinants as a percentage of the total clones screened. Single crossovers of two *cps2E* deletion plasmids, one with 600-bp and the other with 1,200-bp regions of homology (pAFP211 and pAFP212, respectively), were passaged without selection and screened for loss of chloramphenicol resistance. The data plotted represent the average results of three replicate experiments. Gray, 1,200-bp, double-crossover recombinants from pAFP212; white, 600-bp, double-crossover recombinants from pAFP211 single-crossover clones. (B) PCR screening spanning the *cps2E* gene in *S. suis* P1/7 to identify deletion mutants from revertants. MW, Hyperladder I molecular weight marker; Δ*cps2E*, *csp2E* deletion mutant DNA; R, revertant to wild-type DNA; WT, wild-type P1/7 DNA; water, water control. Primers *cps2E*screen F and *cps2E*screen R were used and produced bands of predicted sizes, for parental DNA and mutant DNA, of 4,011 bp and 2,667 bp, respectively.

screening alone a reliable method of identifying double-crossover recombinants, a phenomenon that should be generally applicable to any nonessential open reading frame. We applied this principle to the deletion of *fliC* encoding flagellin from *C. difficile*. Allele exchange plasmids for the deletion of *fliC* from M68 (pAFP182) and the clade 1 strain, 630 (pAFP91), were constructed with 1,200-bp regions of homology and single-crossover integrants isolated. Following five nonselective subcultures, mutants were identified among double crossovers by PCR screening (Fig. 3A and B). One mutant of each strain (accordingly designated M68Δ*fliC* and 630Δ*fliC*) was found to be nonmotile in contrast to its parent strain in plate motility assays (25-ml plates of BHIS medium containing 0.3% Bacto agar [Difco] were inoculated with a single colony and incubated for 48 h and the images captured using a Cannon 600D digital camera) (Fig. 3C, D, E, and F).

The usefulness of long regions of homology and serial subculturing to facilitate mutagenesis seems unlikely to be unique to *C. difficile*. *Streptococcus suis* serotype 2 is a major pathogen of swine that has recently been reported to have crossed the species barrier, causing infections in humans (22, 23). The few reports to date of mutagenesis of *S. suis* used allele exchange without a counterselection marker and necessitated the laborious screening of large numbers of colonies (24–26). A major virulence determinant of *S. suis* is its polysaccharide capsule, and the *cps2E* gene is thought to be essential for capsule formation. In order to generate a Δ*cps2E* mutant for further study, the high-frequency approach to allele exchange was applied.

First, we determined that pMTL82151 was nonreplicative in *S. suis* (see the supplemental material). Two allele exchange cassettes for the deletion of *cps2E*, one with 600-bp regions of homology (pAFP211) and the other with 1,200-bp regions of homology (pAFP212), were then constructed. *S. suis* P1/7 was cultured at 37°C in a 5% CO_2 incubator and grown on BHI medium supplemented with chloramphenicol (5 $\mu g \cdot ml^{-1}$) where appropriate. Transformation of the allele exchange plasmids by electroporation was done as previously described (27), and chloramphenicol-resistant single-crossover clones were obtained. Single crossovers were passaged daily without selection, and each day, colonies were replica plated onto chloramphenicol and nonselective agar to determine whether the second recombination event had occurred. As with *C. difficile*, the frequency of the second recombination was considerably higher with 1,200-bp than with 600-bp regions of homology (Fig. 4A), such that mutants could be easily isolated by the fifth and sixth passage without selection. Chloramphenicol-sensitive clones were screened by PCR to determine whether they contained mutant or revertant alleles (Fig. 4B). Deletion mutants of *cps2E* were isolated at the fifth nonselective passage and confirmed by PCR.

High-frequency allele exchange offers a widely applicable alternative to the use of counterselection markers in the generation of precise, markerless mutations. As the development of conterselection markers can be challenging, this high-frequency approach should be particularly useful in organisms for which counterselection markers have not been established (11). In fact, our findings suggest that there may be less need for counterselection markers than previously envisaged.

ACKNOWLEDGMENTS

REFERENCES

1. **Falkow S.** 2004. Molecular Koch's postulates applied to bacterial pathogenicity—a personal recollection 15 years later. Nat. Rev. Microbiol. **2**:67–72.
2. **Falkow S.** 1988. Molecular Koch's postulates applied to microbial pathogenicity. Rev. Infect. Dis. **10**(Suppl 2):S274–S276.
3. **Cartman ST, Kelly ML, Heeg D, Heap JT, Minton NP.** 2012. Precise manipulation of the *Clostridium difficile* chromosome reveals a lack of association between the *tcdC* genotype and toxin production. Appl. Environ. Microbiol. **78**:4683–4690.
4. **Cartman ST, Minton NP.** 2010. A mariner-based transposon system for in vivo random mutagenesis of *Clostridium difficile*. Appl. Environ. Microbiol. **76**:1103–1109.
5. **Heap JT, Ehsaan M, Cooksley CM, Ng YK, Cartman ST, Winzer K, Minton NP.** 2012. Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. Nucleic Acids Res. **40**:e59. doi:10.1093/nar/gkr1321.
6. **Heap JT, Kuehne SA, Ehsaan M, Cartman ST, Cooksley CM, Scott JC, Minton NP.** 2010. The ClosTron: mutagenesis in Clostridium refined and streamlined. J. Microbiol. Methods **80**:49–55.
7. **Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP.** 2007. The ClosTron: a universal gene knock-out system for the genus Clostridium. J. Microbiol. Methods **70**:452–464.
8. **Link AJ, Phillips D, Church GM.** 1997. Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. J. Bacteriol. **179**:6228–6237.
9. **Al-Hinai MA, Fast AG, Papoutsakis ET.** 2012. Novel system for efficient isolation of clostridium double-crossover allelic exchange mutants enabling markerless chromosomal gene deletions and DNA integration. Appl. Environ. Microbiol. **78**:8112–8121.
10. **Gay P, Le Coq D, Steinmetz M, Ferrari E, Hoch JA.** 1983. Cloning structural gene *sacB*, which codes for exoenzyme levansucrase of *Bacillus subtilis*: expression of the gene in *Escherichia coli*. J. Bacteriol. **153**:1424–1431.
11. **Reyrat JM, Pelicic V, Gicquel B, Rappuoli R.** 1998. Counterselectable markers: untapped tools for bacterial genetics and pathogenesis. Infect. Immun. **66**:4011–4017.
12. **Steinmetz M, Le Coq D, Djemia HB, Gay P.** 1983. Genetic analysis of *sacB*, the structural gene of a secreted enzyme, levansucrase of *Bacillus subtilis* Marburg. Mol. Gen. Genet. **191**:138–144.
13. **Griffiths D, Fawley W, Kachrimanidou M, Bowden R, Crook DW, Fung R, Golubchik T, Harding RM, Jeffery KJ, Jolley KA, Kirton R, Peto TE, Rees G, Stoesser N, Vaughan A, Walker AS, Young BC, Wilcox M, Dingle KE.** 2010. Multilocus sequence typing of *Clostridium difficile*. J. Clin. Microbiol. **48**:770–778.
14. **He M, Sebaihia M, Lawley TD, Stabler RA, Dawson LF, Martin MJ, Holt KE, Seth-Smith HM, Quail MA, Rance R, Brooks K, Churcher C, Harris D, Bentley SD, Burrows C, Clark L, Corton C, Murray V, Rose G, Thurston S, van Tonder A, Walker D, Wren BW, Dougan G, Parkhill J.** 2010. Evolutionary dynamics of *Clostridium difficile* over short and long time scales. Proc. Natl. Acad. Sci. U. S. A. **107**:7527–7532.
15. **Knetsch CW, Terveer EM, Lauber C, Gorbalenya AE, Harmanus C, Kuijper EJ, Corver J, van Leeuwen HC.** 2012. Comparative analysis of an expanded *Clostridium difficile* reference strain collection reveals genetic diversity and evolution through six lineages. Infect. Genet. Evol. **12**:1577–1585.
16. **Marsh JW, O'Leary MM, Shutt KA, Sambol SP, Johnson S, Gerding DN, Harrison LH.** 2010. Multilocus variable-number tandem-repeat analysis and multilocus sequence typing reveal genetic relationships among *Clostridium difficile* isolates genotyped by restriction endonuclease analysis. J. Clin. Microbiol. **48**:412–418.
17. **Stabler RA, Dawson LF, Valiente E, Cairns MD, Martin MJ, Donahue EH, Riley TV, Songer JG, Kuijper EJ, Dingle KE, Wren BW.** 2012. Macro and micro diversity of *Clostridium difficile* isolates from diverse sources and geographical locations. PLoS One **7**:e31559. doi:10.1371/journal.pone.0031559.
18. **Drudy D, Harnedy N, Fanning S, O'Mahony R, Kyne L.** 2007. Isolation and characterisation of toxin A-negative, toxin B-positive *Clostridium difficile* in Dublin, Ireland. Clin. Microbiol. Infect. **13**:298–304.
19. **Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR.** 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene **77**:61–68.
20. **Heap JT, Pennington OJ, Cartman ST, Minton NP.** 2009. A modular system for Clostridium shuttle plasmids. J. Microbiol. Methods **78**:79–85.
21. **Williams DR, Young DI, Young M.** 1990. Conjugative plasmid transfer from Escherichia coli to Clostridium acetobutylicum. J. Gen. Microbiol. **136**:819–826.
22. **Feng Y, Shi X, Zhang H, Zhang S, Ma Y, Zheng B, Han H, Lan Q, Tang J, Cheng J, Gao GF, Hu Q.** 2009. Recurrence of human *Streptococcus suis* infections in 2007: three cases of meningitis and implications that heterogeneous *S. suis* 2 circulates in China. Zoonoses Public Health **56**:506–514.
23. **Tang J, Wang C, Feng Y, Yang W, Song H, Chen Z, Yu H, Pan X, Zhou X, Wang H, Wu B, Zhao H, Lin Y, Yue J, Wu Z, He X, Gao F, Khan AH, Wang J, Zhao GP, Wang Y, Wang X, Gao GF.** 2006. Streptococcal toxic shock syndrome caused by *Streptococcus suis* serotype 2. PLoS Med. **3**:e151. doi:10.1371/journal.pmed.0030151.
24. **Allen AG, Bolitho S, Lindsay H, Khan S, Bryant C, Norton P, Ward P, Leigh J, Morgan J, Riches H, Eastty S, Maskell D.** 2001. Generation and characterization of a defined mutant of *Streptococcus suis* lacking suilysin. Infect. Immun. **69**:2732–2735.
25. **Feng Y, Cao M, Shi J, Zhang H, Hu D, Zhu J, Zhang X, Geng M, Zheng F, Pan X, Li X, Hu F, Tang J, Wang C.** 2012. Attenuation of *Streptococcus suis* virulence by the alteration of bacterial surface architecture. Sci. Rep. **2**:710. doi:10.1038/srep00710.
26. **Li M, Wang C, Feng Y, Pan X, Cheng G, Wang J, Ge J, Zheng F, Cao M, Dong Y, Liu D, Lin Y, Du H, Gao GF, Wang X, Hu F, Tang J.** 2008. SalK/SalR, a two-component signal transduction system, is essential for full virulence of highly invasive *Streptococcus suis* serotype 2. PLoS One **3**:e2080. doi:10.1371/journal.pone.0002080.
27. **Takamatsu D, Osaki M, Sekizaki T.** 2001. Construction and characterization of *Streptococcus suis-Escherichia coli* shuttle cloning vectors. Plasmid **45**:101–113.