# Random genome deletion studies of *Corynebacterium glutamicum*

**Yota Tsuge**

*Laboratory of Molecular Microbiology and Genetics,*
*Graduate School of Biological Sciences,*
*Nara Institute of Science and Technology*

01/29/2007

## Contents

**Summary** ·····4

**Introduction** ·····7

**Chapter 1 Isolation of new transposable elements**

1.1 Background ·····11

1.2 Results ·····12

1.2.1 Isolation of new transposable elements ·····12

1.2.2 Analysis of IS*14999* ·····13

1.2.3 Analysis of Tn*14751* ·····16

1.3 Discussion ·····20

**Chapter 2 Construction of random genome deletion method**

2.1 Background ·····26

2.2 Results ·····27

2.2.1 Construction of random genome deletion method ·····27

2.2.2 Analysis of genome deletion strains ·····33

2.3 Discussion ·····35

**Chapter 3 Identification of genes required for cell separation**

3.1 Background ·····37

3.2 Results ·····38

3.3 Discussion ·····43

**Concluding remarks** ·····51

**Materials and Methods** ·····53

**Acknowledgements** ·····69

**References** ·····71

**Figures and Tables** ·····84

**Summary**

In this thesis, I describe a new method for efficient way of analysis of uncharacterized genes in *Corynebacterium glutamicum*. At first, two new transposable elements, IS*14999* and Tn*14751*, were isolated from *C. glutamicum* strains. IS*14999* comprised a 1,149-bp nucleotide sequence with 22-bp imperfect terminal inverted repeats. This putative transposase appears to have partial homology to IS*642*, an IS*630*/Tc*1-mariner* superfamily element, at the C-terminal region in the amino acid sequence. A phylogenetic tree constructed on the basis of amino acid sequences of transposases revealed that this new transposable element was more similar to eukaryotic Tc*1*/*mariner* family elements than to prokaryotic IS*630* family elements. Tn*14751* is a native composite transposon that comprises two functional copies of a corynebacterial IS*31831*-like insertion sequence (IS) element organized as converging terminal inverted repeats. Tn*14751* carries 17.4 kb of *C. glutamicum* chromosomal DNA containing various genes, including genes involved in purine biosynthesis but not genes related to bacterial warfare, such as genes encoding mediators of antibiotic resistance or extracellular toxins. Both isolated elements showed a random transposition tendency, suggesting that they may be useful for genetic engineering of *C. glutamicum*.

At second, a new genome engineering method for *C. glutamicum*, using the insertion sequence and Cre/*loxP* system, was established. The deletion strains, generated using only two vectors, varied not only in their lengths but also in the location of the deletion along the *C. glutamicum* R genome. This method generated 42 *C. glutamicum* mutants (0.4-186 kb). A total of 393.6 kb (11.9% of the *C. glutamicum* R genome), coding for 331 genes, was confirmed to be non-essential under standard laboratory conditions.

*C. glutamicum* is known to have a unique cell division system called snapping division, which form V-shape after cell separation. Through analysis of deletion strains, *cglR1596* gene was identified to be involved in cellar morphology. Its N-terminal region exhibits a 49-amino acid signal peptide for secrete extracellular. At the C-terminal end, there is an NlpC/p60 domain, which is found in cell wall hydrolases. Single disruptant of *cglR1596* elongated about three-fold compared to wild type, and more than two nucleoids were observed within a single cell. β-galactosidase fusion experiments suggested that the *cglR1596* gene is transcribed mainly during the mid to late exponential phases. Double disruptant of *cglR1596* and *cglR2070*, which also have an NlpC/P60 domain at their C terminus, elongated more than the *cglR1596* single disruptant, although a *cglR2070* single mutant exhibited a

cell shape similar to that of the wild type. Therefore, at least CglR1596 and CglR2070 were considered to be involved in cell separation in *C. glutamicum*. By transmission electron microscopy, *C. glutamicum* is revealed to have a two-layer cell wall, which makes two daughter cells unseparated even after septum formation is complete. In *cglR1596* mutant cells, cell septa curving outward were observed, indicating that the two daughter cells stress one another. These results indicate that the snapping division is accomplished by both hydrolysis of the cell wall in junction point of two daughter cells and pressure of the daughter cells.

## Introduction

In 1908, monosodium glutamate (MSG) was discovered as a major ingredient of the seaweed konbu. MSG has a popular flavor and is known today as *umami*. After its discovery, MSG was marketed as a flavor enhancer. At that time, the sources for MSG were wheat, soybean, and other plant protein material from which MSG was extracted after hydrolysis by hydrochloric acid. However, this method could not produce a large amount of MSG. A breakthrough was achieved in 1956: the isolation of an L-glutamate producing bacterium, *Corynebacterium glutamicum*. Isolation of *C. glutamicum* made it possible to produce a large amount of L-glutamate by biosynthesis, rather than by extraction from natural protein. Strains yielding high quantities of L-glutamate and other amino acids such as L-lysine, for example, which have a feedback resistance, were then constructed by chemical mutagenesis. They have been used industrially to produce amino acids for decades (Kinoshita, 1985; de Graaf *et al*., 2001; Hermann, 2003). Today, *C. glutamicum* is one of the most important organisms in bioindustry and is used to produce about two million tones of amino acids per year, of which more than one million tones are accounted for by MSG, used as a flavor enhancer, and more than 0.6 million tons by L-lysine, employed as a feed additive and as a precursor of drugs, cosmetics, and further pharmaceutical

compounds. This market volume is constantly expanding: for example, current demand for L-lysine is increasing by as much as 10% each year (Eggeling and Bott, 2005).

*C. glutamicum* is a Gram-positive bacterium of high GC content. Its hierarchy leading to the genus *Corynebacterium* is (Fig. 1) class *Actinobacteria* — subclass *Actinobacteridae* — order *Actinomycetales* — suborder *Corynebacterineae* — family *Corynebacteriaceae*. The suborder *Corynebacterineae* consists of the families *Corynebacteriaceae* (consisting of the genera *Corynebacterium* and *Turicella*), *Dietziaceae* (consisting of the genus *Dietzia*), *Gordoniaceae* (consisting of the genus *Gordonia*), *Mycobacteriaceae* (consisting of the genus *Mycobacterium*), *Nocardiaceae* (consisting of the genera *Nocardia* and *Rhodococcus*), *Tsukamurellace1ae* (consisting of the genus *Tsukamurella*), and the genera *Williamsia* and *Skermania*, which are characterized by the presence of long-chain α-alkyl, β-hydroxy fatty acids, the so-called mycolic acid, in their cell wall (Stackebrandt *et al*., 1997). Some bacteira belonging to the order *Actinomycetales* are known to have a unique cell division mechanism called snapping division, in which the daughter cells adopt a V-form after cell separation (Starr and Kuhn, 1962; Krulwich and Pate, 1971; Puech *et al*., 2001).

To date, over 500 bacterial genomes have been sequenced. Among *Corynebacterium* species, five strains, *C. glutamicum* R (Yukawa *et al*., 2007), *C. glutamicum* ATCC13032 (Ikeda and Nakagawa 2003; Kalinowski *et al*., 2003), *C. efficiens* (Nishio *et al*., 2003), *C. diphtheria* (Cerdeno-Tarraga *et al*., 2003) and *C. jeikeium* (Tauch *et al*., 2005) have been sequenced. The global demand for amino acids is still increasing in the 21st century (Fig. 2; Hermann, 2003). One reason is the worldwide problem of Bovine Spongiform Encephalopathy (BSE), which forces breeding companies to use a lot of amino acids as feed additives in place of bone meal. Therefore, it is expected that highly productive strains can be constructed using genome information in the post-genomic era. Unfortunately, however, there are still numerous uncharacterized genes in the *C. glutamicum* genome. The best way to analyze uncharacterized genes is to construct function-deficient strains and determine their phenotype. There are two ways to construct gene disruptants: target disruption and transposon mutagenesis. Both techniques have a time-consuming stage, construction of targeting vectors for each disruptant and identification of transposition sites, respectively. To offer another approach for efficient analysis of gene function, a random genome deletion method, based on the random transposition of IS element and DNA the excision reaction of the Cre/*loxP* system, was used in the present work.

In this thesis, the isolation and characterization of new transposable elements to be used as a molecular tool is shown in the first chapter. In the second chapter, a method for constructing random genome deletions for efficient identification of uncharacterized genes is described. In the final chapter, genes required for cell separation, identified through analysis of genome deletion strains are reported.

## Chapter 1 Isolation of new transposable elements

### 1.1 Background

Transposable element is a mobile genetic element that is present in almost all organisms. It consists of two elements, insertion sequence (IS) element and transposon. IS element is the simplest form of transposable element containing only transposase coding sequence that catalyze transposition reaction, and terminal inverted repeats. On the other hand, transposon generally carries resistance genes against antibiotics or extracellular toxin or heavy metals, along with transposase and inverted repeats. Transposable element is a powerful molecular tool for genetic engineering such as construction of single gene disruption library using their random transposition characteristic (Hutchison *et al*., 1999; Goryshin *et al*., 2000).

To date, hundreds of transposable elements have been identified in many bacteria, but not many of them are known in *C. glutamicum*. (Vertès *et al*., 1994; Bonamy *et al*., 1994, 2003; Jager *et al*., 1995). Moreover, only a few transposable elements have been verified to possess transposition activity. I initiated studies aimed at isolation of new functional transposable elements from *C. glutamicum* to use as a molecular tool. In this chapter, two new functional transposable elements in *C. glutamicum* were isolated and characterized.

## 1.2 Results

### 1.2.1 Isolation of new transposable elements

Transposable elements were isolated by positive selection using suicide vector, pMV5 (Vertès *et al*., 1994) carrying the *Bacillus subtilis sacB* gene. Gram-negative bacterial cells become lethal in the presence of *sacB* gene product, levansucrase, in the sucrose medium (Gay *et al*., 1985). It is thought that the fatality is due to accumulation of high molecular weight polysaccharide, levan in cell envelop, that is synthesized by levansucrase (Steinmetz *et al*., 1983). On the other hand, Gram-positive bacteria do not show sucrose sensitivity in the same situation, because of the lack of outer membrane that hinder the diffusion of levan. However, *C. glutamicum* and several other Gram-positive bacteria show sucrose sensitivity in the presence of the levansucrase because they contain a mycolic acid layer in their cell envelope which play a similar role as outer membrane (Jager *et al*., 1992; Pelicic *et al*., 1996). Overnight culture of *C. glutamicum* cells harboring plasmid pMV5 was plated on the minimal medium containing sucrose. From several sucrose-tolerant colonies that grew, *sacB*-disrupted strains by a transposable element were obtained. Restriction analysis of extracted plasmid DNA revealed that *sacB*-containing fragment was altered when a transposable element was transposed. After screening of 104 *C.*

*glutamicum* ATCC strains, two plasmids (named pCRB512 and pCRA730) that have altered size of DNA fragment with *Sma*I-*Xba*I digestion as described in materials and methods, were obtained from *C. glutamicum* ATCC14751 and ATCC14999 strains, respectively.

### 1.2.2 Analysis of IS*14999*

*Characterization of IS14999*

The nucleotide sequence fragment inserted in plasmid pCRB512 was determined. The DNA fragment was comprised of 1,149 bp, with 22 bp inverted repeats (2-bp mismatches) at both ends, which has one potential ORF. The ORF begins with an ATG at position 81 and ends at position 1,115, and consists of 1,035 nucleotides, corresponding to a product of 345 amino acids with a predicted molecular weight of 39.3 kDa. The deduced amino acid sequence of the ORF has partial homology (29%) in the C-terminal region with the transposase of IS*642* in *Bacillus halodurans* C-125, which belongs to the IS*630* family. This new IS element was named IS*14999*. The 5'-TA-3' dinucleotides flanking the element were duplicated upon insertion of IS*14999* into *sacB* as a direct repeat. IS*630* family elements were verified to also duplicate the 5'-TA-3' dinucleotide (Tenzen *et al*., 1990). These facts indicated that

IS*14999* belonged to the IS*630* family. This is believed to be the first report of a IS*630* family transposable element in corynebacteria or mycobacteria.

*Phylogenetic relationship between IS14999 and IS630/Tc1-mariner superfamily elements*

Transposases exhibit a highly conserved triad DDE motif as a catalytic domain at the C-terminus (Mahillon and Chandler, 1998) and this motif has proved to play a crucial role in transposition (Lohe *et al*., 1997). The IS*630* family comprises part of the IS*630*/Tc*1*-*mariner* superfamily along with the eukaryotic Tc*1*/*mariner* family because of overall sequence similarity and a specific TA dinucleotide insertion target (Doak *et al*., 1994; Shao and Tu, 2001). Multiple alignment based on the transposase of IS*14999* and IS*630*/Tc*1*-*mariner* superfamily elements were conducted. The results showed that the transposase of IS*14999* has a DDE motif at its C-terminal region and that flanking amino acids of this motif were partially conserved (Fig. 3). These facts clearly showed that IS*14999* belonged to the IS*630*/Tc*1*-*mariner* superfamily. A phylogenetic tree was generated for 18 IS elements belonging to the IS*630* family and six Tc*1*/*mariner* family transposable elements based on the amino acid sequences of their transposases (Fig. 4). Interestingly, the phylogenetic tree showed that IS1*4999* is

closer to eukaryotic Tc*1*/ *mariner* family elements than to the prokaryotic IS*630* family elements. Moreover, the distance between the last two residues in the DDE catalytic triad of IS*14999* was 38 amino acids, which is a unique distance compared to other IS*630*/Tc*1*-*mariner* superfamily elements (Fig. 3). These facts indicated that IS*14999* might be transposed from a Tc*1*/*mariner* family element and could form a new subfamily of the IS*630*/Tc*1*-*mariner* superfamily.

*Transposition of IS14999 into C. glutamicum R and its target preference*

To assess whether IS*14999* could be used as a molecular tool for genetic engineering, a mutagenesis vector of IS*14999* carrying kanamycin resistance gene was constructed. The resulting vector named pCRB203, which cannot replicate in *C. glutamicum*, was used to mutate the *C. glutamicum* R genome. All colonies grown on a rich medium containing kanamycin tested showed chloramphenicol sensitivity, suggesting that IS*14999* was transposed into the chromosome but not into the pHSG398 vector region containing the chloramphenicol resistance gene. Transposition efficiency of Tn*14999* was 22 c.f.u. per μg DNA, calculated by the number of kanamycin-resistant clones on the selective plate by counting an average of five experiments. Sixty insertion sites were determined and the results showed that

IS*14999* seemed to transpose at random sites in the *C. glutamicum* R genome (Fig. 5). To investigate whether IS*14999* recognized other sequences besides the duplicated 5'-TA-3' dinucleotide, flanking regions of the target sequence were analyzed in detail. The results revealed that IS*14999* always duplicated the 5'-TA-3' dinucleotide, and moreover, it preferentially recognized the eight-base 5'-AGCTAGCT-3' palindrome sequence (Fig. 6).

### 1.2.3 Analysis of Tn*14751*

*Characterization of Tn14751*

The complete nucleotide sequence of both strands of inserted DNA fragment in plasmid pCRA730 was determined. The 20.3-kb DNA fragment comprised identical 1,453-bp inverted DNA fragments at each end and a 17,392-bp fragment between these fragments (Fig. 7A). Computer analysis of the inverted DNA fragments indicated the presence of one potential ORF. This ORF consisted of 1,311 nucleotides corresponding to 436 amino acids with a predicted molecular weight of 49.6-kDa. The deduced amino acid sequence encoded by the ORF showed high sequence similarity (99.5%) to the sequence encoded by the transposase gene of IS*31831* (Vertès *et al*., 1994), which belongs to the ISL*3* family. Each 1,453-bp inverted DNA fragment had

a 24-bp imperfect inverted repeat (5-bp mismatches) at both ends (IR-L and IR-R) (Fig. 7B), and the 3' end of the transposase gene and IR-R had an 11-bp overlap. These data suggest that the 1,453-bp inverted DNA fragments are IS*31831*-like elements and that the 20.3-kb mobile element, designated Tn*14751*, was a composite transposon that comprised two copies of IS*31831*-like at each end. The IS elements at the ends of Tn*14751* were designated IS*14751L* and IS*14751R* (Fig. 7).

To confirm the phylogenetic position of the transposase of IS*14751* (IS*14751L* or IS*14751R*), the amino acid sequence of the IS*14751* transposase was compared to the sequences of known transposases that belong to the ISL*3* family (Fig. 8). The IS*14751* transposase formed a tight cluster with the transposases of IS*1207* from *C. glutamicum* Bl15, IS*31831* from *C. glutamicum* ATCC 31831, IS*Bli3* from *Brevibacterium linens*, IS*Psp2* from *Pseudomonas* sp. strain EST1001 (pEST1226), IS*Bli1* from *B. linens*, IS*13869* from *Brevibacterium lactofermentum* ATCC 13869, and IS*1096* from *Mycobacterium smegmatis* ATCC 607. Except for *Pseudomonas* sp. strain EST1001 harboring plasmid pEST1226, which contains the transposase gene of IS*Psp2*, all these strains containing IS elements are closely related (*Corynebacterium*, *Brevibacterium* and *Mycobacterium* species). A comparison of the inverted repeat (IR-L and IR-R) sequences of eight IS elements that formed a tight cluster as

determined by phylogenetic analysis (Fig. 8) revealed a high level of conserved sequences, especially the first 8 bp at the 5' end (Fig. 9).

The two IS elements bracket a large piece of chromosomal DNA containing the following 13 open reading frames: *purM* (encoding 5-phosphoribosyl-5-aminoimidazole synthase), *purF* (encoding amidophosphoribosyl transferase), ORFs encoding three hypothetical proteins (orf1, orf2, and orf3), *purL* (encoding 5-phosphoribosyl-formyl- glycinamidine synthase II), *purQ* (encoding 5-phosphoribosyl-formylglycinamidine synthase I), ORFs encoding four hypothetical proteins (orf4 to orf7), *dctA* (encoding aerobic C4-dicarboxylate transporter), and an ORF encoding one hypothetical protein (Fig. 7A). All of the genes are present in the same order in the genomes of *C. glutamicum* R (Yukawa *et al*., 2007) and *C. glutamicum* ATCC 13032 (Ikeda and Nakagawa 2003, Kalinowsky *et al*., 2003), but they are not flanked by two insertion sequences in *C. glutamicum* ATCC13032 genome. The gene cluster resembles a similar gene cluster in *Corynebacterium efficiens* with lower sequence similarity than the similarity in *C. glutamicum* strains, except for the *dctA* gene, which is located at a different locus on the chromosome. On the other hand, analysis of the *Corynebacterium diphtheriae* genome showed that several genes in the cluster are absent and that the remaining genes are scattered on

the chromosome. The differences in gene distribution among the strains mentioned above corresponded to the differences in phylogenetic classification determined by 16S rRNA gene analysis of corynebacteria (Nakamura *et al*., 2003).

*Transposition of Tn14751 derivatives into C. glutamicum*

To clarify the transposition efficiency of Tn*14751*, an artificial mini-composite Tn*14751* transposon (mini-Tn*14751*) was constructed. The 17.4-kb corynebacterial chromosome portion was omitted from Tn*14751* and replaced with kanamycin resistance gene in order to avoid a background for transposition efficiency caused by homologous recombination. The resulting plasmid named pCRA732 (Fig. 10A), which did not replicate in *Corynebacterium*, was electroporated into *C. glutamicum* to transpose the mini-Tn*14751* into the chromosome. The transposition efficiency was $1.8 \times 10^2$ mutants per μg of DNA. In order to verify that the derivative of Tn*14751* transposed randomly, genomic Southern hybridization of nine randomly selected mini-Tn*14751* integrants (CGR732-1 to CGR732-9) was conducted (Fig. 10). Three different kinds of probes (fragments I, II, and III) were used in Southern hybridization to determine whether the whole mini-Tn*14751* or either one of the two IS*14751* elements at both ends of Tn*14751* transposed into chromosomal DNA in these

integrants (Fig. 10B). When the left side of a fragment of the kanamycin resistance gene (fragment I) was used as the probe, hybridization signals were detected at different sizes, indicating a variety of insertion mutations (Fig. 10C). Although utilization of the right side of the fragment of the kanamycin resistance gene (fragment II) as the probe also resulted in different sizes of hybridization signals, the hybridization patterns obtained with fragments I and II as probes did not overlap (Fig. 10C and D). Southern hybridization with fragment III (IS*14751*) as the probe revealed two bands in each lane, and the hybridization pattern was a composite of the two hybridization patterns described above (Fig. 10E), suggesting that the whole mini-Tn*14751* was transposed into the *C. glutamicum* chromosomal DNA.

**1.3 Discussion**

In this chapter, two transposable elements, IS*14999* and Tn*14751* were isolated and characterized. IS*14999* is thought to belong to IS*630*/Tc*1-mariner* superfamily and more similar to eukaryotic Tc*1*/*mariner* family than prokaryotic IS*630* family. Both family elements are known to preferentially transpose to and consequently duplicate upon insertion a 5'-TA-3' dinucleotide sequence (Ohtsubo and Sekine, 1996; Mahillon and Chandler, 1998; Plasterk *et al*., 1999). As same as these family

elements, IS*14999* was verified to always duplicate 5'-TA-3' upon insertion. Tc*1*/*mariner* family elements, like IS*630* family elements, have conserved DDE or DDD triad amino acids as an essential part of their catalytic site, and mutations in the triad abolished transposase activity (Lohe *et al*., 1997) and they share a similar signature sequence or motif in the catalytic domain of their respective transposases (Shao and Tu, 2001; Urasaki *et al*., 2002).

Inverted repeats of IS*630* family elements are not as conserved as other IS element families (data not shown). IRs of Tc*1*/*mariner* family elements, each of unique length, show partial conservation only in the first four IR nucleotides (Plasterk, 1996). These characteristics are part of the reason for the broad diversity of IS*630*/Tc*1* superfamily elements beyond the frame of prokaryotes and eukaryotes. The phylogenetic tree of IS*630*/Tc*1*-*mariner* superfamily elements showed that some elements were clustered together based on the distance between the second and third amino acids in their DDE motifs. IS*14999* is positioned among the eukaryotic Tc*1*/*mariner* family elements in spite of its prokaryotic origin. It should be noted that the distance between the second D and third E residues of the transposase of IS*14999* in the DDE motif, the catalytic triad, was invariably 38 residues. The distances between the first two Ds are variable while the distances between the last two residues

in the DDE motif are mostly invariable for a given IS or transposon family (Fig. 3). Most IS*630*/Tc*1-mariner* superfamily elements show distances of 34, 35 or 37 residues between the latter two residues, and form subfamilies depending on these distances (Shao and Tu, 2001). IS*14999* has a unique distance of 38 residues in IS*630*/Tc*1–mariner* superfamily, which suggests to form a new subfamily of the IS*630*/Tc*1–mariner* superfamily.

Analysis of insertion sites of IS*14999* showed that the 5'-TA-3' dinucleotide was always duplicated upon insertion, and moreover, it preferably transposed into an 8 bp (5'AGCTAGCT-3') palindrome sequence in the *C. glutamicum* R genome. In this sequence, the A at position –3 and the T at position +3 are the most conserved (55.7% and 70.5% respectively) (Fig. 4). A few detailed analyses of preferred insertion sites of the other IS*630*/Tc*1-mariner* superfamily elements have been reported. IS*630* has been reported to transpose preferentially to the 5'-CTAG-3' sequence (Tenzen and Ohtsubo, 1991). Tc*1* and Tc*3*, 5'-AKATATGT-3' (K=G or A) or 5'-AYATATRT-3' (Y=C or T; R=G or A) and 5'-ATATATTT-3' respectively, were preferentially recognized (Mori *et al*., 1988; Korswagen *et al*., 1996; Preclin *et al*., 2003). The length of conserved preference sequence is the same between IS*14999*, Tc*1* and Tc*3*. In Tc*1*, the A at position -3 and the T at position +3 are the more highly conserved

(75% and 73% respectively) (Preclin *et al*., 2003). This high conservation of the A at position 23 and the T at position +3 is identical to the situation in IS*14999*. IS*14999* is close to eukaryotic transposable elements not only for the result of phylogenetic analysis, but also for the similarity of their preferred target sequences.

The other isolated transposable element, Tn*14751*, is a composite transposon which carries two copies of IS*31831*-like elements (IS*14751*) organized in inverted repeats flanking an approximately 17.4-kb chromosomal DNA fragment. The GC content of this chromosomal DNA fragment was determined to be 55.3%. This value is in agreement with the GC contents of non-medical strains of the genus *Corynebacterium*, including *C. glutamicum* R (Yukawa *et al*., 2007) and *C. glutamicum* ATCC 13032 (Kalinowski *et al*., 2003; Ikeda and Nakagawa, 2003), whose GC contents were calculated to be 54.1 and 53.8%, respectively. The amino acid sequences deduced from the open reading frames that are present in Tn*14751* exhibit a high degree of similarity with amino acid sequences encoded by genes from *C. glutamicum* R and ATCC 13032. These observations corroborate the view that the DNA fragment carried by transposon Tn*14751* originates from the *Corynebacterium* chromosome and does not result from a horizontal gene transfer event.

To our knowledge, transposons carrying large chromosomal DNA fragments

have not been isolated frequently. Genes involved in microbial warfare, such as antibiotic resistance genes or biosynthetic genes for resistance against extracellular toxins or heavy metals, have been encountered frequently in mobile genetic elements, whereas housekeeping genes, like those of Tn*14751*, have not. However, transposons containing a more limited number of chromosomal genes have been reported, as exemplified by an IS*10* derivative carrying sequences from the *E. coli gal* operon flanked by two IS*10* in a direct repeat structure (Raleigh and Kleckner, 1984). Tn*14751* probably formed by insertion of a copy of an IS*31831*-like element (IS*14751*) in an inverted repeat fashion 17.4-kb from an initial copy of IS*14751*, thus generating a composite transposon. The fact that these two copies of IS*14751* functioned as a composite transposon, not as individual insertion sequences, is interesting. Individual sequencing of these two insertion elements confirmed that they encode a functional transposase and are flanked by functional inverted repeats and that they are thus likely to be capable of individual transposition.

To verify the distribution of both transposable elements in *C. glutamicum* strains, dot blot hybridization of chromosomal DNA was performed by using a transposase sequence of each IS element as a probe. The results show that IS*14999* is a rare IS element in *C. glutamicum* strains, whereas IS*14751* or IS*31831* is abundant (Fig. 11).

The difference reflects the transposition efficiency of both IS element.

Both IS*14999* and Tn*14751* have not only interesting characteristics, but also a capability of a useful tool for genetic engineering, because they are able to transpose randomly into the chromosome. Especially, Tn*14751* is useful in the case of integration of a long DNA fragment into chromosome.

## Chapter 2 Construction of random genome deletion method

### 2.1 Background

At present, almost half of the open reading frames (ORFs) identified in each of the sequenced *Corynebacterium* species is annotated either as hypothetical protein or protein of unknown function. Despite recent advances in high-throughput techniques such as DNA microarray for analysis of bacterial genomes, the process of gene characterization is still not very fast. Moreover, the interrelationships between cellular processes and gene function remain poorly understood.

A popular approach to identify the function of non-essential genes is to construct gene-disruption mutants and analyze the phenotypes of the resultant mutants. This strategy potentially offers many clues to understand the genes' functions. Though this approach has been useful in identifying the roles of various important genes, it is tedious, requiring, for example, the construction of targeting vectors for each disruption mutant. In contrast, an alternative strategy of transposon-insertion mutagenesis can easily generate large mutant pools (Hutchison *et al*. 1999, Goryshin *et al*. 2000), but with the major drawback that it is difficult to identify transposition sites. The other problem is that it is hard to check the influence of multiple-genes disruption. To conquer this problem, a new genomic engineering method was

developed.

In this chapter, I describe another strategy for functional analysis of unknown genes based on the *C. glutamicum* insertion sequence (IS) element, IS*31831* (Vertès *et al*. 1994), and the Cre/*loxP* system (Sternberg *et al*. 1986). The Cre/*loxP* system is comprises of a simple two-component excision system with Cre recombinase and two 34-bp *loxP* sequences. Combining IS*31831* and Cre/*loxP* system, a new approach named random genome deletion method that makes it possible to obtain many strains that were deleted their chromosome randomly was developed. The method is rapid and efficient in the sense that it obviates the construction of all but two vectors. It encompasses a simple and a useful system for easily identifying non-essential regions and new genes of unknown function through comparison and analysis of deletion strains.

## 2.2 Results

### 2.2.1. Construction of random genome deletion method

*Construction of mini-transposon vectors*

In order to delete genomic region with Cre/*loxP* system, two *loxP* sequences integrated in chromosome were required. Although the transposable element that has

almost random transposition tendency, it is not perfectly random. Therefore, two different IS elements were firstly tried to integrate *loxP* sequence into the chromosome. It is important for selection of transposable element that does not exist in parental strain, *C. glutamicum* R, because there is a possibility of a second transposition caused by an endogenous same IS element after genome deletion had occurred.

To date, four functional IS elements; IS*31831*, IS*1206*, IS*1207* and IS*14999* have been identified in *C. glutamicum* (Vertès *et al*. 1994, Bonamy *et al*. 1994, Bonamy *et al*. 2003). The complete genome sequence reveals no other IS-like genes of these IS elements except for IS*1206* in *C. glutamicum* R. Because IS*1207* has a high homology with IS*31831*, IS*14999* and IS*31831* were firstly used to transpose two *loxP* sequences into the chromosome. However, owing to low transposition efficiency of IS*14999* compared to IS*31831*, this approach was unsuccessful. Therefore, IS*31831* was employed as a sole IS element in this method.

IS*31831* is another *C. glutamicum* IS element that is 1,453-bp in length and has 24-bp inverted repeats (5-bp mismatches). This IS element has a very high homology (99.5%) with IS*14751* as described above, and is known to have high transposition efficiency and also a high random transposition tendency (Vertès *et al*. 1994, Suzuki

*et al*. 2006). To excise large segments of *C. glutamicum* R genome, two *loxP* sequences were integrated into the chromosome by IS*31831*, while Cre expression vector was used to insert Cre recombinase into *C. glutamicum* cells. Three kinds of vectors were constructed to delete the genome. The first vector, named pCRB554 (Tn-Km) carried a kanamycin resistance gene, *lacZ* gene and one *loxP* sequence between two inverted repeats of IS*31831* (Fig. 12A). The second vector, named pCRB555 (Tn-Cm) carried a chloramphenicol resistance gene and one *loxP* sequence between two inverted repeats (Fig. 12B). The efficiency of transposition of the two mini-transposon vectors into *C. glutamicum* R chromosome was $2.1\times10^{3}$ (cfu per μg DNA) and $2.9\times10^{2}$ (cfu per μg DNA), respectively. The third vector, named pCRB505, was used to express Cre recombinase in *C. glutamicum* cells.

*Outline of the random deletion method*

A representation of the deletion protocol is shown in Figure 13. First, Tn-Cm was electroporated into the *C. glutamicum* R cells. Positive integrants (*C. glutamicum* R::Tn-Cm) were selected by their chloramphenicol resistance. In the next step, the chloramphenicol resistant integrants were electroporated with Tn-Km and the resultanting mutants (*C. glutamicum* R::Tn-Cm::Tn-Km) selected by their combined

chloramphenicol and kanamycin resistance. Finally, *C. glutamicum* R::Tn-Cm::Tn-Km cells were transformed by pCRB505 Cre expression vector, and positive transformants selected by their spectinomycin resistance. To note here is that the two integrated *loxP* sequences may align in the same direction or in opposite directions; the former case results in an excision event whereas the latter case results in merely *loxP* inversion (Kuhn and Torres, 2002). Deletion strains were identified based on their β-galactosidase activity and sensitivity to antibiotics. Tn-Km integrants were blue on the X-gal plate due to the presence of the *lacZ* gene, whereas deletion strains were yellow due to loss of *lacZ* (Fig. 14A). Deletion strains were also kanamycin- and chloramphenicol- sensitive because of the emission of antibiotic resistance genes from chromosome (Fig. 14B). Successful deletion strains retained only a 243-bp DNA fragment containing inverted repeats and one *loxP* site after this reaction that enable us to determine the deleted region (Fig. 13).

*Transposition features of Tn31831*

Because the same IS element was used in this experiment, the possibility exists of re-transposition of Tn-Cm region upon Tn-Km transposition. Southern hybridization was used to verify not just the random transposition of mini-transposons

but also whether or not this re-transposition occurred. Using 0.7-kb PCR fragments containing either chloramphenicol or kanamycin resistance genes as probe, southern hybridization of each of eight randomly selected mutants was performed. Firstly, random transposition of constructed mini-transposons was investigated. In *C. glutamicum* R::Tn-Cm mutants, eight different bands were detected with a chloramphenicol probe, indicating that each mutant contained a single insertion and that there were no obvious insertion hotspots (Fig. 15A). Similarly, eight different bands were detected with a kanamycin probe in *C. glutamicum* R::Tn-Km mutants, also indicating that each mutant contained a single insertion with no obvious insertion hotspots (Fig. 15B). This random transposition of IS*31831* has been further verified by the experiment of construction of a single-gene disruption library (Suzuki *et al*. 2006). Using the *C. glutamicum* R::Tn-Cm mutant of Figure 15A, lane 8 as parental strain for subsequent Tn-Km transpositions, eight bands of the same size were detected with the parental strain using the chloramphenicol probe, indicating that the transposition region of Tn-Cm did not transpose into different loci in Tn-Km transposition (Fig. 15C). The implication here was that even though the same IS element was used, the first transposition region of the mini-transposon was stable. Using kanamycin probe on the same *C. glutamicum* R::Tn-Cm::Tn-Km strains

similarly gave eight different bands, suggesting random transposition of Tn-Km into *C. glutamicum* R::Tn-Cm strains (data not shown).

*Construction of random deletion strains*

After the two transpositions, Cre recombinase was introduced into mutants using plasmid pCRB505. Transformation efficiency of pCRB505 into *C. glutamicum* was very high ×(more than $10^7$ mutants per μg DNA). Yellow colonies considered to be a deletion strain were obtained about 1.5% of the total number of colonies that grew. Deletion strains after Cre recombinase expression were screened by their color on X-gal-containing plates and their antibiotic sensitivity (Fig. 14), and colony PCR using inverted repeats of IS*31831* as primers (data not shown). Deleted region of 42 strains were identified (see Materials and Methods). PCR using primers designed for locations 100-bp from each terminal inverted repeat were performed to verify genome deletion (Fig. 16, Table 3).

### 2.2.2 Analysis of deletion strains

In total 394-kb DNA (11.9% of total genome) was deleted. 331 genes in the deleted region were revealed to be non-essential under standard laboratory conditions

(Table 6). The deletions ranged from 0.4 to 186-kb (Table 4). After determination of the sequence of the missing regions, Cre recombinase expression vector pCRB505 was cured from each deletion strain. Of the 42 deletion strains, 8 (Random Deletion (RD) 6, 7, 9, 16, 19, 22, 32 and 33) did not grow on minimal medium (Table 4). Comparing the growth rates of the deletion strains to that of the wild-type *C. glutamicum* R on rich and minimal media containing glucose as a sole carbon source revealed six strains (RD15, 16, 26, 27, 38 and 40) with severe growth deficiency (less than 0.5 of wild type) (Table 5). It is unclear at the moment what genes are responsible for this growth defect, but new insight into the regulatory mechanisms of important metabolic pathways could be gleaned from further analysis of these deletion strains.

Two strains, RD22 (50.7-kb deletion strain from *cglR0179* to *cglR0231*) and RD41 (10.1-kb deletion strain from c*glR1595* to *cglR1604*) showed severe morphological change. Both cells elongated compared to wild type in stationary phase and exponential growth phase, respectively (Fig. 18). Complementation experiment revealed that *cglR0197* and *cglR1596* were responsible genes of these phenotypic changes (Fig. 18). Insertional inactivation of both genes into wild type showed identical morphological phenotype as corresponding deletion strains.

## 2.3 Discussion

In this chapter, I describe an efficient method for random deletion that employs IS*31831* in conjunction with the Cre/*loxP* excision system. The merits of this system are that: (i) the function of deleted gene(s) is completely lost. This contrasts with transposon mutagenesis or targeted disruption, where the target gene may still be functional after disruption in some cases. (ii) Collective analysis of a group of genes of unknown function is facilitated, because many genes are deleted in one experiment. (iii) The possibility to identify new genes because many ORFs in the genome, not just those that are already predicted, can be deleted. (iv) Multiple non-essential genes can be identified at once, as the 42 deletion strains representing a total of 394-kb genomic region containing 331 genes deleted in this study show. I further demonstrated here that a 280-kb sub-region of the 394-kb deleted was not part of the *C. glutamicum* R strain specific regions which was identified by comparative genome between strain R and strain ATCC13032 (Suzuki *et al*, 2005a-d), indicating that this method achieved random deletions.

Whereas some IS elements always transpose into the same nucleotide sequence (e.g. IS*630*-family into TA dinucleotides) (chapter 1), others transpose into the same-length nucleotide sequence, that may not necessarily be identical. IS*31831* and

other ISL*3*-family elements are known to transpose into eight-nucleotide sequence as DRs (Mahillon and Chandler, 1998; IS Finder (http://www-is.biotoul.fr/)). Scanning a total of 84 neighboring sequences of Tn-Cm and Tn-Km transposition regions of deletion strains revealed an AT content at 3rd, 4th, 5th and 6th nucleotide of DR were 88%, 100%, 100% and 88%, respectively, indicating that transposase of IS*31831* recognizes AT rich regions in the central domain of DR (Fig. 19). Furthermore, the majority of nucleotide sequences in the DR formed a 5'-GGTTAACC-3' eight-palindrome sequence (Fig. 19), same as the preference sequence of IS*14999* in the first chapter. These observations are consistent with the fact that the mechanism of insertion of IS elements, nicking and cleaving double strand of DNA, is similar to the recognition mechanism of restriction enzyme. It is interesting that IS*31831* transposes globally into the chromosome, but it prefers to transpose into particular AT rich sequence in the central region, though the reason is not clear.

Several large-segment deletion strains (over 100-kb) were obtained within the three-minute-to-eight-minute region of the circular chromosome (Fig. 16). GC-content of this region (52.4%) is not so different from GC-content of the total genome (54.2%). Suzuki *et al.* (Suzuki *et al*. 2006) used IS*31831* to construct a single gene disruption library, showing that IS*31831* transposed randomly throughout the *C.*

*glutamicum* R genome. The fact that the deleted regions in this study appear to be biased toward one particular region is due not to IS*31831* preferentially transposing into this region but simply to a lack of essential genes in this region. As 42 deletion strains were obtained from only two series of experiments, it can be reasonably expected that other deletion strains from different regions could be obtained. Precise and fundamental investigation about *C. glutamicum* will be able to achieve through analysis of strains in the library of genome deletion strains.

## Chapter 3 Identification of genes required for cell separation

### 3.1 Background

In most prokaryotic and eukaryotic species, cell division occurs by the formation of a division septum and the subsequent formation of two equivalent daughter cells. After completion of chromosome replication and segregation of the daughter chromosomes to the two halves of the cell, the septum assembles at a predetermined site and two progeny cells are produced. In most bacteria, cell division is achieved by the simultaneous constriction of both the cytoplasmic membrane, the peptidoglycan layer and any other cell envelope layers, such as the outer membrane of Gram-negative bacteria (Rogers *et al*., 1980; Nanninga, 1998).

In *E. coli* or *B*. *subtilis*, separation of the daughter cells by cleavage of the central part of the septal cell wall can occur together with septation and constriction, although the sequences of these processes are slightly different (Nanninga *et al*., 1979; Lutkenhaus, 1993; Heidrich *et al*., 2001). However, *C. glutamicum* is different that cells have a pair of daughter cells that sealed off from each other by a double membrane but joined by a layer of wall material after septum formation is complete. Cell separation then occurs by the unique action called snapping division (Starr and Kuhn, 1962), results in forming “V” shape (Fig. 20). This unique cell division system

has been observed in the other Actinomycetales species such as *Arthrobacter* or *Mycobacterium* (Krulwich and Pate, 1971; Dahl, 2004). Although this phenomenon was firstly described over a century ago (Kurth, 1898) and morphological observations by phase contrast and electron microscopes have been performed, a detailed mechanism of the division system involved is still obscure.

In the previous chapter, a genome deletion strain RD41 showed morphological change and it was revealed that *cglR1596* was responsible for the phenotype. In this final chapter, a detailed analysis of *cglR1596* was performed and it was revealed that this gene is required for cell separation at the final stage of cell division, together with *cglR2070*. A possible mechanism for snapping division is also discussed.

## 3.2 Results

*Disruption of cglR1596 causes severe morphological change*

Through the analysis of genome deletion strain in the previous chapter, *cglR1596* gene was identified to be involved in morphological change. The *cglR1596* single mutant elongated about three-fold compared to wild type, and several lines, considered to be the cell septa were observed on the cell surface by phase contrast microscope (Fig. 21). By using SYTO16, more than two nucleoids were observed

within a single cell, indicating that this gene is involved in cell separation at the final step of cell division. The line on cell surface lay between nucleoids, indicating cell sputum. CglR1596 has 28% homology with p45 of *Leifsonia xyli*, a peptidoglycan lytic protein. CglR1596 has 611 amino acid residues and is predicted that there is a 49-amino acid signal peptide at the N terminus by SignalP (http://www.cbs.dtu.dk/services/SignalP/) (Nielsen *et al*., 1997); there is also an NlpC/P60 domain at the C terminus (Fig. 22A). This domain is contained in some cell wall hydrolases such as the *lyt* and *cwl* genes in *B. subtilis*, and *p60* in *Listeria monocytogenes* (Yamamoto *et al*., 2003; Machata *et al*., 2005; Fukushima *et al*., 2006).

*Secretion and β-galactosidase assay*

According to SignalP (Nielsen *et al*., 1997) analysis, CglR1596 contains a putative signal peptide, suggesting that it assists the export of the protein from the cell. To test this hypothesis, the signal sequence of *cglR1596* was fused to the α-amylase from *Geobacillus stearothermophilus* on the plasmid Lsv-PtacAmi-EcoRV (Watanabe, unpublished). The resulting vector, L-1596SP-Ami, was transformed into *C. glutamicum* R. Following growth, the transformants were overlaid on a

starch-containing plate, allowing detection of extracellular amylase activity by visualization of a clear halo after addition of iodine solution. A clear halo was visualized around colonies of transformants (Fig. 22B), indicating that the fusion protein was extracellularlly secreted.

To determine the period of expression of *cglR1596*, I constructed a *cglR1596-lacZ* gene fusion vector, L-1596PlacZ was constructed. Figure 22C shows the β-galactosidase activity of *C. glutamicum* R carrying L-1596PlacZ. As shown in Figure 22C, β-galactosidase was constitutively produced, indicating that the *cglR1596* promoter is active throughout the growth. Expression rose gradually during the exponential phase, became maximal at the late exponential phase, and then fell in the late stationary phase. This results indicates that cell separation is mainly occurred in mid to late exponential phase. This hypothesis is consistent with the result of counting the average number of nucleoids of cells between early and late exponential phases (data not shown).

*Double disruption of cglR1596 and cglR2070 resulted in a more severe morphological defect than observed for cglR1596 single disruptant*

*C. glutamicum* R has three genes that have an NlpC/P60 domain at their C terminus as

well as *cglR1596*: *cglR0802*, *cglR2069* and *cglR2070*, which have signal sequence at their N terminus except c*glR0802* (Fig. 24A). However, single disruptants of these genes did not show any morphological phenotype, unlike the *cglR1596* mutant. In *B. subtilis*, mutants having multiple disruptions of cell wall hydrolases possessed longer cell shapes than single disruptants, according to the number of disrupted genes (Ohnishi *et al*., 1999). Therefore, it was expected that multiple disruption of these genes along with *cglR1596* would cause a more elongated cell shape than the *cglR1596* single mutant depending on the number of the hydrolases inactivated, as observed in *B. subtilis*. Indeed, double disruption of *cglR2070* and *cglR1596* produced greater elongation of the cell than the *cglR1596* single mutant (Fig. 24B). CglR2070 showed a secretion activity like CglR1596 (Fig. 22B). The growth rate of *cglR1596* disruptant and *cglR1596* and *cglR2070* double disruptant showed slightly delay compared to wild type (data not shown). These growth defects seemed to be caused by the lag of cell extension and chromosome segregation of both mutant compared to wild type, because the distance of two septa in the mutant was shorter than the distance of two septa by Van-FL analysis (Figs. 23A and 24C). The average number of nucleoids within a single cell was 4.11 for double disruptant and 3.11 for *cglR1596* single disruptant (Fig. 24D). Successive disruption of *cglR0802* and *cglR2069*

together with *cglR1596* and *cglR2070* did not led any to more pronounced morphological difference than observed for the double mutant of *cglR1596* and *cglR2070*. CglR2069 and also showed a secretion activity (data not shown), indicating that *cglR2069* also could function as a cell wall hydrolase, seemingly this mutant did not have an enough hydrolase activity to affect cell morphology, because the mutant was forced to snap.

*Test of CglR1596 for murein hydrolase activity*

To determine whether CglR1596 represents a murein hydrolase of *C. glutamicum*, purified His6×CgR1596 fusion protein was tested in a Zymogram using purified *C. glutamicum* R cell walls as substrate. However, the His6×CgR1596 fusion protein showed no lytic activity (data not shown). To avoid the denaturing conditions of SDS-PAGE, the purified His6×CgR1596 protein was tested for autolytic activity using *C. glutamicum* cell walls in a native assay. Here again, the His6×CgR1596 fusion protein was not able to hydrolyze *C. glutamicum* cell walls (data not shown), indicating that either the absence of murein hydrolase activity by CglR1596 or that CglR1596 requires other factors to show hydrolase activity. Because CglR1596 does not have LysM domain, peptidoglycan-binding domain, in N-terminus, which is

usually found in cell wall hydrolases showing lytic activity by Zymogram assay, it is likely that the absence of hydrolase activity on Zymogram is due to the absence of binding activity to peptidoglycan by CglR1596 itself. In Group B streptococcus, putative cell wall hydrolase PcsB also lacked hydrolase activity under the same conditions, although it is obvious required for the cell wall separation (Reinscheid *et al*., 2001). This gene also does not have LysM domain at N terminus, indicating that there might not a little for cell wall hydrolase that do not have hydrolase activity on its own.

### 3.3. Discussion

In this chapter, two genes, *cglR1596* and cglR2070, were identified as being required for cell separation in *C. glutamicum*. A *cglR1596* single mutant and a double disruptant of *cglR1596* and *cglR2070* showed morphological defects, whereas a *cglR2070* single disruptant did not show any mutant phenotype. Phenotypic complementation of the double disruptant to wild type was observed with plasmids carrying only *cglR1596*. These results imply that CglR1596 and CglR2070 may form a multienzyme complex, and that CglR1596 plays a major role in cell separation supported by CglR2070, although the detailed relationship of these genes is still

unclear.

C. glutamicum *as a MreB-lacking bacterium*

It is also noteworthy that there is no MreB protein, an actin-like cytoskeleton protein found in most bacterial species, in *C. glutamicum* or its relatives. Addition of A22, which inhibits MreB function and induces spherical cells, did not cause any morphological change in *C. glutamicum* (Iwai *et al*., 2002; data not shown). In *B.* subtilis, cell shape is maintained by the combined action of MreB and its helical filament-forming relative Mbl, which are mainly responsible for width control and linear axis control, respectively. Disruption of *mreB* resulted in a round cell shape with loss of cell width control, and the effect was ultimately lethal, whereas the *mbl* disruptant cells were bent and twisted at irregular angles (Jones *et al*., 2001). MreB is also thought to act as a basement for new peptidoglycan synthesis, because the localization pattern of MreB protein and of nascent peptidoglycan synthesis is linked with helix formation (Cabeen and Jacobs-Wagner, 2005). Therefore, the result that *mreB* and *mbl* genes are absent in *C. glutamicum* raises the possibility that another mechanism is responsible for maintaining cell shape and also polar growth of the cell wall.

It might be expected that CglR1596 affects cell shape determination in *C. glutamicum*. By phase contrast and scanning electron microscopes (SEM), it was observed that *cglR1596* single mutant was fatter than wild-type cell (Figs. 21 and 25). To determine whether the cytoplasm became enlarged or the cell wall became thicker, the plasma membrane was stained by FM4-64 and TEM analysis were performed. This showed that the cytoplasm of the mutant became larger than wild type, indicating that *cglR1596* may indeed have a role in determination of the width size, although the mechanism is still not clear (Figs. 23B and 26).

*Possible mechanism of snapping division*

Most *cglR1596* mutant cells had two or three septa within a single cell, and the sizes of each compartment were different. The distance between cell pole and adjoining cell septum was greater than that between one cell septum and the next septum (Fig. 21). However, nascent peptidoglycan visualized by Van-FL is mainly synthesized at cell septa in *cglR1596* mutant cells (Fig. 23A), although wild-type *C. glutamicum* cells synthesize new peptidoglycan at the cell pole at cell extension time (Fig. 23A; Daniel and Errington, 2003; Cabeen and Jacobs-Wagner, 2005). This indicates that the intermediate compartment of a mutant cell, which is sandwiched

between cell septa, cannot elongate well compared to the compartments at both ends. This hypothesis was supported by time-lapse analysis of the *cglR1596* mutant (Fig. 20B).

Actinomycetales species such as *Arthrobacter* or *Mycobacterium*, which have a two-layered cell wall, also show snapping division (Krulwich and Pate, 1971; Dahl, 2004), suggesting that this two-layered cell wall plays an important role in this unique cell division. By transmission electron microscopy (TEM) analysis, it is confirmed here that *C. glutamicum* also has a two-layered cell wall, and that the two daughter cells cross-link with each other by the outer layer of the cell wall during septum formation and even after it is complete (Figs. 26A and B). In other wards, the separation of daughter cells has not yet started, even though septum formation is complete. In *E. coli*, separation of the daughter cells by cleavage of the central part of the septal cell wall can occur at almost the same time as septation and cell separation, while in *B. subtilis*, start of septation and cell separation are different. In *E. coli*, cell separation caused by AmiC and EnvC occurs at almost the same time as septum formation (Heidrich *et al*., 2001; Bernhardt and de Boer, 2004). In *B. subtilis*, there is a time lag between separation and septation; however, cell separation clearly starts before septum formation is complete (Nanninga *et al*., 1979). The mechanism of cell

separation by constriction of cell wall is conserved within these model bacteria. *C. glutamicum*, however, has a pair of daughter cells that are sealed off from each other by a cell wall but joined by a layer of outer wall material after septum formation is complete. Then, when the cell wall at the junction point starts to be hydrolyzed, cell separation suddenly occurs, resulting in formation "V" shape (Fig. 20). *cglR1596* mutant cells do not separate even after the septum starts to bend, showing that CglR1596 is required for cell separation (Fig. 26B).

Three interesting phenomena were observed by TEM analysis. The first is that two septa were formed at the same time (Fig. 26C). Together with the fact that there were no anucleate cells (minicells) in the *cglR1596* mutant, *cglR1596* appers not to be related to the genes involved in cell extension, chromosome segregation and septum formation at the molecular level. The second observation is that after cell separation, one side of the junction was still cross-linked, and but the other side of junction was broken and two scars were visible (Figs. 26A and B). The final observation is that the cell septa were curved outward in some mutant cells (Figs. 26D and E). Considering the latter two observations, together with the finding that *C. glutamicum* cells divide into two daughter cells extremely rapidly, on an agarose pad observed by time-lapse analysis (Fig. 20), it is speculated that the two daughter cells press against each other

with considerable force, and that this force is important for snapping division. Snapping division may therefore be accomplished by hydrolysis of the joining point of the two-layered cell wall requiring at least *cglR1596* and *cglR2070*, and pressure in the daughter cells.

In *C. glutamicum*, a specific factor for determining polar growth and producing stress force in cell elongation may exist. I searched the database and selected candidates meeting the following three requirements: (i) conserved among the four sequenced *C. glutamicum* strain genomes; (ii) no isolated gene disruptants [i.e., candidates for essential genes (Suzuki *et al*., 2006)]; (iii) has a transmembrane-spanning region with coiled-coil domain to interact with other proteins. Among the candidate genes was *cglR2032*. This gene has some homology with *divIVA* in *B. subtilis*, which is involved in selection of cell division sites and is also required for proper segregation of chromosomes into developing spores (Edwards and Errington, 1997; Errington, 2001). The former function is similar to that of *minE* in *E. coli*, which is one of the players of determination systems of Z ring formation site, called the Min system (Errington *et al*., 2003). The Min system, consisting of three genes, *minC*, *minD* and *minE*, forming the min operon, is widely conserved in bacterial species; however, some bacteria, including *C. glutamicum,* do not have

homologs of these genes except for *minD* (cglR2985). Disruption of *cglR2985* produces minicells (data not shown), indicating that there must be a determination system for division-site selection involving *cglR2985*, although its mechanism is still obscure.

In Min system-lacking bacteria, *divIVA* homologs are thought not to primarily affect cell division, but rather have a different function. In *Streptomyces coelicolor*, for example, $DivIVA_{Sc}$ protein localizes at the tip of hyphae and is likely to have a crucial role in hyphal polar growth and cell shape determination, whereas the DivIVA in *B. subtilis* localizes at cell septa and is involved in division site selection (Flärdh, 2003). In *Brevibacterium lactofermentum*, a species very closely related to *C. glutamicum*, $divIVA_{Bl}$ could not be disrupted, the gene product was polarly localized, and overexpression resulted in oval swollen cells (Ramos *et al*., 2003). I therefore considered that CglR2032 protein was likely to be involved in elongation of cell pole in *C. glutamicum* and that overexpression of *cglR2032* would result in the formation of separated cells, like wild type. Wild type cells overexpressing *cglR2032* showed one enlarged cell pole (Fig. 27A). The protein mainly localized at one cell pole, although slight fluorescence was also observed at the other cell pole and at the cell septa (Fig. 27A). On the other hand, overexpression of *cglR2032* in the *cglR1596*

mutant or in the *cglR1596* and *cglR2070* double mutant could not be obtained, although cells of normal cell shape with no fluorescence signal were observed, indicating that overexpression of *cglR2032* in both mutants is lethal (data not shown). I next constructed a strain in which the expression level of *cglR2032* was reduced. It has been reported previously that the promoter of the lactose operon of *E. coli* (P*lac*) is not well recognized by the *C. glutamicum* RNA polymerase (Ramos *et al*., 2005). A strain in which the native *cglR2032* promoter was replaced with the *E. coli lac* promoter by double crossover showed a round cell shape in both wild type and the *cglR1596* mutant, and also showed a short chain form in the *cglR1596* mutant (Figs. 27B and C). This result shows that CglR2032 is required for maintaining the coryneform cell shape. Taken together, these observations suggest that *cglR2032* correlates for cell extension to create pressure in the cells for snapping division. Therefore, snapping division may be accomplished by hydrolysis of the joining point of the two-layered cell wall and require pressure in the daughter cells.

## Concluding remarks

In this thesis, new molecular tools of transposable element were developed, and a genome engineering method for *Corynebacterium glutamicum* was established. Through analysis of genome deletion strains, genes related to the unique cell division system called snapping division were also identified. By electron microscopic analysis, snapping division was thought to be accomplished by hydrolysis of the two-layered cell wall and pressure in the daughter cells.

*C. glutamicum* has been an important industrial microorganism because of its high production of amino acids such as glutamate and lysine. Owing to its high growth rate and resistance to lysis, *C. glutamicum* is expected to be a universal host for production of useful materials, not only other amino acids but also organic acids or ethanol, in the 21st century. There are still many questions to solve for understanding metabolic regulation to create highly productive strains. The random genome deletion method developed in this work should provide many clues for identifying unknown genes related to various cell processes including metabolic regulation.

As well as its uniqueness as a material-producing bacterium, *C. glutamicum* has distinctive characteristics in its cell division and cytoskeleton systems. *C. glutamicum*

is thought to be a good model of a bacterium lacking *mreB*, a widely conserved cytoskeleton protein in bacteria.

Finally, two genes required for cell separation in *C. glutamicum* were identified in this thesis. Technological advances in cellular bacteriology will reveal the detailed molecular mechanism of snapping division, and how the unique characteristics of this bacterium have evolved.

## Materials and Methods

### Bacterial strains, plasmids and culture conditions

Bacterial strains and plasmids used in this thesis are listed in Table 1. *C. glutamicum* strains were grown in minimal or A medium as a rich medium at 33 °C with aeration. Minimal medium contained 2 g $l^{-1}$ urea, 7 g $l^{-1}$ $(NH_4)_2SO_4$, 0.5 g $l^{-1}$ $K_2HPO_4$, 0.5 g $l^{-1}$ $KH_2PO_4$, 0.5 g $l^{-1}$ $MgSO_4 \cdot 7H_2O$, 6 mg $l^{-1}$ $FeSO_4 \cdot 7H_2O$, 6 mg $l^{-1}$ $MnSO_4 \cdot 7H_2O$, 200 μg $l^{-1}$ biotin, and 200 μg $l^{-1}$ thiamine-HCl. A medium contained 2 g $l^{-1}$ yeast extract, and 7 g $l^{-1}$ casamino acids in addition to the components of minimal medium. Four percent (w/v) glucose was added to both mediums as a sole carbon source. *E. coli* strains were grown in Luria-Bertani (LB) medium at 37 or 30 °C with aeration. When necessary, spectinomycin was added to a final concentration of 200 μg $mL^{-1}$, ampicillin to 100 μg $mL^{-1}$, kanamycin to 50 μg $mL^{-1}$ and X-gal to 200 μg $mL^{-1}$. Chloramphenicol was used at 50 μg $mL^{-1}$ in *E. coli* strains, and at 5 μg $mL^{-1}$ in *C. glutamicum* strains.

### DNA techniques and PCR methodology

*E. coli* plasmid DNA was extracted using QIAprep® spin Miniprep Kit (QIAGEN) and *C. glutamicum* genomic DNA was extracted using GenomicPrep™

Cells and Tissue DNA Isolation Kit (GE Healthcare) according to the manufacturer's instructions. PCR reactions were performed using *TAKARA LA Taq*™ DNA polymerase (TaKaRa) or *Pyrobest*® DNA Polymerase (TaKaRa) in GeneAmp PCR System 9700 (Applied Biosystems). PCR products were electrophoresed on 1% agarose gel and recovered by using QIAquick® Gel Extraction Kit (QIAGEN). DNA ligation was performed using ligation kit version 2.1 (TaKaRa).

**Transformation of *C. glutamicum***

All plasmid DNA used in the transformation of *C. glutamicum* was extracted from *E. coli* JM110 (*dam*$^-$ *dcm*$^-$). Plasmid DNA extracted from a *dam*$^+$ *dcm*$^+$ *E. coli* strain cannot efficiently transform *C. glutamicum* because of the presence of a methyl-specific restriction system in *C. glutamicum* (Vertès *et al*., 1993). One microgram of unmethylated plasmid was used to transform *C. glutamicum* cells using a GenePulser II (Bio-Rad). Electroporated cells were added to 1 mL of A medium supplemented with glucose and incubated for 2 h at 33 °C. An appropriate volume of culture was plated on medium containing the appropriate antibiotic to select transformants.

**DNA sequencing and sequence analysis**

Nucleotide sequence determination was performed by the dideoxy chain termination method using ABI PRISM 3100 genetic analyzer (Applied Biosystems). DNA sequence data was analyzed with the GENETYX program (Software Development). Nucleotide sequences were determined on both strands independently. Comparison searches of DNA and deduced protein sequences were performed with IS FINDER (http://www-is.biotoul.fr/is.html) and with the BLAST search program (Altschul *et al*., 1997) provided by the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/blast). Multiple alignment was done using CLUSTAL W version 1.83 (Thompson *et al*., 1994). Based on the amino acid sequence of the transposase, a phylogenetic tree was generated using TreeView version 1.6.0 (Page, 1996).

**Dot blot and Southern hybridization**

One microgram of chromosomal DNA from various *Corynebacterium* strains was denatured at 95°C for 10 min and spotted onto Hybond-N+ nylon membrane (GE Healthcare). After the membranes were baked at 80°C for 30 min, membranes were pre-hybridized for 30 min, hybridized overnight at 60°C, and then washed at high

stringency as described elsewhere (Sambrook and Russell, 2001). Southern hybridization was carried out as followed. Genomic DNA of *C. glutamicum* strains was digested with appropriate restriction enzyme and electrophoresed on 1% agarose gel. After electrophoresis, genomic DNA was transferred to positively charged Hybond-N+ nylon membrane (GE Healthcare) using Probe Tech (Oncor). DNA probes were prepared using Gene Image™ Random Prime Labeling Module (GE Healthcare). For dot blot hybridization and Southern hybridization for Tn*14751* experiment, a 1.5-kb DNA fragment containing either IS*14751L* or IS*14751R* (which were identical) in Tn*14751* (fragment III), amplified by PCR performed with primers 1 and 2, and pCRA730 as the template, was used. The same DNA probe was also adapted for use in the transposition experiments with the *C. glutamicum* chromosome. DNA fragment I (0.4 kb; left part of $Km^r$), amplified with primers 7 and 8 with pUC4K as a template, and DNA fragment II (0.4 kb; right part of $Km^r$), amplified with primers 9 and 10 and with pUC4K as the template, were also utilized for transposition into the *C. glutamicum* chromosome. Hybridization signals were detected with a LAS-1000 image analyzer system (Fuji film).

**Detection of transposition event**

*C. glutamicum* cells harboring plasmid pMV5 were grown overnight in A medium supplemented with spectinomycin. A total of 100 mL of the overnight culture was plated on minimal medium containing sucrose supplemented with spectinomycin. From several sucrose-tolerant colonies that grew after 48 h incubation at 33 °C, *sacB*-disrupted strains carrying a transposable element in their *sacB* genes were obtained. The resistant colonies were cultured in liquid minimal sucrose medium and plasmid DNA was extracted. Then plasmid DNA was used to transform *E. coli*, and plasmid DNA was extracted again. Restriction analysis of extracted plasmid DNA using *Sma*I and *Xba*I revealed that the band *sacB*-containing fragment was altered when a transposable element was transposed into the *sacB* gene.

**Sequencing of IS*14751L* and IS*14751R***

The random sequencing of the 20.3-kb mobile element Tn*14751* described above revealed the presence of two copies of IS*31831*-like elements (IS*14751L* and IS*14751R*) at both ends as inverted repeats. To clarify the difference between the nucleotide sequences of IS*14751L* and IS*14751R*, we individually cloned two DNA fragments, as follows: an 8.3-kb *Hpa*I-*Bgl*II DNA fragment containing IS*14751L* was

subcloned into the *Sma*I and *Bam*HI sites of pUC119, and a 1.9-kb *Hpa*I-*Sca*I DNA fragment containing IS*14751R* was subcloned into the *Sma*I site of pUC119. Both inserts on the plasmids were sequenced by primer walking methods.

**Construction of transposon vector**

To construct transposon vector for IS*14999*, plasmid pCRB512 was digested with *Hpa*I and *Dra*I, and subcloned into the *Hin*dIII site of plasmid pHSG398, which cannot replicate in *C. glutamicum*, to yield plasmid pCRB201. A kanamycin-resistance cassette was amplified using PCR with primers P1 and P2 from template pUC4K DNA. The PCR product was subcloned into *Hin*dIII-digested and bluntended pCRB201 to construct plasmid pCRB203. For construction of transposon vector of Tn*14751* is described as follows. A 20.3-kb *Hpa*I-*Hap*I DNA fragment containing the entire Tn*14751* transposon (Fig. 8) was recovered from pCRA730 and was ligated to the *Sma*I site of pHSG398, resulting in plasmid pCRA731. Inverse PCR was performed by using primers 3 and 4 and pCR731 as the template for amplifying a 5.1-kb DNA fragment containing IS*14751L*, pHSG398, and IS*14751R*. The amplified DNA fragment was digested with *Bgl*II and ligated to a *Bgl*II-digested 1.2-kb kanamycin cassette, which was amplified by PCR by using primers, 5 and 6

and pUC4K as the template. The resultant plasmid, pCRA732 (Fig. 10A) containing a mini-composite Tn*14751* transposon, was used for further study.

**Construction of mini-transposon vectors**

Transposase sequence of IS*31831* was amplified using PCR with primers 157 and 158 from plasmid pCRB504 as a template. The PCR product was subcloned into pCRB506 digested with *Eco*RI and *Sac*I to construct plasmid pCRB507. Chloramphenicol and kanamycin resistance cassettes were amplified with primer 113 and 114 or 121 and 122, respectively, using plasmids pHSG398 and pUC4K as templates, respectively. Both PCR products were subcloned into plasmid plox3 digested with *Eco*RI and *Bam*HI or *Hin*dIII to yield plasmid pCRB535 and pCRB539. *lacZ* gene was amplified with primer LacZF and LacZR and digested with *Xba*I, and then ligated with pCRB535 to construct plasmid pCRB536. Finally, the transposition region was amplified with primer 163 and 164 or 165 and 166 using plasmids pCRB536 and pCRB539 as templates, respectively. Primers 163 and 165 contain left inverted repeats and 164 and 166 contain right inverted repeats at their 5' ends. The PCR products were digested with *Spe*I and ligated with pCRB536 and pCRB539, respectively, and digested with *Xba*I, to construct plasmids pCRB554 (Tn-Km) and

pCRB555 (Tn-Cm) (Fig. 12). Nucleotide sequences of plasmids were verified on both strands independently. *E. coli* cells carrying the plasmid which contained the transposase sequence were grown at 30 °C.

**Determination of insertion sites and deleted regions**

After genomic DNA of recombinants or deletion strains was extracted and digested with appropriate restriction enzyme, it was circularized by self-ligation. The flanking region of insertion sites was amplified by inverse PCR with primers P3 and P4. The PCR product was sequenced to determine insertion sites. To identify deleted regions, each of the deletion strains was digested with appropriate restriction enzymes that left the 243-bp nucleotide sequence that remained after the deletion reaction intact. After self-ligation of the digested genomic DNA, the flanking DNA of the terminal inverted repeats was amplified by inverse PCR with primers RDSF and RDSR. The PCR product was sequenced with the same primers to determine deleted regions of the genome.

**Growth characteristics of deletion strains**

Deletion strains were cultured in A medium or minimal medium containing 4%

(w/v) glucose as a sole carbon source. Overnight growth culture was inoculated into fresh medium to an optical density (OD) 610 nm of 0.1. $OD_{610}$ was measured every hour and used to calculate specific growth rate. We defined the relative growth rate of deletion strains as the per hour growth rate divided by the per hour growth rate of wild-type strain.

**α-Amylase assay on agar plates**

Extracellular production of CglR1596 and CglR2070 proteins was detected by the starch iodine reaction of α-amylase from *Geobacillus stearothermophilus*. Signal sequences of both genes were amplified with PCR using primers 1596SPF-EcoRV and 1596SPF-EcoRV, and 2070SPF-EcoRV and 2070SPR-EcoRV, respectively. PCR fragments were cloned into plasmid Lsv-PtacAmiEcoRV (Watanabe, unpublished), which contains the start codon-less α-amylase gene under the *tac* promoter. The resulting plasmid, named Lsv-1596SP-Ami and Lsv-2070SP-Ami, were transformed into *C. glutamicum* R and transformants were overlaid on an A medium plate containing 4% (w/v) starch for 2 days at 33 °C. After 3 mL of iodine solution (1.3 mM iodine and 40 mM potassium iodide) was dropped on the plate, the presence of a white halo around transformants on a purple background indicates extracellular

α-amylase activity. As a negative control, Lsv-PtacAmiEcoRV, lacking a signal sequence, was used.

**β-Galactosidase assay**

The promoter region of *cglR1596* was amplified by PCR using primers 1596PF-SmaI and 1596PR-SmaI and cloned into plasmid L-lacZ, which carries a promoter less *lacZ* gene (Suda, unpublished). The resulting vector, named L-1596PlacZ, and control vector L-lacZ were transformed into *C. glutamicum* R. Transformants were grown in A medium containing 4% (w/v) glucose as a sole carbon source supplemented with chloramphenicol. Cells were harvested every 90 min and optical density (OD) was measured at 610 nm. Cells were washed with Z buffer (30 mM $Na_2HPO_4$, 20 mM $NaH_2PO_4 \cdot 2H_2O$, 5 mM KCl, 0.5 mM $MgSO_4 \cdot 7H_2O$ and 2.8 μl/ mL β-mercaptoethanol) and then stored at –80 °C. The frozen cells were suspended in Z buffer of appropriate volume to give an OD at 610 nm of 1.5. After the cells were permeabilized by toluene, 4 mg/mL of o-nitrophenyl-β-D-galactopyranoside (ONPG) was added. Degradation of ONPG was measured by DU® 800 (Beckman Coulter) at a wavelength of 420 nm.

**Light and fluorescence microscopy**

Microscopy was performed on an Olympus AX70 microscope equipped with a 100× DIC objective and appropriate filter sets (Chroma Technology) and Photometric Cool Snap™ HQ camera (Nikon). Images were processed with Metamorph 5.0 (Universal Imaging) and Adobe Photoshop 5.0. Time-lapse analysis was performed as described elsewhere with modifications (Matroule *et al*., 2004). *C. glutamicum* cells were placed on a 0.5% agar-padded slide containing A medium with 4% glucose. DAPI (Wako) and SYTO16 (Molecular Probes) were used to stain nucleoids. FM4-64 (Molecular Probes) was used to stain the plasma membrane. Fluorescent vancomycin (Van-FL) (Molecular Probes) was used to stain nascent peptidoglycan. Van-FL was added to growing cultures to a final concentration of 10 μg/ml. The culture was then incubated for 10–20 min to allow absorption of the antibiotic. The cells were then fixed for examination at a later time. For fixation of a sample, cells from the culture were washed in phosphate-buffered saline (PBS), and then suspended in 1.6% (w/v) formaldehyde in PBS and left on ice for one hour. The fixed cells were washed three times with PBS and then mounted on slides. Unlabeled vancomycin (Wako) was mixed with an equal amount of Van-FL. The final vancomycin/Van-FL concentration was again 10 μg/ml.

**Electron microscopy**

For scanning electron microscopy analysis, wild type and *cglR1596* mutant cultures in exponential phase ($OD_{610}$ ≈1.0) were centrifuged and fixed with 2.5% (w/v) glutaraldehyde, washed three times with 50 mM PBS, and then dehydrated sequentially in 60, 80, and 99% ethanol at –30 °C for 30 min, 30 min and overnight, respectively. Samples were then immersed in thiobarbituric acid and were dehydrated using a Hitachi ES-203. Vapor deposition of platinum was performed with a Hitachi E-1010. Samples were examined in a Hitachi S-4700 scanning electron microscope at an accelerating voltage of 1.5 kV. For transmission electron microscopy analysis, 10 mL of wild type and *cglR1596* mutant cells were grown to exponential phase ($OD_{610}$ ≈1.0) and fixed in 2.5% (w/v) glutaraldehyde in 50 mM PBS followed by a 2-hour incubation at 4 °C. After cells were washed with distilled water for 10 min at 4 °C five times, they were fixed with 1% (w/v) osmium tetroxide diluted with distilled water at 4 °C overnight. Cells were then washed with distilled water for 10 min at 4 °C three times, and dehydrated with a graded series of ethanol (25, 60, 80, 99, and 100%) using two hours for the first four steps at 4 °C and 20 min for the last step at room temperature three times. Next, cells were incubated successively in mixtures of ethanol and PIOx with a proportion of two to one and one to two, respectively, at

room temperature for 20 min each, then in mixtures of PIOx and Spurr medium with a proportion of three to one, one to one, and one to three at room temperature for one hour each, and fiannly in Spurr medium for two hours twice and overnight. Polymerization was performed at 60 °C for 2 days. Ultrathin sections were cut with a diamond knife, collected onto copper grids, counterstained with 4% (w/v) uranyl acetate for 5 min, washed with distilled water and air dried. Samples were examined in a Hitachi H-7100 transmission electron microscope at an accelerating voltage of 75 kV.

**Preparation of fusion proteins**

The *cglR1596* and *cglR2070* genes lacking a start codon were amplified by PCR using 1596MLF-SacI and 1596R-KpnI, and 2070MLF-SacI and 2070R-KpnI, respectively. The resulting PCR products were digested with *Sac*I and *Kpn*I, and cloned into *Sac*I/*Kpn*I-digested pCold-I vector (TaKaRa). The resulting plasmids (pCold1596, pCold2070) were sequenced to ensure that no mutations had been introduced during cloning, and transformed into *E. coli* strain BL21/pLysS. Transformants were inoculated into 100-mL cultures of LB medium containing 100 μg $mL^{-1}$ ampicillin, cultivated at 37 °C to an $OD_{610}$ of 0.5, and transferred to 15 °C for

30 min, after which 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the medium. Twenty-four hours after induction, cells were harvested by centrifugation (10,000 *g*, 10 min, 4 °C), and the pellet was frozen at –70 °C. The pellet was resuspended in PBS, sonicated, and the lysate was centrifuged at 10,000 g for 10 min at 4 °C. The His-tagged proteins were recovered using His-trap HP (GE Healthcare) according to the manufacturer's instructions.

**Purification of *C. glutamicum* cell walls**

*C. glutamicum* cell walls were prepared as described elsewhere with modifications (Besra, 1998). *C. glutamicum* R was grown in 1L of A medium containing 4% (w/v) glucose until stationary phase. Cells were harvested by centrifugation (5,000 *g*, 10 min, 4 °C). Wet cells (0.5 g) were resuspended in breaking buffer (2% (w/v) Triton X-100 in PBS)) and sonicated for 5 min at 4 °C (using thirty 5-sec cycles of sonication and cooling). After sonication, the suspension was centrifuged (10,000 *g*, 10 min, 4 °C) and the pellet was again resuspended in breaking buffer. After incubation at room temperature overnight, cells were centrifuged (10,000 *g*, 10 min, 4 °C). The pellet was resuspended in 2% (w/v) sodium dodecyl sulfate (SDS) and incubated at 95 °C for one hour twice. The suspension was centrifuged

(10,000 *g*, 10 min, 4 °C) and the pellet was washed successively with distilled water, 80% (v/v) acetone, and acetone. The final pellet was dried for 30 min in a Speed-Vac and ground in a mortar. The cell wall was resuspended to 2% (w/v) in distilled water containing 0.02% (w/v) sodium azide. The suspension was stored at 4 °C.

**Zymogram assay**

Purified His6-CglR1596 and His6-CglR2070 were subjected to SDS-PAGE with gels containing 0.1% (w/v) purified *C. glutamicum* cell walls as a substrate. SDS-PAGE gels were run at 15 mA on ice. Following electrophoresis, gels were rinsed in distilled water, transferred to 300 mL of renaturation solution (25 mM Tris-HCl, pH 7.2, 1% (v/v) Triton X-100), and incubated at 37 °C for 16 hours with gentle shaking. Gels were rinsed with distilled water, stained with 0.01% (w/v) methylene blue in 0.01% (w/v) KOH for 3 hours, and destained with distilled water. Lysozyme and BSA were used as positive and negative controls, respectively. For native assay to detect hydrolase activity in cell wall preparation, fusion proteins were spotted onto an agar plate containing 0.1% (w/v) purified *C. glutamicum* cell walls, and the plate was incubated at 37 °C for 3 hours. Staining and de-staining procedures were carried out as described above.

**Construction of multiple disruptant of cell wall hydrolases**

A multiple disruptant of cell wall hydrolases was constructed using the mutant *lox* system as described previously (Suzuki *et al*., 2005d). The *cglR2070*, *cglR0802* and *cglR2069* genes were successfully disrupted from strain *cglR1596*::Tn*5*. Each disruptant was screened by the phenotype of spectinomycin resistance. Two homologous regions of each gene hindered by a spectinomycin cassette sandwiched by mutant *lox* sequences, *lox66* and *lox71*, were integrated into the chromosome by a double crossover event, followed by deletion of the spectinomycin cassette by Cre expression (Albert *et al*., 1995).

**Acknowledgements**

I would like to thank Dr. Hideaki Yukawa, Research Institute of Innovative Technology for the Earth (RITE), for respect to my will and providing abundant research expenditure.

I thank Professors Ahkiho Yokota, Hisaji Maki and Hiroshi Takagi, Nara Institute and Science and technology (NAIST), for reviewing my thesis.

I sincerely acknowledge Dr. Masayuki Inui, RITE, for continuous encouragement and many supports.

I am deeply grateful to Dr. Nobuaki Suzuki, RITE, for valuable advises about scientific matter and future plan.

I appreciate to Dr. Megumi Iwano, NAIST, for kindly teaching me techniques for electron microscope.

I also appreciate to Dr. Ian Smith, NAIST, for critical reading of this thesis.

I thank Dr. José A. Gil (Universidad de León), Dr. Keiro Watanabe (RITE) and Dr. Masako Suda (RITE) for kindly providing plasmid pEAG2, Lsv-PtacAmiEcoRV and L-lacZ, respectively.

I thank Dr. Hiroshi Nonaka, RITE, Present Mie University, Dr. Hideo Kawaguchi, RITE, Dr. Shohei Okino, RITE, Dr. Masako Suda, RITE, Dr. Shougo Yamamoto,

RITE, Sho Ishii, NAIST, and all the other members of Microbiology Research Group (MMG) in RITE for general talking; Kana Ninomiya, RITE, for general and scientific talking; Dr. Hidetaka Ogino, RITE, for scientific talking and useful discussions; research assistants of MMG for technical assistance.

Finally, I would like to express my deepest appreciation to my parents for their numerous supports in my entirely life.

## References


Albert, H., Dale, E.C., Lee, E., and Ow, D.W. (1995) Site-specific integration of DNA into wild-type and mutant *lox* sites placed in the plant genome. *Plant J* **7**: 649-659.

Bernhardt, T.G., and de Boer, P.A. (2004) Screening for synthetic lethal mutants in *Escherichia coli* and identification of EnvC (YibP) as a periplasmic septal ring factor with murein hydrolase activity. *Mol Microbiol* **52**: 1255-1269.

Bonamy, C., Labarre, J., Reyes, O., and Leblon, G. (1994) Identification of IS*1206*, a *Corynebacterium glutamicum* IS*3*-related insertion sequence and phylogenetic analysis. *Mol Microbiol* **14**: 571-581.

Bonamy, C., Labarre, J., Cazaubon, L., Jacob, C., Le Bohec, F., Reyes, O., and Leblon, G. (2003) The mobile element IS*1207* of *Brevibacterium lactofermentum* ATCC21086: isolation and use in the construction of Tn*5531*, a versatile transposon for insertional mutagenesis of Corynebacterium glutamicum. *J Biotechnol* **104**: 301-309.

Cabeen, M.T., and Jacobs-Wagner, C. (2005) Bacterial cell shape. *Nat Rev Microbiol* **3**: 601-610.

Cerdeno-Tarraga, A.M., Efstratiou, A., Dover, L.G., Holden, M.T., Pallen, M.,

Bentley, S.D., Besra, G.S., Churcher, C., James, K.D., De Zoysa, A., Chillingworth, T., Cronin, A., Dowd, L., Feltwell, T., Hamlin, N., Holroyd, S., Jagels, K., Moule, S., Quail, M.A., Rabbinowitsch, E., Rutherford, K.M., Thomson, N.R., Unwin, L., Whitehead, S., Barrell, B.G., and Parkhill, J. (2003) The complete genome sequence and analysis of *Corynebacterium diphtheriae* NCTC13129. *Nucleic Acids Res* **31**: 6516-6523.

Dahl, J.L. (2004) Electron microscopy analysis of *Mycobacterium tuberculosis* cell division. *FEMS Microbiol Lett* **240**: 15-20.

Daniel, R.A., and Errington, J. (2003) Control of cell morphogenesis in bacteria: two distinct ways to make a rod-shaped cell. *Cell* **113**: 767-776.

de Graaf, A.A., Eggeling, L., and Sahm, H. (2001) Metabolic engineering for L-lysine production by *Corynebacterium glutamicum*. *Adv Biochem Eng Biotechnol* **73**: 9-29.

Doak, T.G., Doerder, F.P., Jahn, C.L., and Herrick, G. (1994) A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common "D35E" motif. *Proc Natl Acad Sci U S A* **91**: 942-946.

Edwards, D.H., and Errington, J. (1997) The Bacillus subtilis DivIVA protein targets to the division septum and controls the site specificity of cell division. *Mol*

*Microbiol* 24: 905-915.

Eggeling, L., and Bott, M. (ed.). 2005. *Handbook of Corynebacterium glutamicum*. CRC Press, Boca Raton, Fla.

Errington, J. (2001) Septation and chromosome segregation during sporulation in *Bacillus subtilis*. *Curr Opin Microbiol* **4**: 660-666.

Errington, J., Daniel, R.A., and Scheffers, D.J. (2003) Cytokinesis in bacteria. *Microbiol Mol Biol Rev* **67**: 52-65.

Flärdh, K. (2003) Essential role of DivIVA in polar growth and morphogenesis in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **49**: 1523-1536.

Fukushima, T., Afkham, A., Kurosawa, S., Tanabe, T., Yamamoto, H., and Sekiguchi, J. (2006) A new D,L-endopeptidase gene product, YojL (renamed CwlS), plays a role in cell separation with LytE and LytF in *Bacillus subtilis*. *J Bacteriol* **188**: 5541-5550.

Gay, P., Le Coq, D., Steinmetz, M., Berkelman, T., and Kado, C.I. (1985) Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. *J Bacteriol* **164**: 918-921.

Goryshin, I.Y., Jendrisak, J., Hoffman, L.M., Meis, R., and Reznikoff, W.S. (2000) Insertional transposon mutagenesis by electroporation of released Tn*5* transposition complexes. *Nat Biotechnol* **18**: 97-100.

Heidrich, C., Templin, M.F., Ursinus, A., Merdanovic, M., Berger, J., Schwarz, H., de Pedro, M.A., and Holtje, J.V. (2001) Involvement of N-acetylmuramyl-L-alanine amidases in cell separation and antibiotic-induced autolysis of *Escherichia coli*. *Mol Microbiol* **41**: 167-178.

Hermann, T. (2003) Industrial production of amino acids by coryneform bacteria. *J Biotechnol* **104**: 155-172.

Hutchison, C.A., Peterson, S.N., Gill, S.R., Cline, R.T., White, O., Fraser, C.M., Smith, H.O., and Venter, J.C. (1999) Global transposon mutagenesis and a minimal *Mycoplasma* genome. *Science* **286**: 2165-2169.

Ikeda, M., and Nakagawa, S. (2003) The *Corynebacterium glutamicum* genome: features and impacts on biotechnological processes. *Appl Microbiol Biotechnol* **62**: 99-109.

Iwai, N., Nagai, K., and Wachi, M. (2002) Novel S-benzylisothiourea compound that induces spherical cells in *Escherichia coli* probably by acting on a rod-shape-determining protein(s) other than penicillin-binding protein 2.

*Biosci Biotechnol Biochem* **66**: 2658-2662.

Jager, W., Schafer, A., Puhler, A., Labes, G., and Wohlleben, W. (1992) Expression of the *Bacillus subtilis sacB* gene leads to sucrose sensitivity in the gram-positive bacterium *Corynebacterium glutamicum* but not in *Streptomyces lividans*. *J Bacteriol* **174**: 5462-5465.

Jager, W., Schafer, A., Kalinowski, J., and Puhler, A. (1995) Isolation of insertion elements from gram-positive *Brevibacterium*, *Corynebacterium* and *Rhodococcus* strains using the *Bacillus subtilis sacB* gene as a positive selection marker. *FEMS Microbiol Lett* **126**: 1-6.

Jones, L.J., Carballido-Lopez, R., and Errington, J. (2001) Control of cell shape in bacteria: helical, actin-like filaments in *Bacillus subtilis*. *Cell* **104**: 913-922.

Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkovski, A., Dusch, N., Eggeling, L., Eikmanns, B.J., Gaigalat, L., Goesmann, A., Hartmann, M., Huthmacher, K., Kramer, R., Linke, B., McHardy, A.C., Meyer, F., Mockel, B., Pfefferle, W., Puhler, A., Rey, D.A., Ruckert, C., Rupp, O., Sahm, H., Wendisch, V.F., Wiegrabe, I., and Tauch, A. (2003) The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J*

*Biotechnol* **104**: 5-25.

Kinoshita, S. (1985) *Glutamic acid bacteria* In: Demain AL, Solomon NA (eds) Biology of Industrial Microorganisms Cummings, London, 115-146.

Korswagen, H.C., Durbin, R.M., Smits, M.T., and Plasterk, R.H. (1996) Transposon Tc*1*-derived, sequence-tagged sites in *Caenorhabditis elegans* as markers for gene mapping. *Proc Natl Acad Sci U S A* **93**: 14680-14685.

Krulwich, T.A., and Pate, J.L. (1971) Ultrastructural explanation for snapping postfission movements in *Arthrobacter crystallopoietes*. *J Bacteriol* **105**: 408-412.

Kuhn, R., and Torres, R.M. (2002) Cre/*loxP* recombination system and gene targeting. *Methods Mol Biol* **180**: 175-204.

Kurth, H. (1898) Über die Diagnose des Diphtheriebacillus unter Berücksichtigung abweichender Culturformen desselben. *Z. Hyg. Infektinonskr. Med. Mikrobiol. Immunol. Virol.* **28**: 409-439.

Lohe, A.R., De Aguiar, D., and Hartl, D.L. (1997) Mutations in the mariner transposase: the D,D(35)E consensus sequence is nonfunctional. *Proc Natl Acad Sci U S A* **94**: 1293-1297.

Lutkenhaus, J. (1993) FtsZ ring in bacterial cytokinesis. *Mol Microbiol* **9**: 403-409.

Machata, S., Hain, T., Rohde, M., and Chakraborty, T. (2005) Simultaneous deficiency of both MurA and p60 proteins generates a rough phenotype in Listeria monocytogenes. *J Bacteriol* **187**: 8385-8394.

Mahillon, J., and Chandler, M. (1998) Insertion sequences. *Microbiol Mol Biol Rev* **62**: 725-774.

Matroule, J.Y., Lam, H., Burnette, D.T., and Jacobs-Wagner, C. (2004) Cytokinesis monitoring during development; rapid pole-to-pole shuttling of a signaling protein by localized kinase and phosphatase in *Caulobacter*. *Cell* **118**: 579-590.

Mori, I., Benian, G.M., Moerman, D.G., and Waterston, R.H. (1988) Transposable element Tc*1* of *Caenorhabditis elegans* recognizes specific target sequences for integration. *Proc Natl Acad Sci U S A* **85**: 861-864.

Nakamura, Y., Nishio, Y., Ikeo, K., and Gojobori, T. (2003) The genome stability in *Corynebacterium* species due to lack of the recombinational repair system. *Gene* **317**: 149-155.

Nanninga, N., Koppes, L.J., and de Vries-Tijssen, F.C. (1979) The cell cycle of *Bacillus subtilis* as studied by electron microscopy. *Arch Microbiol* **123**: 173-181.

Nanninga, N. (1998) Morphogenesis of *Escherichia coli*. *Microbiol Mol Biol Rev* **62**: 110-129.

Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Int J Neural Syst* **8**: 581-599.

Nishio, Y., Nakamura, Y., Kawarabayasi, Y., Usuda, Y., Kimura, E., Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikeo, K., and Gojobori, T. (2003) Comparative complete genome sequence analysis of the amino acid replacements responsible for the thermostability of *Corynebacterium efficiens*. *Genome Res* **13**: 1572-1579.

Ohtsubo, F., and Sekine, Y. (1996) Bacterial insertion sequences. *Curr Top Microbiol Immunol* **204**: 1-26.

Page, R.D. (1996) TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* **12**: 357-358.

Pelicic, V., Reyrat, J.M., and Gicquel, B. (1996) Expression of the *Bacillus subtilis sacB* gene confers sucrose sensitivity on mycobacteria. *J Bacteriol* **178**: 1197-1199.

Plasterk, R.H. (1996) The Tc1/mariner transposon family. *Curr Top Microbiol*

*Immunol* **204**: 125-143.

Plasterk, R.H., Izsvak, Z., and Ivics, Z. (1999) Resident aliens: the Tc*1*/*mariner* superfamily of transposable elements. *Trends Genet* **15**: 326-332.

Preclin, V., Martin, E., and Segalat, L. (2003) Target sequences of Tc*1*, Tc*3* and Tc*5* transposons of *Caenorhabditis elegans*. *Genet Res* **82**: 85-88.

Puech, V., Chami, M., Lemassu, A., Laneelle, M.A., Schiffler, B., Gounon, P., Bayan, N., Benz, R., and Daffe, M. (2001) Structure of the cell envelope of corynebacteria: importance of the non-covalently bound lipids in the formation of the cell wall permeability barrier and fracture plane. *Microbiology* **147**: 1365-1382.

Raleigh, E.A., and Kleckner, N. (1984) Multiple IS*10* rearrangements in *Escherichia coli*. *J Mol Biol* **173**: 437-461.

Ramos, A., Honrubia, M.P., Valbuena, N., Vaquera, J., Mateos, L.M., and Gil, J.A. (2003) Involvement of DivIVA in the morphology of the rod-shaped actinomycete *Brevibacterium lactofermentum*. *Microbiology* **149**: 3531-3542.

Ramos, A., Letek, M., Campelo, A.B., Vaquera, J., Mateos, L.M., and Gil, J.A. (2005) Altered morphology produced by *ftsZ* expression in *Corynebacterium glutamicum* ATCC 13869. *Microbiology* **151**: 2563-2572.

Reinscheid, D.J., Gottschalk, B., Schubert, A., Eikmanns, B.J., and Chhatwal, G.S. (2001) Identification and molecular analysis of PcsB, a protein required for cell wall separation of group B Streptococcus. *J Bacteriol* **183**: 1175-1183.

Rogers, H.J., Perkins, H.R., and Ward, J.B. (1980) *Microbial Cell Walls and Membranes*. London: Chapman & Hall.

Saitou, N., and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**: 406-425.

Sambrook, J., and Russell, D.W.W. (2001) *Molecular Cloning: a Laboratory Manual,* 2nd edn Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Stackebrandt, E., Rainey, F.A., and Ward-Rainey, N.L. (1997) Proposal for a new hierarchal classification system, *Actinobacteria* classisnov. *Int J Syst Bacteriol* **47**: 479-491.

Starr, M.P., and Kuhn, D.A. (1962) On the origin of V-forms in Arthrobacter atrocyaneus. *Arch Mikrobiol* **42**: 289-298.

Shao, H., and Tu, Z. (2001) Expanding the diversity of the IS*630*-Tc*1*-*mariner* superfamily: discovery of a unique DD37E transposon and reclassification of the DD37D and DD39D transposons. *Genetics* **159**: 1103-1115.

Steinmetz, M., Le Coq, D., Djemia, H.B., and 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.

Sternberg, N., Sauer, B., Hoess, R., and Abremski, K. (1986) Bacteriophage P1 cre gene and its regulatory region. Evidence for multiple promoters and for regulation by DNA methylation. *J Mol Biol* **187**: 197-212.

Suzuki, N., Tsuge, Y., Inui, M., and Yukawa, H. (2005a) Cre/*loxP*-mediated deletion system for large genome rearrangements in *Corynebacterium glutamicum*. *Appl Microbiol Biotechnol* **67**: 225-233.

Suzuki, N., Nonaka, H., Tsuge, Y., Okayama, S., Inui, M., and Yukawa, H. (2005b) Multiple large segment deletion method for *Corynebacterium glutamicum*. *Appl Microbiol Biotechnol* **69**: 151-161.

Suzuki, N., Okayama, S., Nonaka, H., Tsuge, Y., Inui, M., and Yukawa, H. (2005c) Large-scale engineering of the *Corynebacterium glutamicum* genome. *Appl Environ Microbiol* **71**: 3369-3372.

Suzuki, N., Nonaka, H., Tsuge, Y., Inui, M., and Yukawa, H. (2005d) New multiple-deletion method for the *Corynebacterium glutamicum* genome, using a mutant *lox* sequence. *Appl Environ Microbiol* **71**: 8472-8480.

Suzuki, N., Okai, N., Nonaka, H., Tsuge, Y., Inui, M., and Yukawa, H. (2006)

High-throughput transposon mutagenesis of *Corynebacterium glutamicum* and construction of a single-gene disruptant mutant library. *Appl Environ Microbiol* **72**: 3750-3755.

Tauch, A., Kaiser, O., Hain, T., Goesmann, A., Weisshaar, B., Albersmeier, A., Bekel, T., Bischoff, N., Brune, I., Chakraborty, T., Kalinowski, J., Meyer, F., Rupp, O., Schneiker, S., Viehoever, P., and Puhler, A. (2005) Complete genome sequence and analysis of the multiresistant nosocomial pathogen *Corynebacterium jeikeium* K411, a lipid-requiring bacterium of the human skin flora. *J Bacteriol* **187**: 4671-4682.

Taylor, L.A., and Rose, R.E. (1988) A correction in the nucleotide sequence of the Tn*903* kanamycin resistance determinant in pUC4K. *Nucleic Acids Res* **16**: 358.

Tenzen, T., Matsutani, S., and Ohtsubo, E. (1990) Site-specific transposition of insertion sequence IS*630*. *J Bacteriol* **172**: 3830-3836.

Urasaki, A., Sekine, Y., and Ohtsubo, E. (2002) Transposition of cyanobacterium insertion element ISY*100* in *Escherichia coli*. *J Bacteriol* **184**: 5104-5112.

Vertès, A.A., Inui, M., Kobayashi, M., Kurusu, Y., and Yukawa, H. (1993) Presence of *mrr*- and *mcr*-like restriction systems in coryneform bacteria. *Res Microbiol*

**144**: 181-185.

Vertès, A.A., Inui, M., Kobayashi, M., Kurusu, Y., and Yukawa, H. (1994) Isolation and characterization of IS*31831*, a transposable element from *Corynebacterium glutamicum*. *Mol Microbiol* **11**: 739-746.

Yamamoto, H., Kurosawa, S., and Sekiguchi, J. (2003) Localization of the vegetative cell wall hydrolases LytC, LytE, and LytF on the *Bacillus subtilis* cell surface and stability of these enzymes to cell wall-bound or extracellular proteases. *J Bacteriol* **185**: 6666-6677.

Yukawa, H., Omumasaba, C.A., Nonaka, H., Kos, P., Okai, N., Suzuki, N., Suda, M., Tsuge, Y., Watanabe, J., Ikeda, Y., Vertes, A.A., and Inui, M. (2007) Comparative analysis of the *Corynebacterium glutamicum* Group and complete genome sequence of strain R. *Microbiology* (in press).

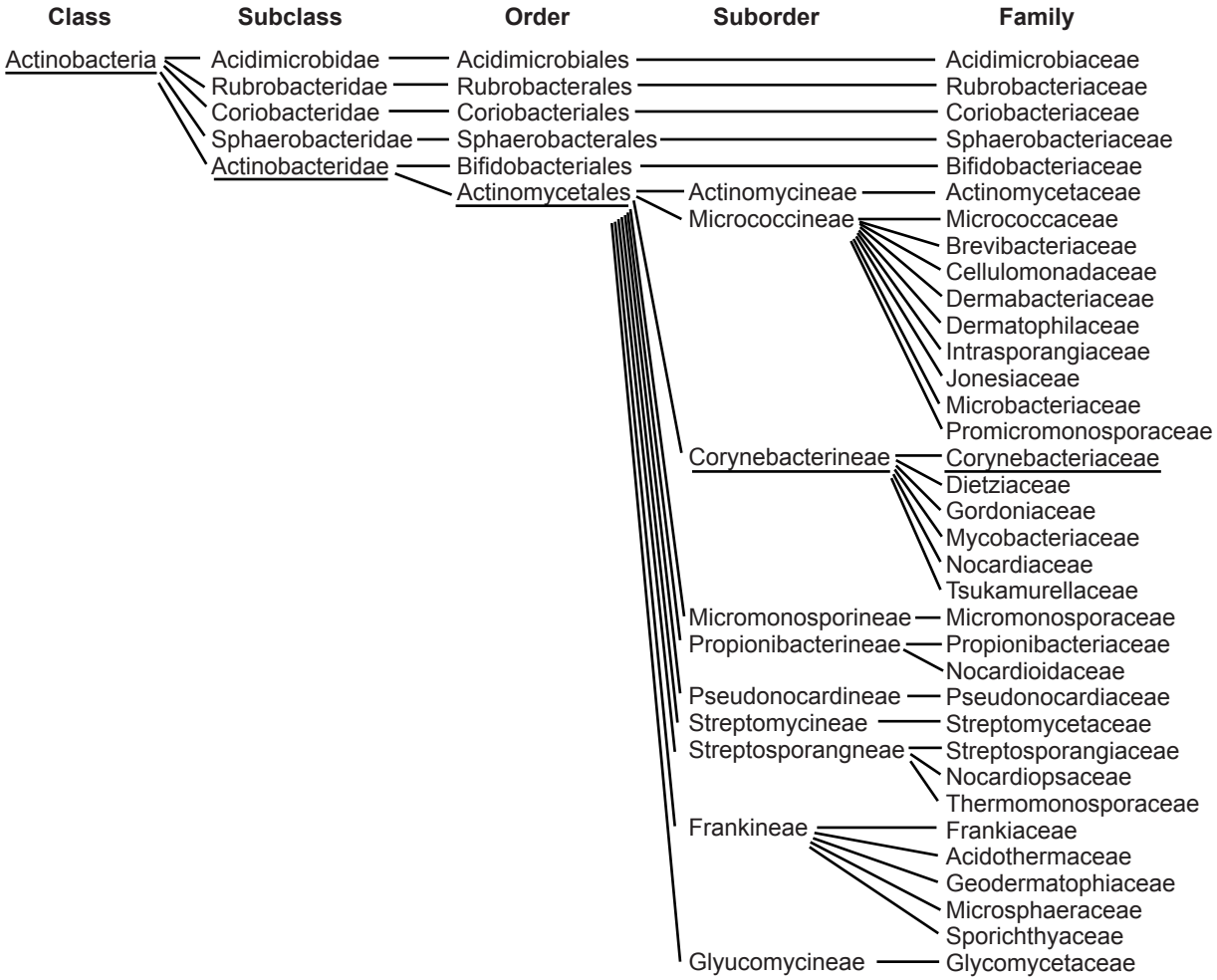



**Fig. 1.** Hierarchic classification system of the class Actinobacteria according to Stackebrandt *et al*., 1997.

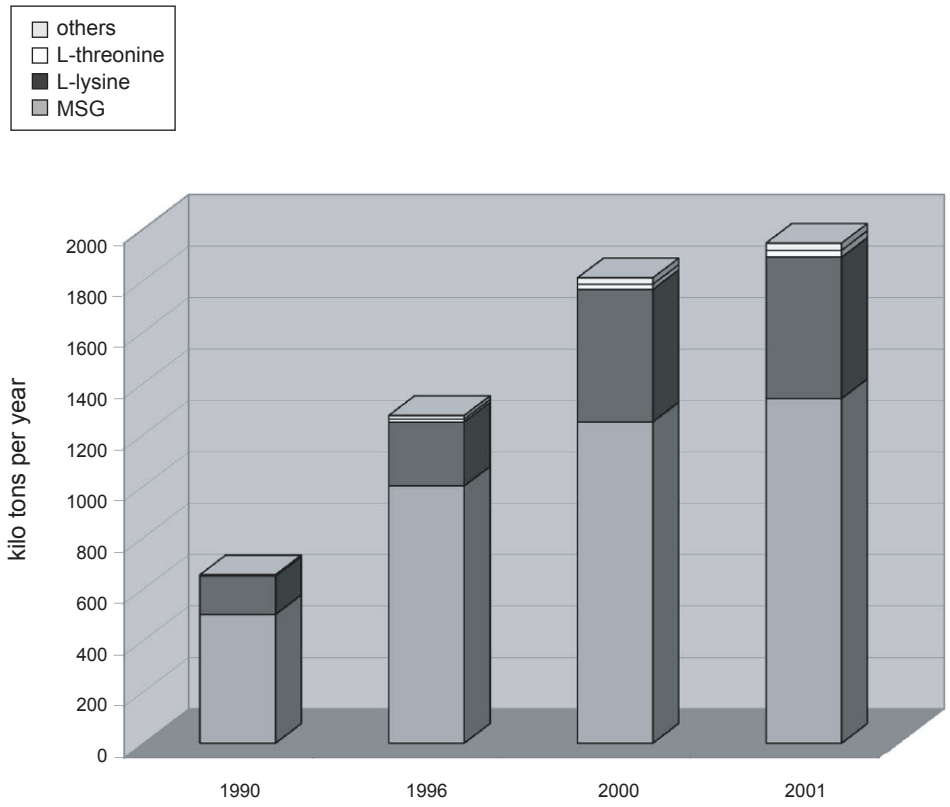



**Fig. 2.** Market development of biotechnological produced amino acids according to Hermann, 2003.

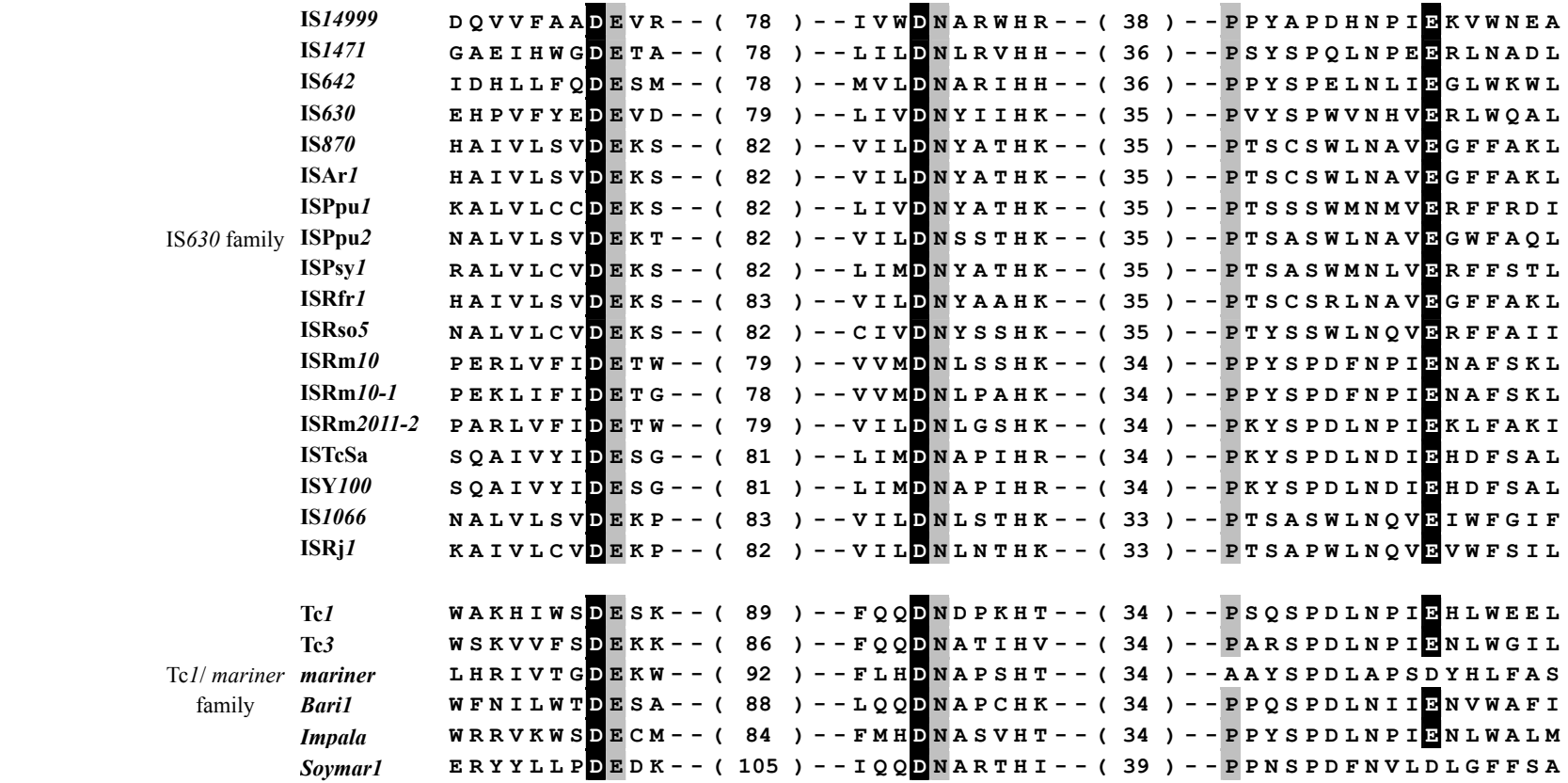



**Fig. 3.** An alignment of segments of transposases encoded by IS*630*/Tc*1-mariner* superfamily elements. The D, D and E of the DDE motif are shown with black backgrounds. Grey backgrounds indicate identical amino acids. The numbers in parentheses show the distance between amino acid sequences of the DDE motif.

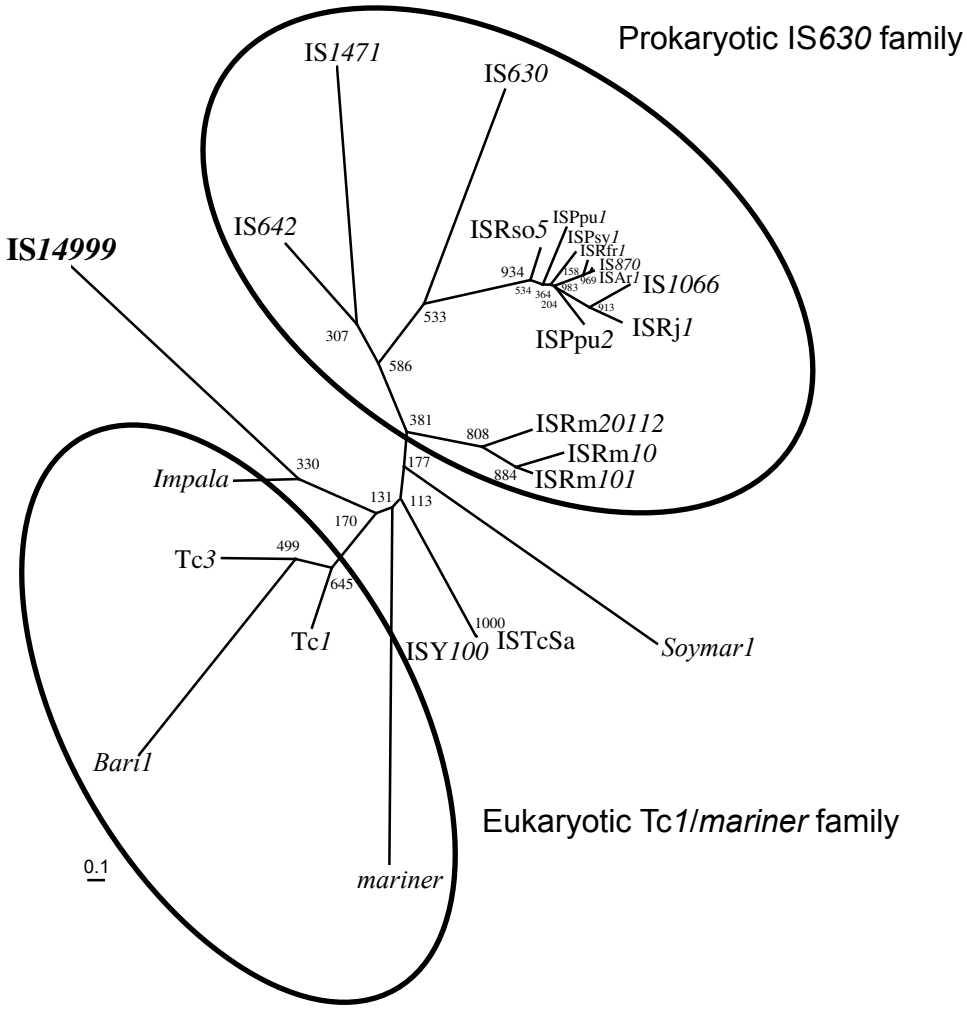



**Fig. 4.** Phylogenetic tree analysis of IS*630*/Tc*1-mariner* superfamily elements. The tree was constructed by the neighbour-joining method (Saitou and Nei, 1987) based on amino acid sequences of transposases. The numbers indicate bootstrap values for 1000 replicates. The scale bar equals a distance of 0.1.The accession numbers or source of information for the sequences used are as follows: IS*14999*, AB186419; IS*1471*, U67938; IS*642*, AP001515; IS*630*, X05955; IS*870*, Z18270; ISAr*1*, K03313; ISPpu*1*, AJ245436; ISPpu*2*, AJ233397; ISPsy*1*, AF169828; ISRfr*1*, M73698; ISRso*5*, CAD16890; ISRm*10*, AF143444; ISRm*10-1*, AJ242573; ISRm*2011-2*, U22370; ISTcSa, U38915; ISY*100*, D90899; IS*1066*, M61114; ISRj*1*, X02581; Tc*1*, X01005; Tc*3*, AF025458; *Impala*, AF282722; *Bari1*, S33560; *mariner*, X78906; *Soymar1*, AF078934.

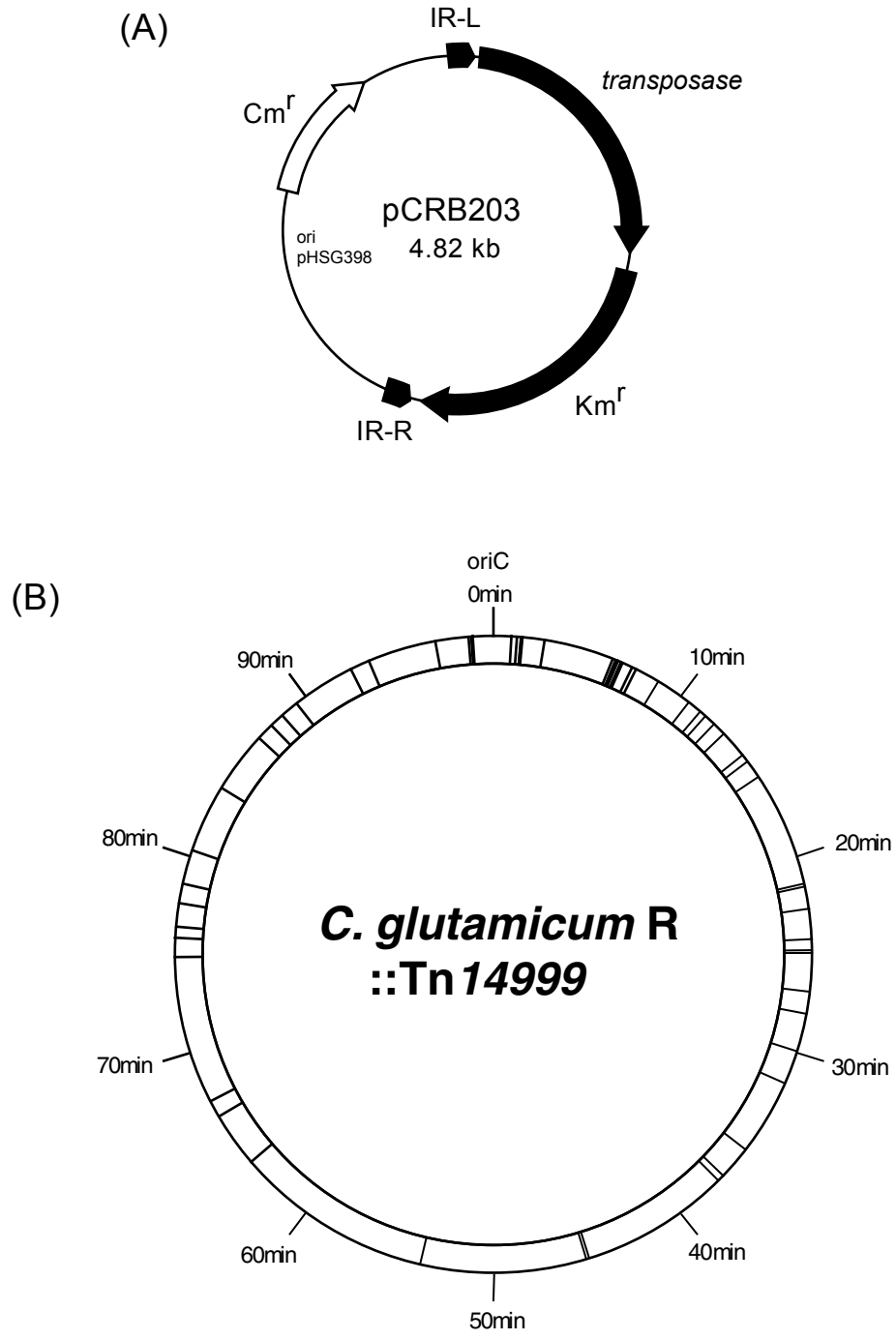



**Fig. 5.** (A) Physical and genetic map of pCRB203. $Km^r$, kanamycin resistance gene; $Cm^r$, chloramphenicol resistance gene; pHSG398 ori, origin of replication from pHSG398.The arrows indicate the direction of transcription. (B) Physical map of insertion sites of Tn*14999* in the *C. glutamicum* R genome.

| | -5 | -4 | -3 | -2 | -1 | Direct Repeat | | +1 | +2 | +3 | +4 | +5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 23 | 25 | 56 | 39 | 11 | 0 | 100 | 8 | 7 | 21 | 30 | 25 |
| T | 20 | 33 | 33 | 3 | 31 | 100 | 0 | 16 | 36 | 70 | 18 | 16 |
| G | 36 | 21 | 5 | 48 | 11 | 0 | 0 | 56 | 2 | 2 | 26 | 30 |
| C | 21 | 21 | 7 | 10 | 46 | 0 | 0 | 20 | 56 | 7 | 26 | 30 |
| | | | **A** | **G** | **C** | **T** | **A** | **G** | **C** | **T** | | |

**Fig. 6.** Target preference of IS*14999*. Numbers represent the percentage occurrence of the preferred base at the positions indicated. The analysis is based on 61 independent insertions (including insertion into *sacB*). Black backgrounds show duplicated bases. Grey backgrounds indicate preferred bases. Bases below the numbers show the most preferred base at positions –3 to +3.

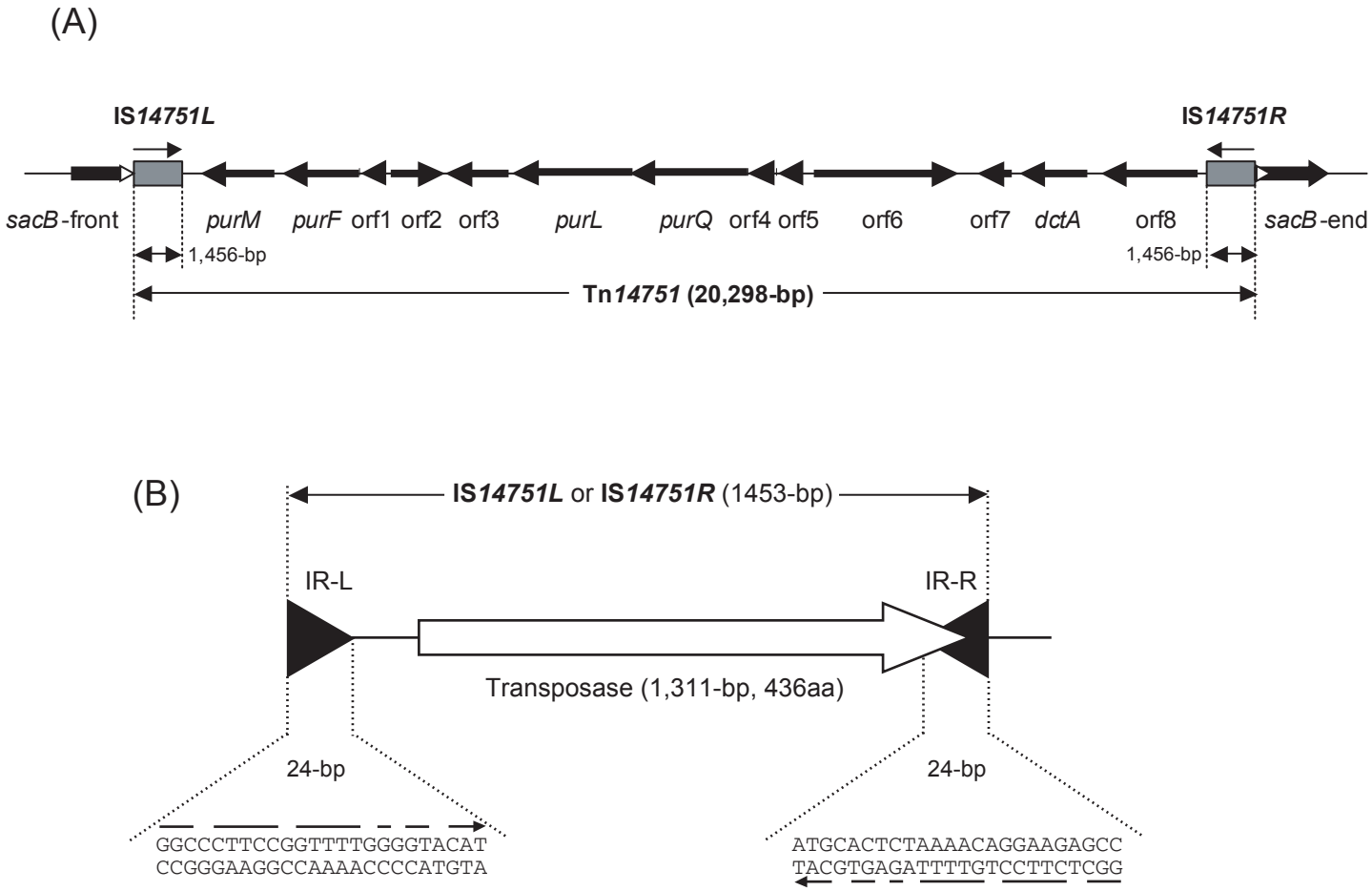



**Fig. 7.** Genetic and physical maps of the composite transposon Tn*14751*. (A) The deduced amino acid sequences encoded by the 13 open reading frames in Tn*14751* were identified. The copies of IS*31831* elements constitute inverted repeats (represented by gray boxes). Eight-base-pair direct repeats are indicated by open triangles. The arrows indicate the directions of transcription. The *sacB* gene is indicated by a cross-hatched arrow. (B) IS*14751L* and IS*14751R*, which are identical, contain the transposase gene. The solid triangles indicate the 24-bp imperfect inverted repeats (IR-L and IR-R). The nucleotide sequences of IR-L and IR-R are indicated below the arrowheads. Nucleotides that were identical in two sequences are indicated by converging arrows. aa, amino acids.

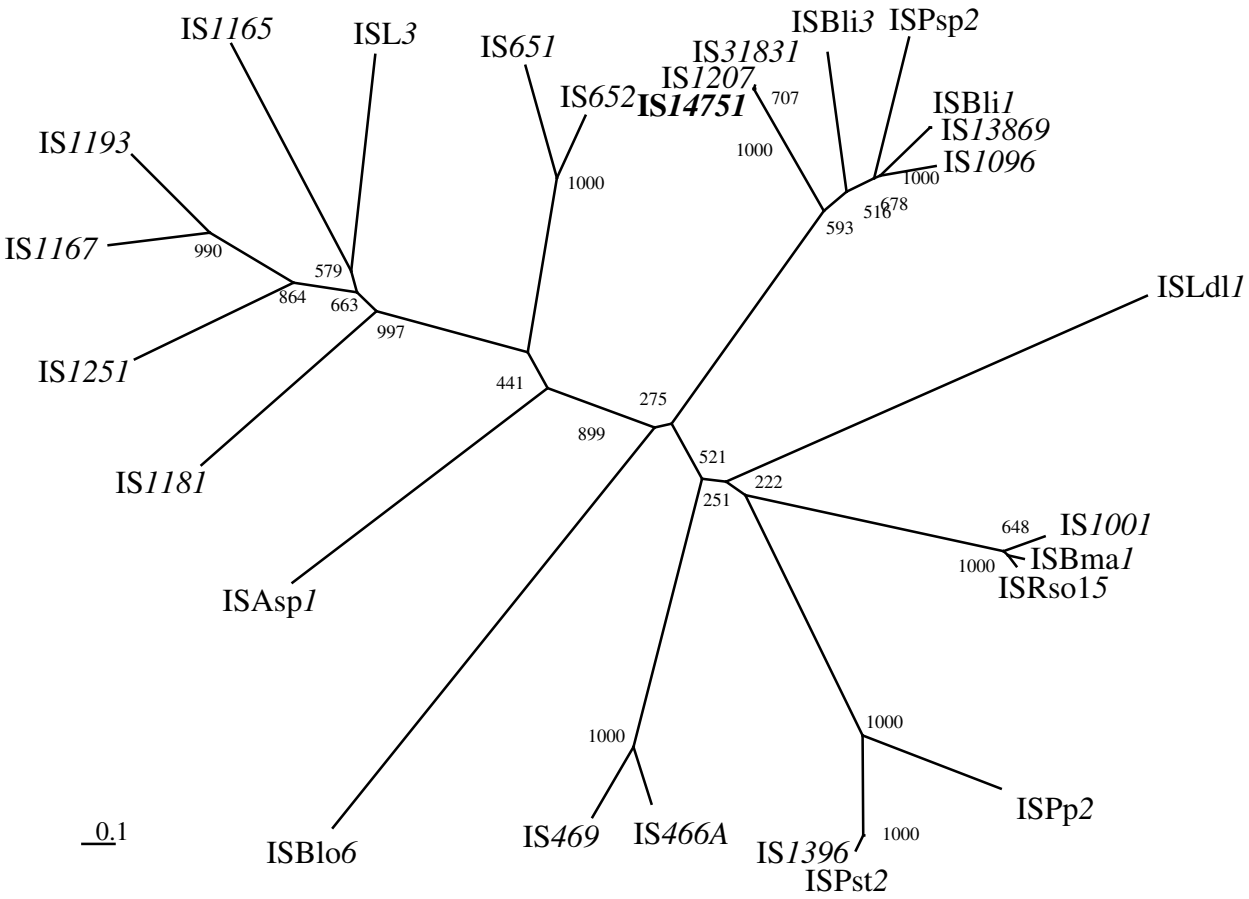



**Fig. 8.** Phylogenetic relationships of known insertion elements belonging to the ISL3 family based on the corresponding transposases. The topology of the phylogenetic tree was evaluated by performing a bootstrap analysis with 1,000 replicates. The scale bar equals a distance of 0.1.The accession numbers or source of information for the sequences used are as follows: IS*1207*, X96962; IS*31831*, D17429; ISBli*3*, http://www-is.biotoul.fr/is.html; ISPsp2, M57500; ISBli*1*, AF052055; IS*13869*, Z66534; IS*1096*, aM76495; ISLdl*1*, AJ302652; IS*1001*, X66858; ISBma*1*, AF285635; ISRso*15*, NC_003295; ISPp*2*, U25434; ISPst*2*, AJ012352; IS*1396*, AF027768; IS*466A*, AB032065; IS*469*, AB032065; ISBlo*6*, NC_004307; ISAsp*1*, U13767; IS*1181*, L14544; IS*1251*, L34675; IS*1167*, M36180; IS*1193*, Y13713; IS*1165*, aX62617; ISL*3*, X79114; IS*651*, NC_002570; and IS*652*, NC_002570.

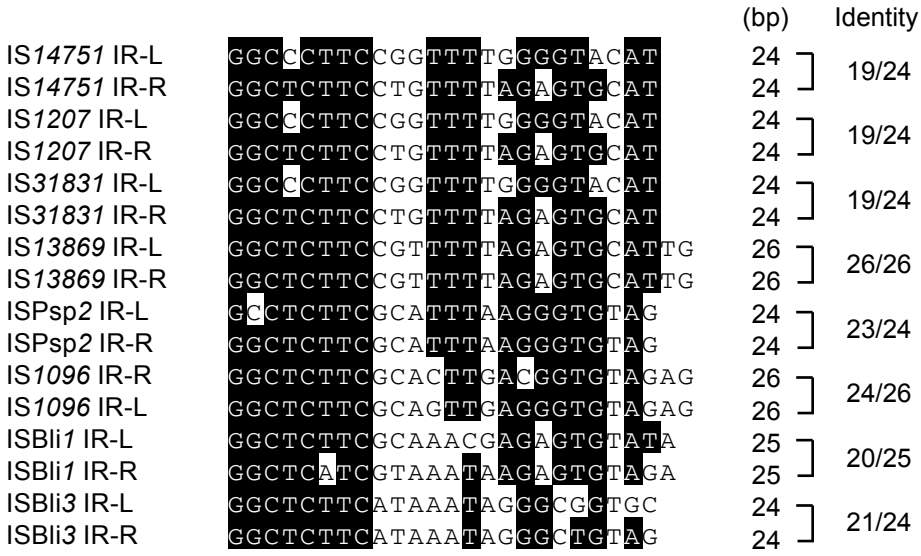



**Fig. 9.** Comparison of the inverted repeat sequences of IS*14751*, IS*1207*, IS*31831*, IS*13869*, ISPsp*2*, IS*1096*, ISBli*1*, and ISBli*3*. Nucleotides that were identical in at least 9 of 16 sequences are enclosed in boxes. IR-L and IR-R indicate the inverted repeats at the 5' end upstream and the 3' end downstream of the transposase gene, respectively. The length of each inverted repeat sequence and the identity ratio (number of identical nucleotides /length of inverted repeat) for IR-L and IR-R of each insertion element are indicated on the right.

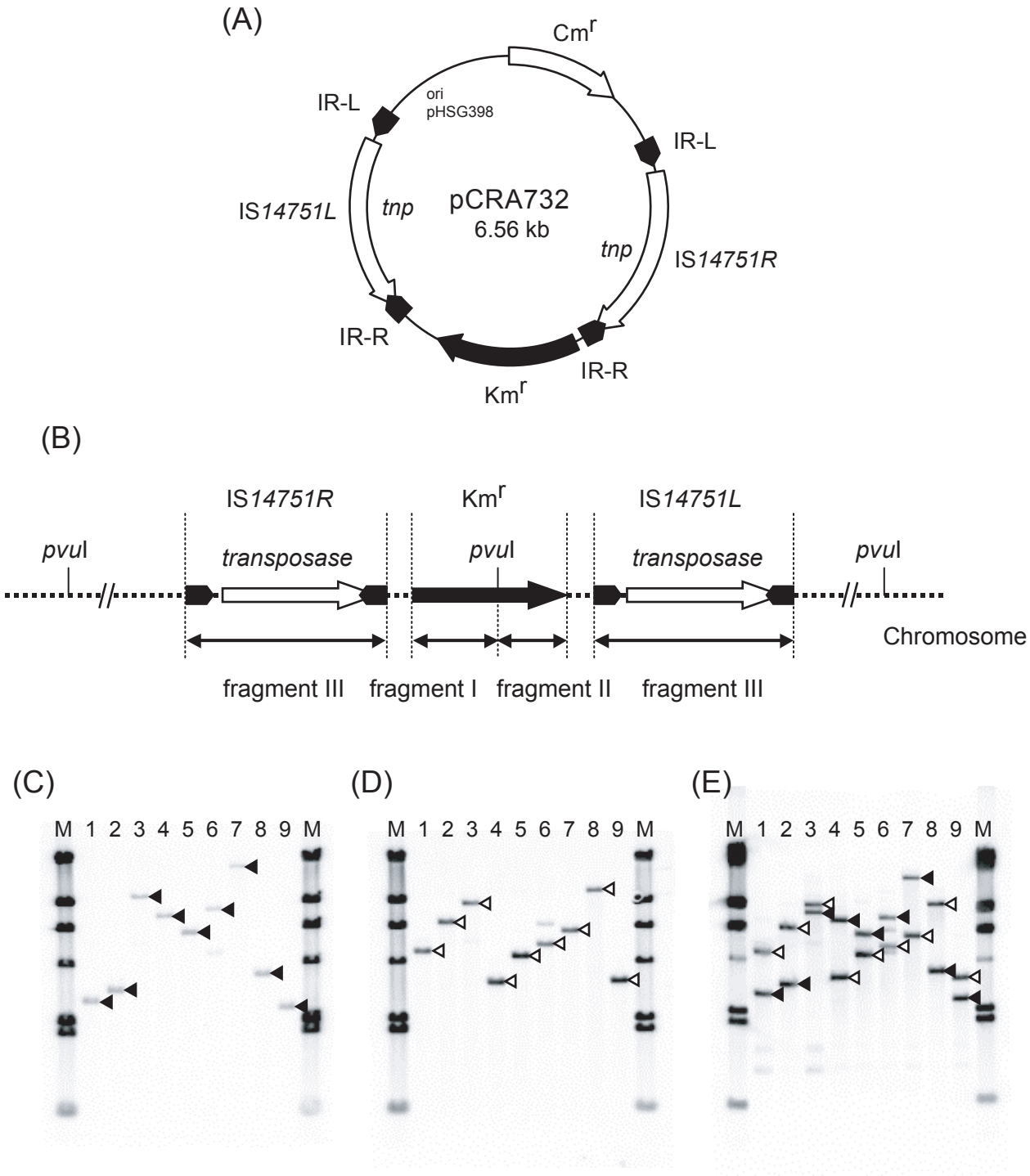
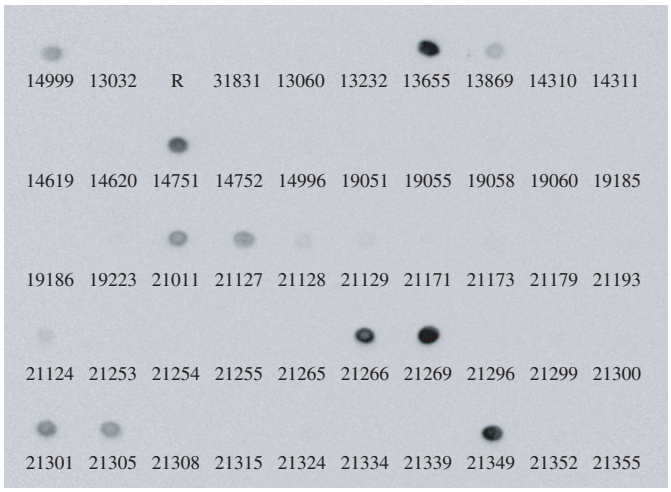



**Fig. 10.** (A) Physical and genetic map of pCRA732. Km$^r$, kanamycin resistance gene; Cm$^r$, chloramphenicol resistance gene; pHSG398 ori, origin of replication from pHSG398; *tnp,* transposase gene. The arrows indicate the direction of transcription.Transposition of mini-composite Tn*14751* into *C. glutamicum*. (B) Schematic physical map of the mini-composite Tn*14751* transposon in the chromosome of *C. glutamicum* mutants. The positions of Southern hybridization probes (fragments I, II, and III) are indicated by bidirectional arrows below the map. The transposase genes and inverted repeat sequences of IS*14751L* and IS*14751R* are indicated by open arrows and solid arrowheads, respectively. The kanamycin resistance gene is represented by a solid arrow. (C to E) Southern hybridization of *Pvu*I-digested chromosomal DNA from nine mini-composite Tn*14751* integrated *C. glutamicum* mutants with fragment I (C), fragment II (D), and fragment III (E) as the probes. M, *Hin*dIII molecular weight marker; lanes 1 to 9, nine mini-composite Tn*14751* integrated *C. glutamicum* mutants. (C) The solid arrowheads indicate the migration positions of hybridization signals when fragment I was the probe. (D) The open arrowheads indicate the migration positions of hybridization signals when fragment II was the probe. (E) Southern hybridization with fragment III as the probe resulted in two bands in each lane (lanes 1 to 9). The migration positions of signals corresponding to signals shown in panel C are indicated by solid arrowheads, and the migration positions of signals corresponding to signals shown in panel D are indicated by open arrowheads.

(A)




(B)

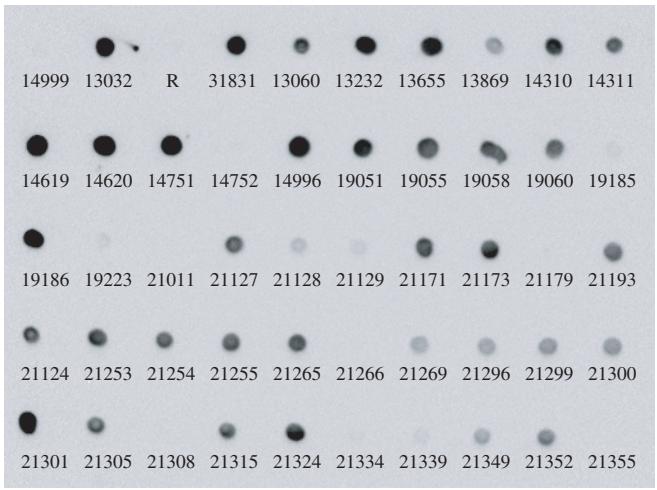



**Fig. 11 .** Dot blot analysis of chromosomal DNA from various corynebacterial strains hybridized with transposase sequence of IS*14999* (A) and IS*14751* (B), respectively. The numbers below the signals indicate the *C. glutamicum* strains. R means *C. glutamicum* R, while all other strains are American Type Culture Collection strains.

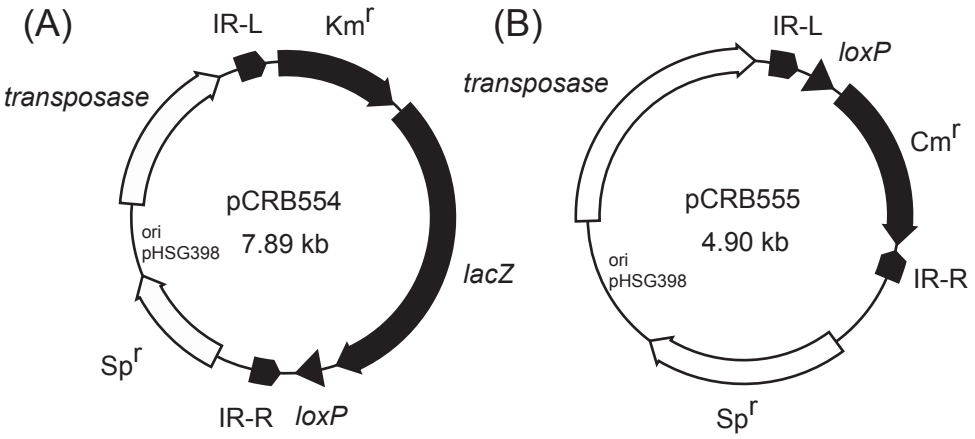



**Fig. 12.** Relevant features of mini-transposons, pCRB554 (A) and pCRB555 (B). $Cm^r$, chloramphenicol-resistance gene; $Sp^r$, spectinomycin-resistance gene; $Km^r$. kanamycin-resistance gene; IR-L, left direct repeat; IR-R, right direct repeat. Nucleotide sequence located between IR-L and IR-R transpose into chromosome.

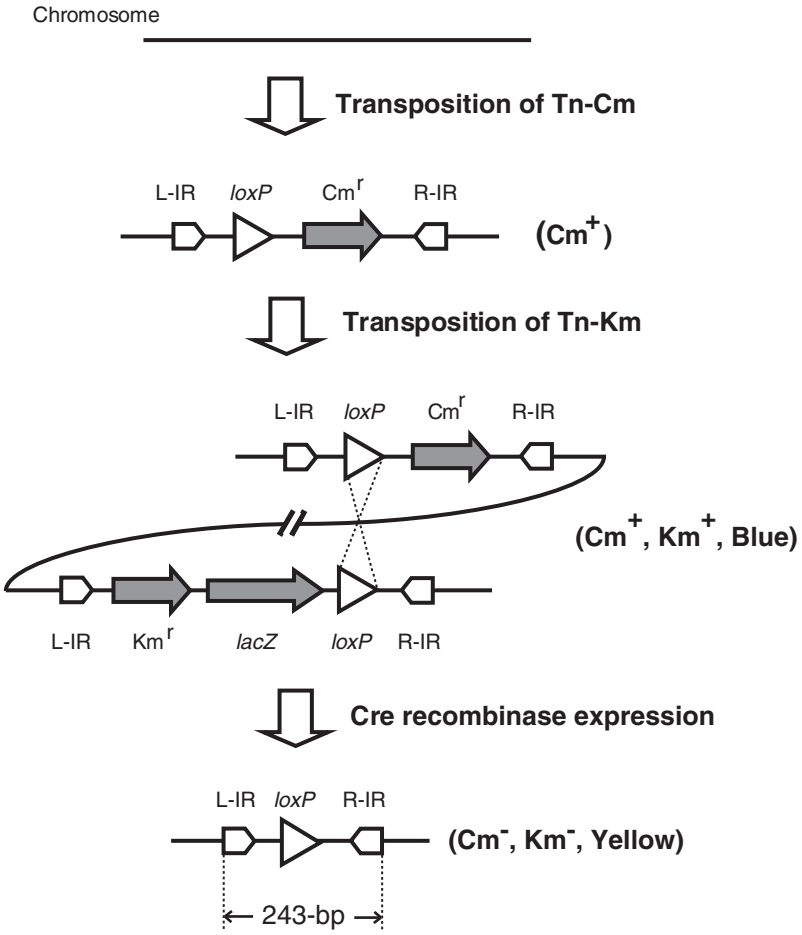



**Fig. 13.** Flow scheme for genome deletion. Resultant mutants of Tn-Cm and Tn-Km double transpositions show chloramphenicol and kanamycin resistance, respectively, and appeared blue on X-gal containing medium. Genome deletion strains after Cre recombinase expression show both chloramphenicol and kanamycin sensitivities, and appeared yellow on X-gal containing medium.

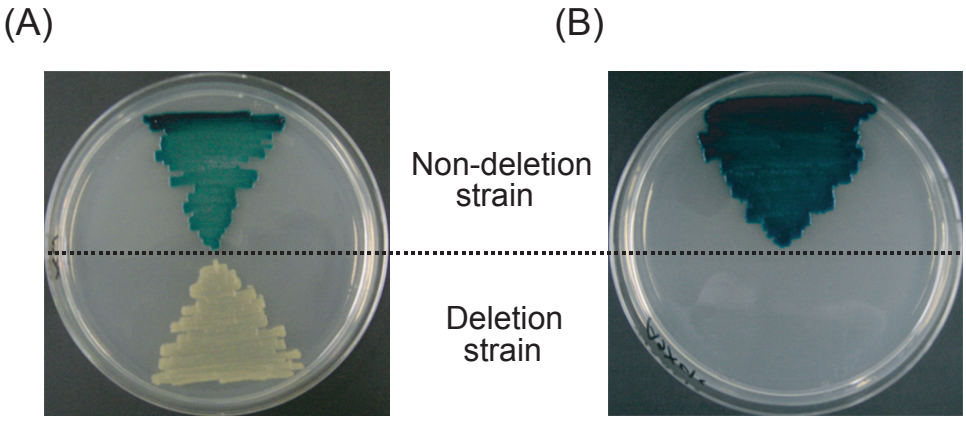



**Fig. 14.** Genome deletion strains were selected by chloramphenicol or kanamycin resistance, and change in colony color. Deletion strains and non-deletion strains were plated on A medium containing spectinomycin and X-gal (A); chloramphenicol, kanamycin, spectinomycin and X-gal (B).

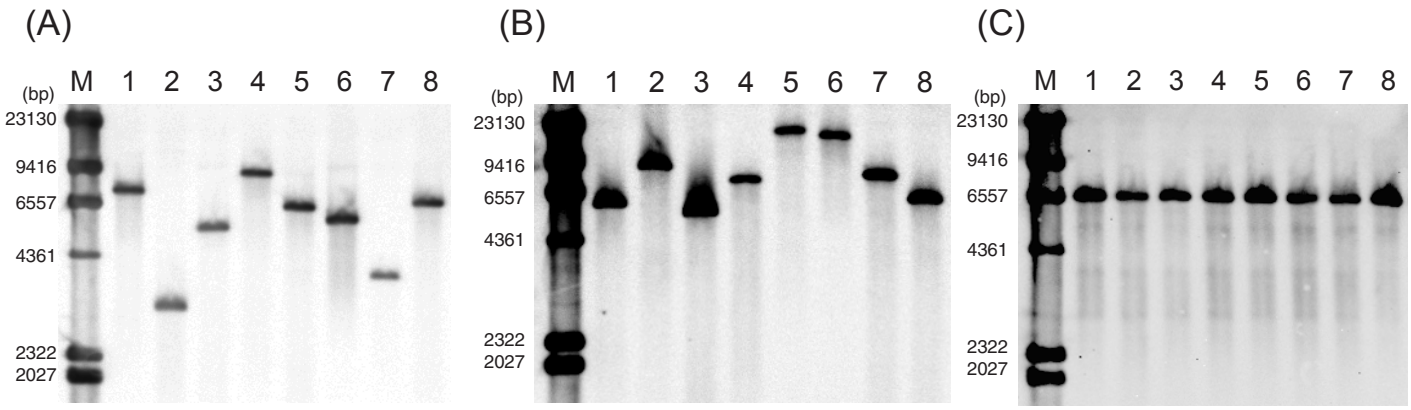



**Fig. 15.** Southern hybridization analysis for random transposition of mini-transposons, Tn-Cm (A) and Tn-Km (B), and stable existence of first transposition region after second transposition (C). In (C), chromosomal DNA of Tn-Km transposition strains into the strain of lane 8 in (A) were used with a chloramphenicol probe. M, standard molecular size marker.

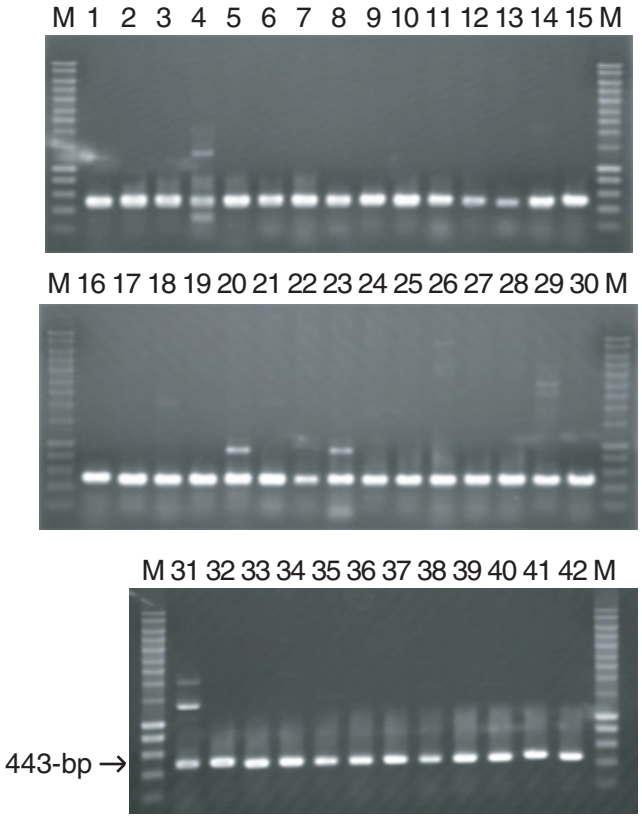



**Fig. 16.** Genome deletion confirmation by PCR. Chromosomal DNA of deletion strains was used as template. Each primer was designed based on the sequence 100-bp from the deletion region. A total of 443-bp nucleotide sequence (243-bp remaining after deletion plus 100-bp for each primer) was amplified.

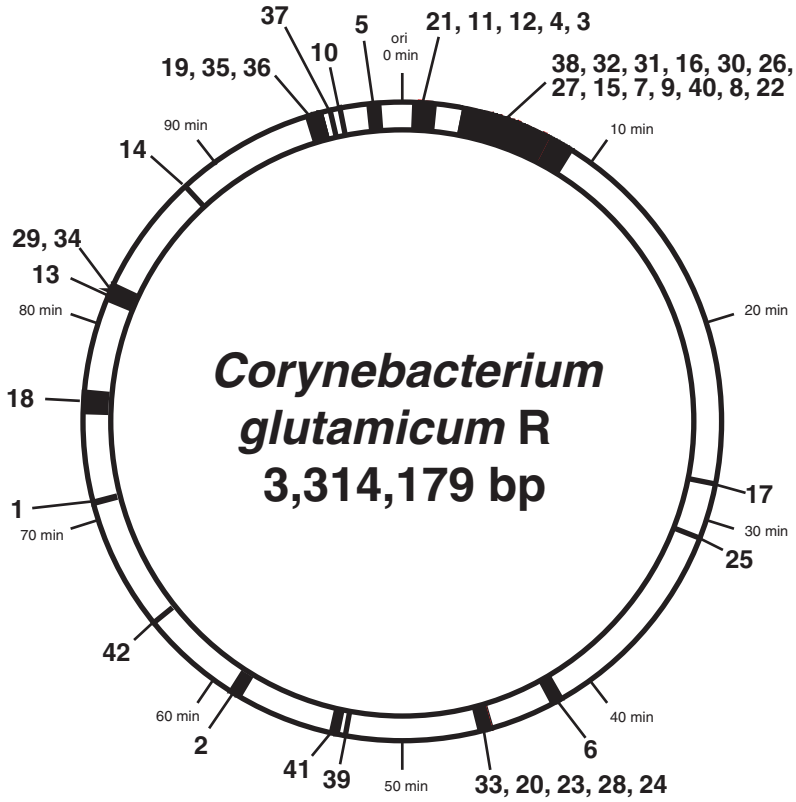



**Fig. 17.** A circular map of *C. glutamicum* R chromosome, showing deleted regions (filled) and resultant strains (numbered).

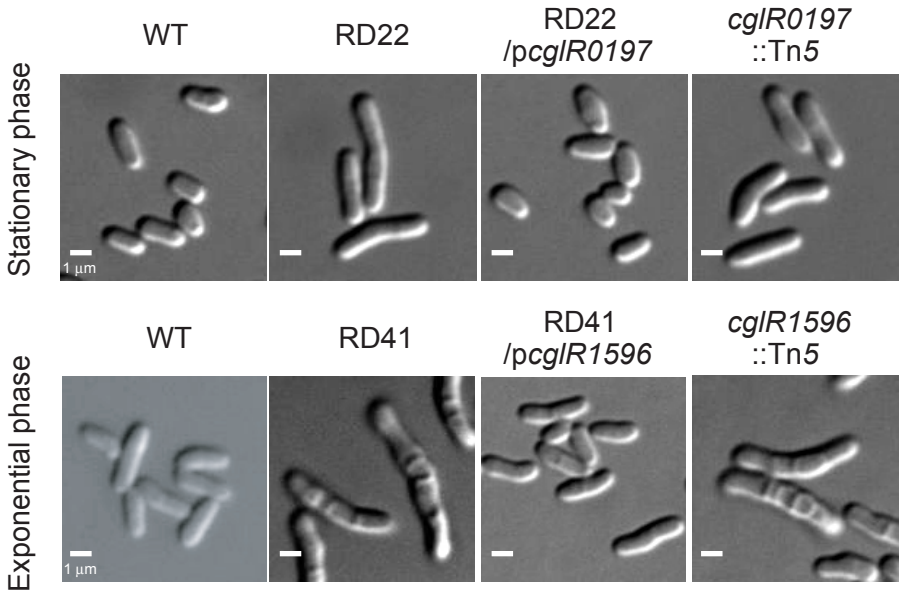



**Fig. 18.** *cglR0197* and *cglR1596* are responsible for morphological change of each deletion strains, RD22 and RD41.

| | +1 | +2 | +3 | +4 | +5 | +6 | +7 | +8 |
|---|---|---|---|---|---|---|---|---|
| A | 31 | 11 | 15 | 36 | 61 | 69 | 14 | 18 |
| T | 14 | 21 | 73 | 64 | 39 | 19 | 8 | 29 |
| G | 31 | 36 | 9 | 0 | 0 | 10 | 31 | 18 |
| C | 24 | 31 | 4 | 0 | 0 | 3 | 48 | 36 |
| A+T (%) | 45 | 33 | 88 | 100 | 100 | 88 | 21 | 46 |
| G+C (%) | 55 | 68 | 13 | 0 | 0 | 13 | 79 | 54 |
| | **G** | **G** | **T** | **T** | **A** | **A** | **C** | **C** |

**Fig. 19.** Eight-bp direct repeat sequence of IS*31831*. Numbers represent the percentage occurrence of the preferred base at the positions indicated. The analysis is based on 82 independent insertions sites in 42 genome deletion strains. Black backgrounds show duplicated bases. Grey backgrounds indicate preferred bases. Bases below the numbers show the most preferred base at each position.

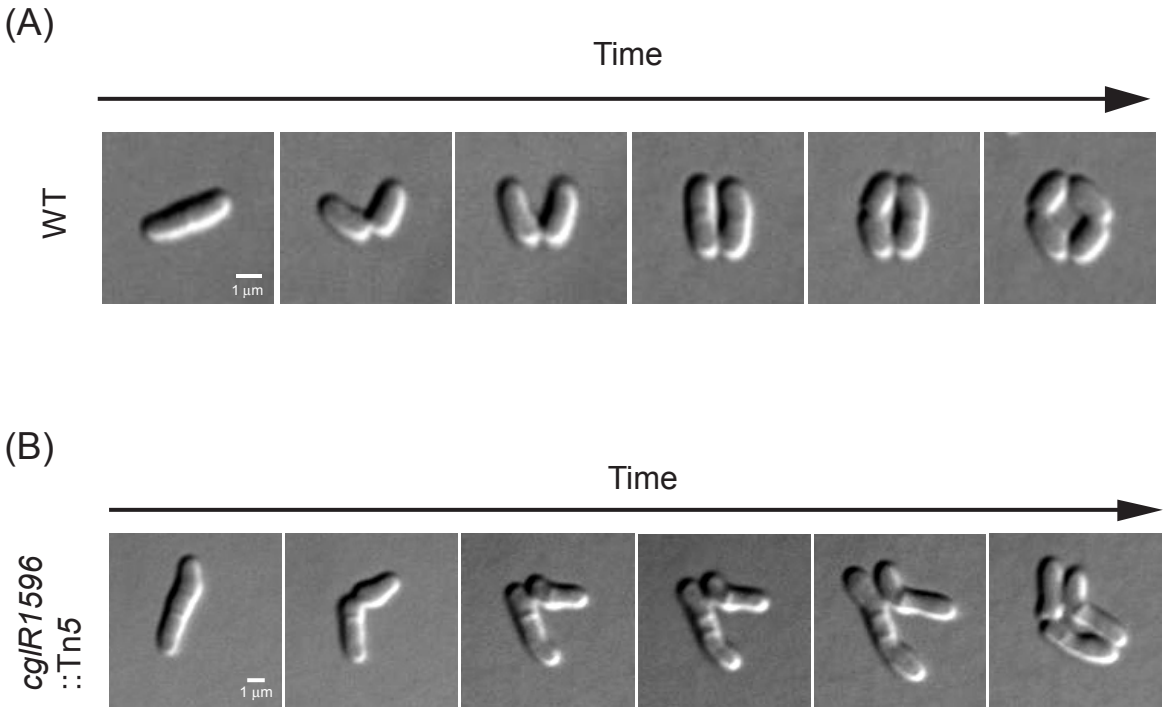



**Fig. 20.** Time lapse analysis of wild type (A) and c*glR1596*::Tn*5* (B) mutant showing snapping division. Predivisional cell was splited and rapidly divided into two daughter cells.

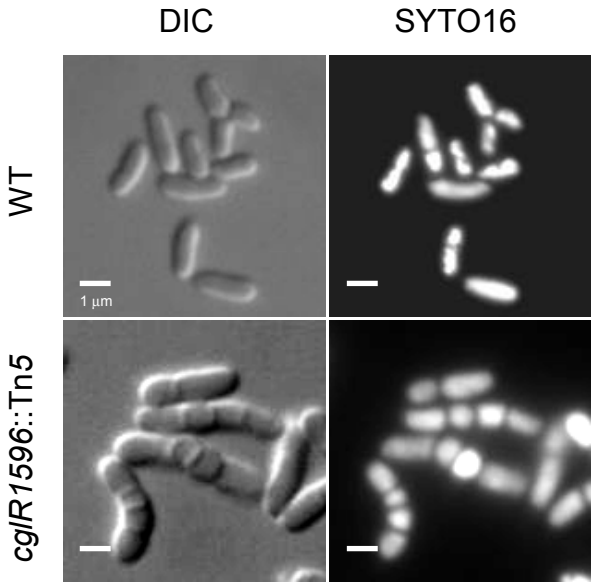



**Fig. 21.** Inactivation of *cglR1596* causes severe morphological change. Left shows phase contrast images and, right, corresponding fluorescence images. c*glR1596* mutant cells contain more than two nucleoids, implying a defect of cell wall separation. The exposure times were 0.1 s for phase-contrast microscopy, 1 s for SYTO16.

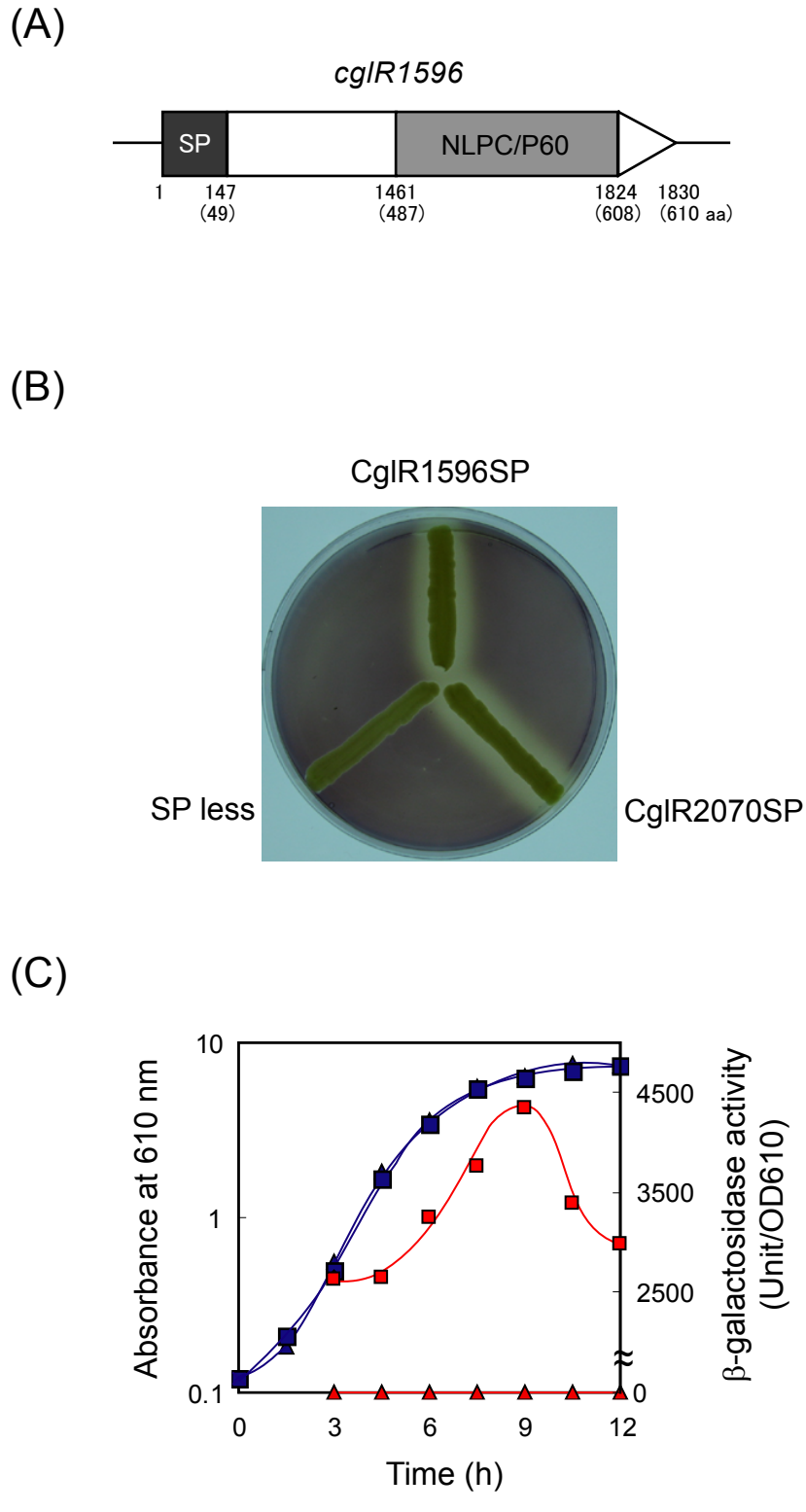



**Fig. 22.** (A) Prediction of gene organization of the *cglR1596* gene. CglR1596 has a signal peptide, and a NLPC/P60 domain at the C-terminus. This domain is found in some cell wall hydrolases. (B) CglR1596 and CglR2070 secretes outside of the cell. Promoter less α-amylase from *Geobacillus stearothermophilus* was used to check secretion activity of signal peptide. Amylase activity was detected by iodo-starch reaction. (C) The β-galactosidase assay using *cglR1596-lacZ* fusion and promoterless *lacZ* gene showing squares and triangles, respectively. Blue and red symbols indicate cell growth and β-galactosidase activity, respectively. A high activity was observed in late exponential phase, implying that many cell separation events are occurred at this phase.

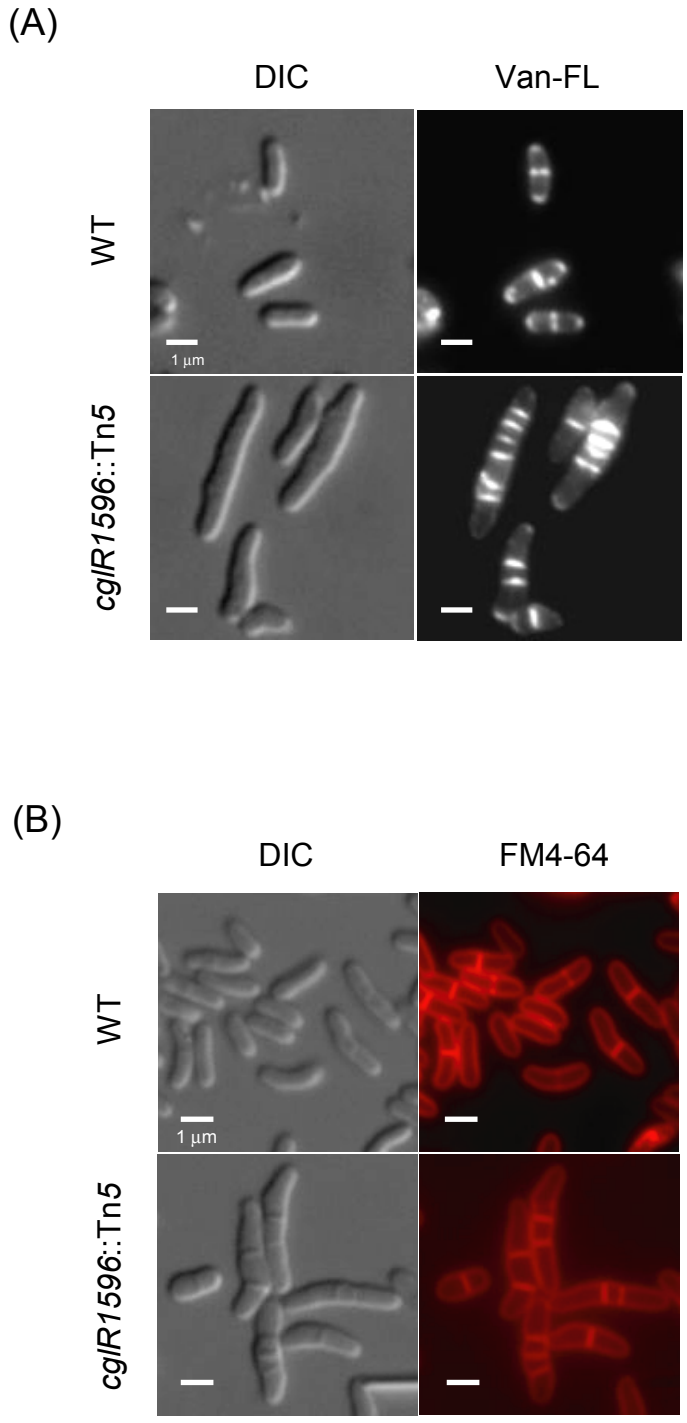



**Fig. 23.** (A) Localization pattern of nascent peptidoglycan synthesis with Van-FL staining. Left shows phase contrast images and, right, corresponding fluorescence images. In the wild type, peptidoglycan is synthesized at cell poles and midcell. On the contrary, *cglR1596*::Tn*5* cells mainly synthesize peptidoglycan at septa. (B) Plasma membrane staining using FM4-64. Left shows phase contrast images and, right, corresponding fluorescence images.Thickness *cglR1596*::Tn*5* mutant cell was caused by the enlargement of cytoplasm. CglR1596 may be responsible for determination of cell width. The exposure times were 0.1 s for phase contrast microscopy, 5 s for Van-FL and FM4-64.

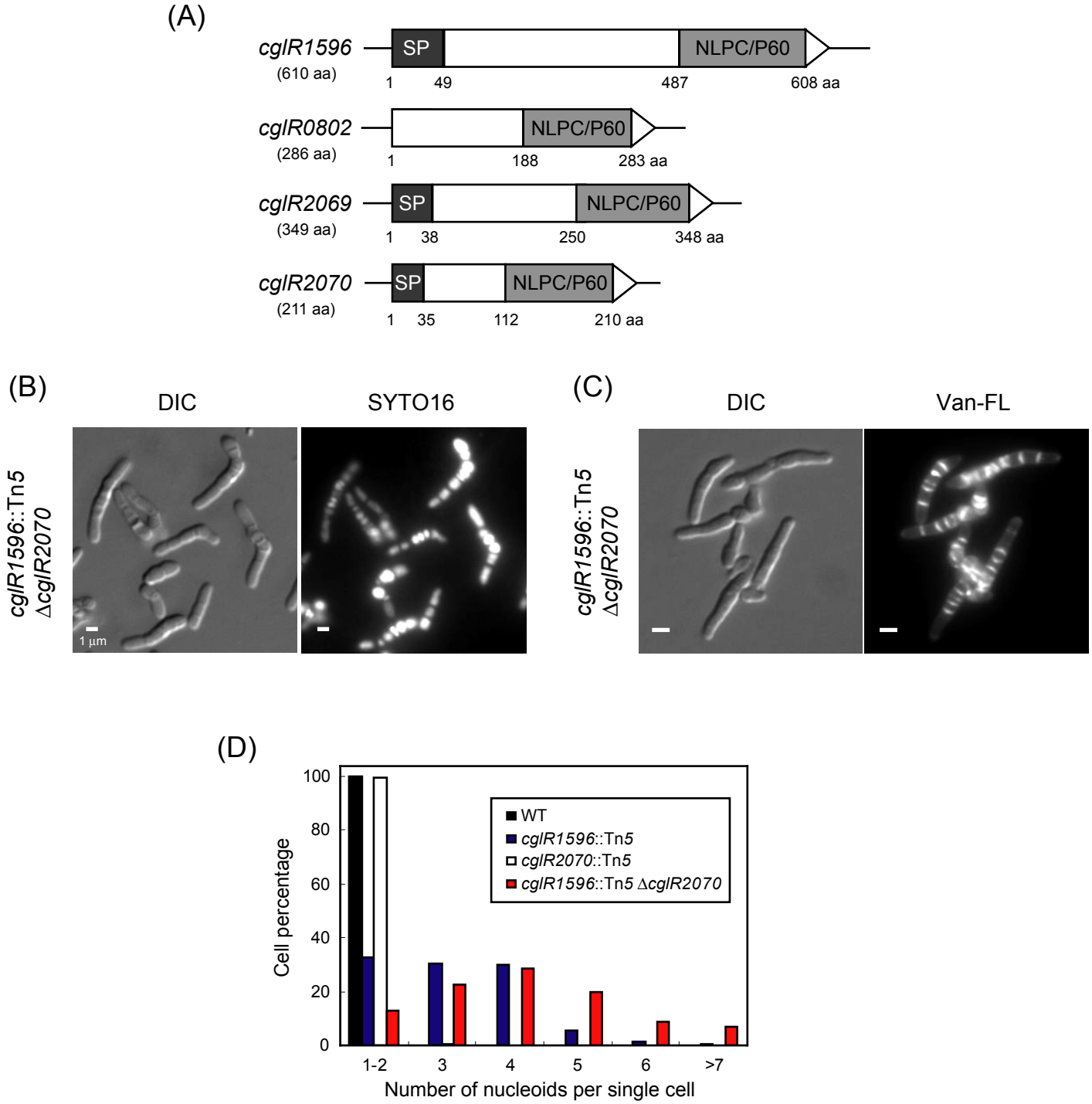



**Fig. 24.** (A) Schematic representation of all four cell wall hydrolases in *C. glutamicum* R including *cglR1596*. Whole amino acid length of each gene shows under gene name. Numbers under the picture represent amino acid number of each motif. aa, amino acid. (B) SYTO16 staining of double inactivation of *cglR1596* and *cglR2070*. (C) Van-FL staining of double inactivation of *cglR1596* and *cglR2070*. (D) Number of nucleoids per single cell of wild type and mutants. A total of 1,000 exponential growth cells were counted. Average number of nucleoids is 1.39 (WT), 1.41 (*cglR2070*::Tn*5*), 3.11 (*cglR1596*::Tn*5*), 4.11 (*cglR1596*::Tn*5* Δ*cglR2070*). The exposure times were 0.1 s for phase-contrast microscopy, 1 s for SYTO16 and 5 s for Van-FL.

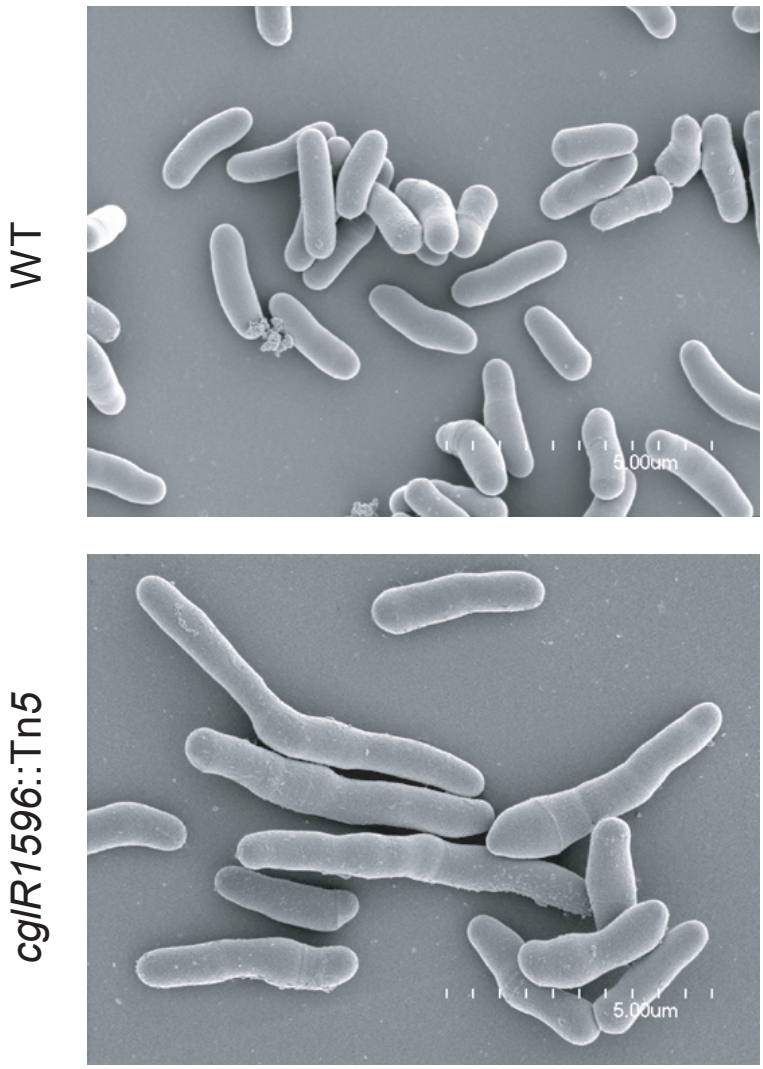



**Fig. 25.** Scanning electron microscopy analysis of wild type and *cglR1596*::Tn*5* .

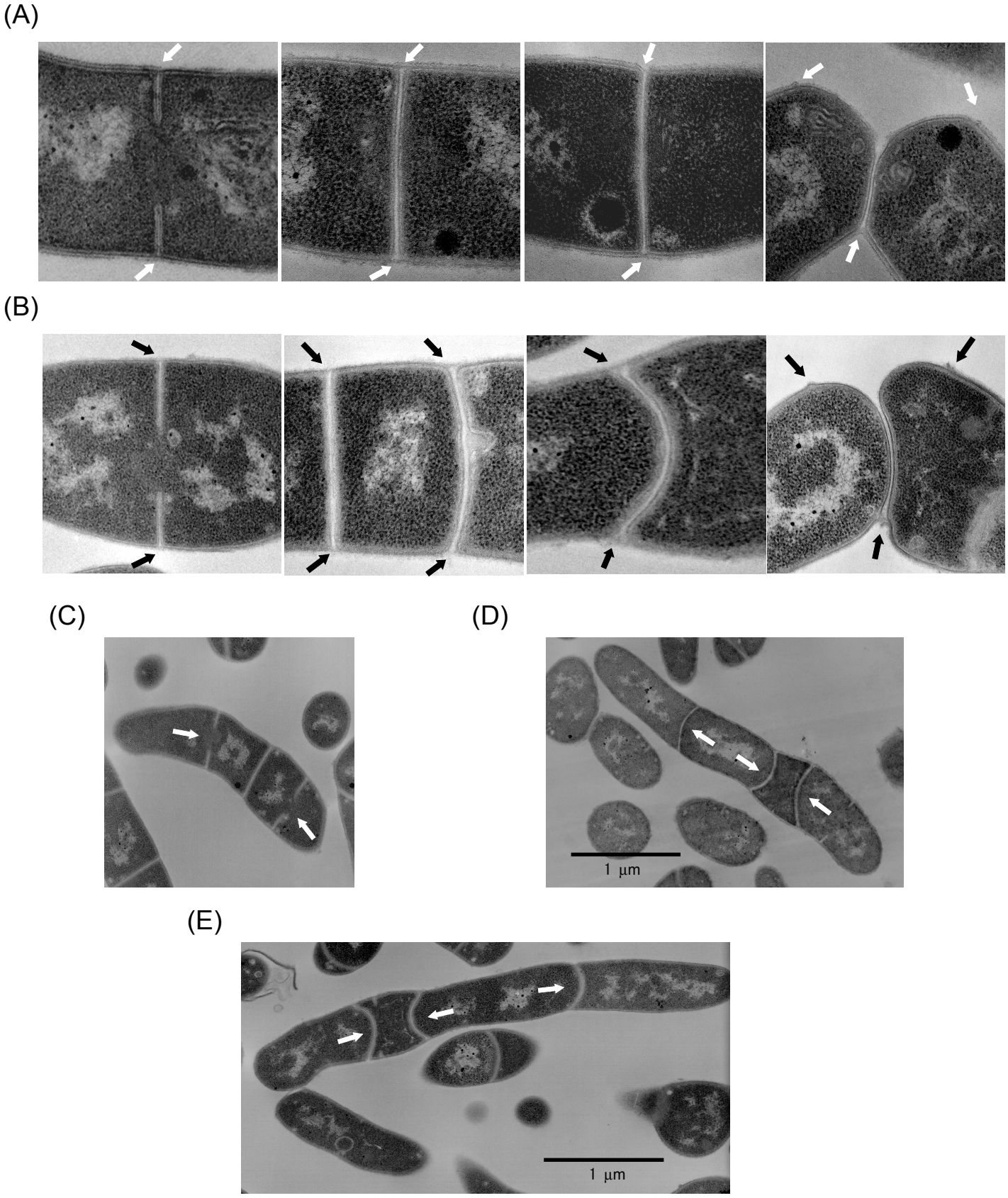



**Fig. 26.** Transmission electron microscopy analysis of cell separation event in wild type (A) and *cglR1596* mutant (B). During and even after septum formation is complete, two sister cells were cross-linked with cell wall (arrows). After septum formation is complete, cell wall at junction point starts to be hydrolyzed (black arrow in A). Even when septum starts to bend, sister cells are still connected in *cglR1596* mutant (B), indicating that CglR1596 is required for cell separation. After cell separation, one side of the junction point was still cross-linked, but, the other side was separated and scars were observed in both wild type (A) and *cglR1596* mutant (B). Two-septa formation is occurred in one cell, indicating that cell separation is independent of chromosome segregation and septum formation (C). In some c*glR1596* mutants, bending septa were observed, indicating that polar growth of *C. glutamicum* stresses the cell membrane (D, E).

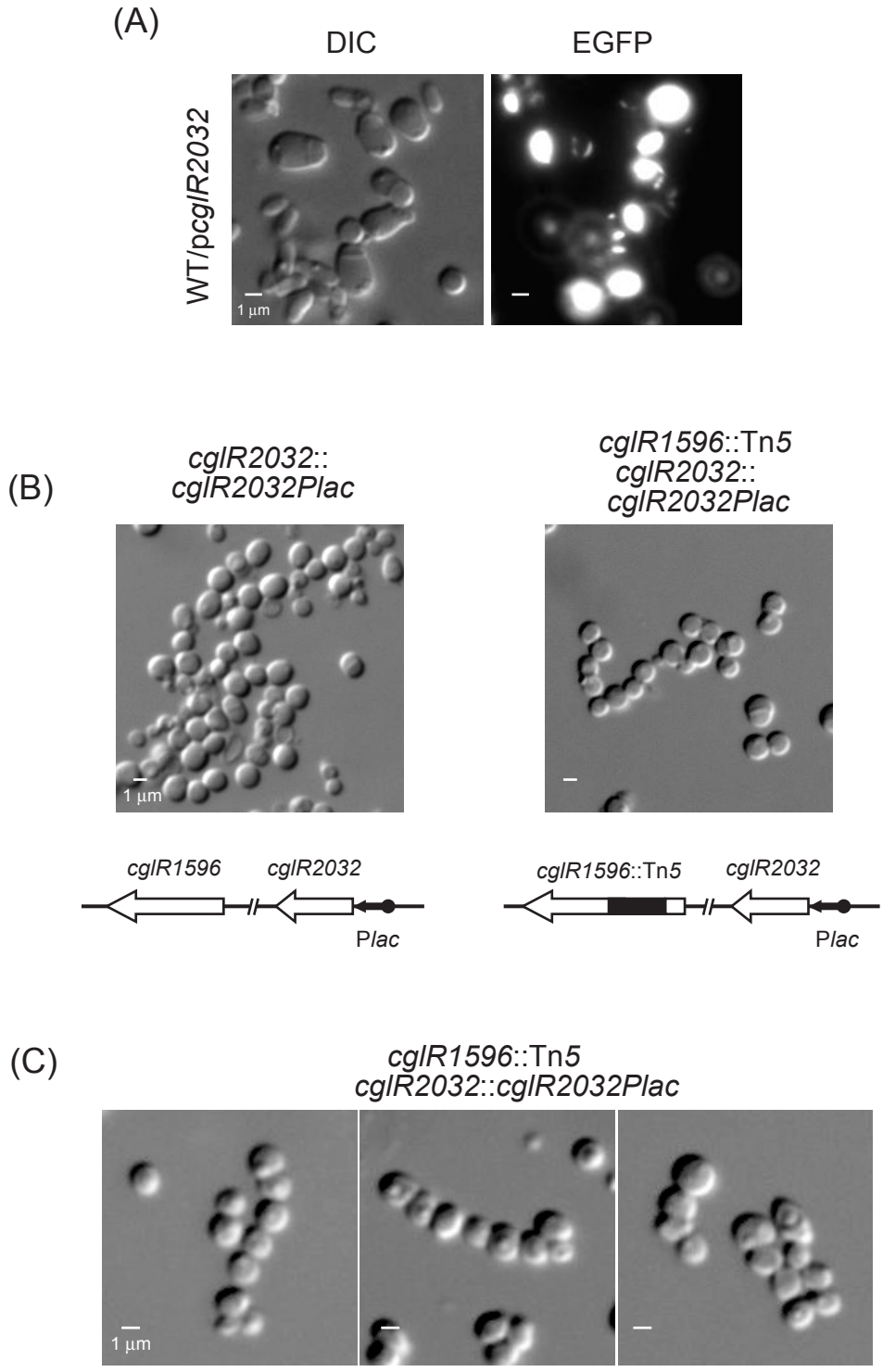



**Fig. 27.** (A) Overexpression of *cglR2032* results in (B) Reduced expression of *cglR2032* gene results in round cell shape in wild type and *cglR1596*::Tn*5* mutant. Promoter of *cglR2032* was replaced with *lac* promoter. CglR2032 is seemed to play a critical role for maintenance of rod shape of *C. glutamicum* cell. Some *cglR1596*::Tn*5 cglR2032*:: *cglR2032Plac* mutants show chained phenotype implying that CglR2032 may be involved in snapping division (C).

**Table 1.** Strains and plasmids used in this study

| Strains and Plasmids | | Relevant Genotype or Description | Reference or Source |
|---|---|---|---|
| Strains | | | |
| *E. coli* | | | |
| | JM109 | *recA1, endA1, gyrA96, thi, hsdR17, supE44, relA1, D(lac-proAB), F'[traD36, proAB+, lacIqZDM15]* | TAKARA |
| | JM110 | *rpsL (Strr), thr, leu, thi-1, lacy, galK, galT, ara, tonA, tsx, dam, dcm, supE44, D(lac-proAB), F'[traD36 proAB lacIqZ Δ M15]tsx, D(lac-proAB)/F'[traD36, proAB+, lacIq, lacZDM15]* | TAKARA |
| | BL21DE3/pLysS | *F⁻ ompT hsdSB(rB⁻ mB⁻) gal dcm (DE3) pLysS (CmR)* | TAKARA |
| *C. glutamicum* | | | |
| | R | Wild-type strain | Kotrba *et al* ., 2001 |
| | ATCC14751 | Wild-type strain | ATCC |
| | ATCC14999 | Wild-type strain | ATCC |
| | RD41 | Genome deletion strain from *cglR1585 -cglR1604* | This study |
| | CHM1 | *cglR1596* ::Tn*5* | Suzuki *et al* ., 2006 |
| | CHM2 | *cglR1596* ::Tn5/L-1596SP-Ami | This study |
| | CHM3 | *cglR15* 96::Tn5/L-1596P-lacZ | This study |
| | CHM4 | *cglR15* 96::Tn5 Δ*cglR2070* | This study |
| | CHM5 | *cglR2070* ::Tn*5* | This study |
| | CHM6 | *cglR1596::* Tn*5 ΔcglR2070 ΔcglR802 ΔcglR2069* | This study |
| | CHM7 | WT/pEAG-2032egfp2 | This study |
| | CHM8 | *divIVAcg* native promoter replaced with *E. coli* lac promoter | This study |
| | CHM9 | *cglR1596* ::Tn*5, divIVAcg* native promoter replaced with *E. coli lac* promoter | This study |
| Plasmids | | | |
| | pHSG398 | $Cm^r$, Cloning vector for *E. coli* | TAKARA |
| | pUC4K | $Km^r$, Cloning vector for *E. coli* | Taylor and Rose, 1988 |
| | Lsv5-9 | $Sp^r$, *E. coli* / *C. glutamicum* shuttle vector; low copy number | Vertes *et al* ., 1994 |
| | pEAG2 | $Cm^r$, $Km^r$, *E . coli* / *C. glutamicum* shuttle vector containing $divIVA_{Bl}$ fused with *egfp2* ; high copy number | Ramos *et al* ., 2003 |
| | pMV5 | $Sp^r$, Lsv5-9 containing *sacB* gene | Vertes *et al* ., 1994 |
| | plox3 | $Cm^r$, α*lac* /MCS (with *loxP* ) | Suzuki *et al* ., 2005a |
| | LsvLsv-PtacAmiEcoRV | $Cm^r$, α-amylase gene with *tac* promoter | Watanabe, unpublished |
| | L-lacZ | $Cm^r$, promoter less *lacZ* gene | Suda, unpublished |
| | pCRB512 | $Sp^r$, pMV5 derivative carrying the *sacB* gene interrupted by IS*14999* | This study |
| | pCRB201 | $Cm^r$, 3.1-kb *Hpa* I-*Dra* I fragment from pCRB512 inserted into *Hin* dIII digested and blunted fragment of pHSG398 | This study |
| | pCRB203 | $Cm^r$, $Km^r$, 1.2-kb PCR fragment (pUC4K *kan* region) digested with *Eco* RV and inserted into *Hin* dIII digested and blunted fragment of pCRB201; Tn*14999* | This study |
| | pCRA730 | $Sp^r$, pMV5 with 20.3-kb Tn14751 This study | This study |
| | pCRA731 | $Cm^r$, pHSG398 with a 20.3-kb *Hpa*I-*Hap* I DNA fragment containing the entire Tn*14751* transposon | This study |
| | pCRA732 | $Cm^r$, $Km^r$; pHSG398 with IS*14751* L, Kmr, and IS*14751* R (mini-composite Tn14751) | This study |
| | pCRB504 | $Cm^r$, IS*31831* | This study |
| | pCRB505 | $Sp^r$, Cre expression vector | This study |
| | pCRB506 | $Sp^r$, pHSG398 modified | This study |
| | pCRB507 | $Sp^r$, pCRB506 carrying transposase of IS*31831* | This study |
| | pCRB535 | $Km^r$, $Cm^r$, plox3 carrying kanamycin resistance cassette | This study |
| | pCRB536 | $Km^r$ $Cm^r$, pCRB535 carrying *lacZ* gene | This study |
| | pCRB539 | $Cm^r$, plox3 carrying chloramphenicol resistance cassette | This study |
| | pCRB554 | $Km^r$, $Sp^r$, pCRB507 with a 4.3-kb PCR fragment with a *loxP,* L-IR, R-IR, kanamycin resistance cassette and *lacZ* gene (Tn-Km) | This study |
| | pCRB555 | $Cm^r$, $Sp^r$, pCRB507 with a 1.3-kb PCR fragment with a *loxP* , L-IR, R-IR and chloramphenicol resistance cassette (Tn-Cm) | This study |
| | L-1596SP-Ami | $Cm^r$, *cglR1596* signalsequence fused with α-amylase gene | This study |
| | L-1596PlacZ | $Cm^r$, promoterless *lacZ* gene under control of *cglR1596* promoter | This study |
| | pCold-His6-1596 | $Ap^r$, pCold-I vector derivative, cloning *cglR1596* | This study |
| | pEAG-2032egfp2 | $Cm^r$, $Km^r$, cglR2032fused with *egfp2* based on pEAG2 | This study |

Table. 2 Primers used in this study

| Name | Sequence (5'-3') | Application |
|---|---|---|
| P1 | AATTGATATCCTGAGGTCTGCCTCGTGAA | Km resistance gene, cloning to pCRB201 |
| P2 | TAAGATATCTGTGTCTCAAAATCTCTGA | Km resistance gene, cloning to pCRB201 |
| P3 | GAAGGATCAGATCACGCATC | Inverse PCR primer for determination of insertion site of Tn*14999* |
| P4 | CAGGTGACATGGATCAGCGT | Inverse PCR primer for determination of insertion site of Tn*14999* |
| 1 | GGCCCTTCCGGTTTTGGGGTACAT | IS*14751* L or IS*14751* R (fragment III) |
| 2 | GGCTCTTCCTGTTTTAGAGTGCAT | IS*14751* L or IS*14751* R (fragment III) |
| 3 | AGTCAGATCTAAGTGGAGCACCTAGATCGC | IS*14751* L and IS*14751* R, cloning to pHSG398 |
| 4 | AGTCAGATCTAGTCACGCACATCTTCTGCA | IS*14751* L and IS*14751* R, cloning to pHSG398 |
| 5 | AGTCAGATCTTGTGTCTCAAAATCTCTGA | Km resistance gene, cloning to pHSG398 |
| 6 | AGTCAGATCTCTGAGGTCTGCCTCGTGAA | Km resistance gene, cloning to pHSG398 |
| 7 | ATGAGCCATATTCAACGGGA | Left part of Km resistance gene (fragment I) |
| 8 | GGACAATTACAAACAGGAAT | Left part of Km resistance gene (fragment I) |
| 9 | CGTATTTCGTCTCGCTCAGG | Right part of Km resistance gene (fragment II) |
| 10 | TTAGAAAAACTCATCGAGCA | Right part of Km resistance gene (fragment II) |
| 157 | GAGAATTCCACAGAACCTGGGCTAGC | IS*31831* , cloning to pCRB506 |
| 158 | GAGAGCTCTTAGAGTGCATTGATCTT | IS*31831* , cloning to pCRB506 |
| 113 | GAGAATTCAAGCCACGTTGTGTCTCAAA | Km resistance gene, cloning to plox3 |
| 114 | GAGGATCCCAACTCAGCAAAAGTTCGAT | Km resistance gene, cloning to plox3 |
| 121 | GAAAGCTTTCCGTCGAACGGAAGATCAC | Cm resistance gene, cloning to plox2 |
| 122 | GAAAGCTTTGCGCTCACTGCCCGCTTTC | Cm resistance gene, cloning to plox2 |
| LacZF | GATCTAGAAGGCTTTACACTTTATGCTTCCGGC | *lacZ* gene, cloning to pCRB535 |
| LacZR | GATCTAGACGACGGCCAGTAAGCTTGCATGCCT | *lacZ* gene, cloning to pCRB535 |
| 163 | GAACTAGTGGCCCTTCCGGTTTTGGGGTACATCACAGAAAGCCACGTTGTGTCTCAAAATCTC | Transposition region of pCRB536 containing L-IR, cloning to pHSG398 |
| 164 | GAACTAGTGGCTCTTCCTGTTTTAGAGTGCATTGATCTTGTAAAACGACGGCCAGTGCCAAGC | Transposition region of pCRB536 containing R-IR, cloning to pHSG398 |
| 165 | GAACTAGTGGCCCTTCCGGTTTTGGGGTACATCACAGAAAGGCTTTACACTTTATGCTTCCGGC | Transposition region of pCRB539 containing L-IR, cloning to pHSG398 |
| 166 | GAACTAGTGGCTCTTCCTGTTTTAGAGTGCATTGATCTTTGCGCTCACTGCCCGCTTTCCA | Transposition region of pCRB539 containing R-IR, cloning to pHSG398 |
| RDSF | ACCGTTCGTATAGCATACATTATACGAAGTTATG | Inverse PCR primer for determination of deleted region |
| RDSR | CGGGTACCGAGCTCGAATTCGTAATCATGG | Inverse PCR primer for determination of deleted region |
| 1596SPF-SmaI | GACCCGGGAAAGTGATAAACATCACAAA | *cglR1596* promoter, cloning to L-lacZ |
| 1596SPR-SmaI | GACCCGGGTAGATTGGCCACATGTTCTC | *cglR1596* promoter, cloning to L-lacZ |
| 1596SP1F-EcoRV | GAGATATCCGGATTGAACAGGAGAACAT | CglR1596 SP1, 2 cloning to Lsv-PtacAmiEcoRV |
| 1596SP1R-EcoRV | GAGATATCCTGTGGCTGCGCCACCGCCG | CglR1596 SP1 cloning to Lsv-PtacAmiEcoRV |
| 1596SP2R-EcoRV | GAGATATCATTTTCCTCTGCCTGTGCAA | CglR1596 SP2 cloning to Lsv-PtacAmiEcoRV |
| 1596SP3F-EcoRV | GAGATATCCCCAGGTAAAAGGTGTCGCG | NCgl1480 SP3 cloning to Lsv-PtacAmiEcoRV |
| 1596SP3R-EcoRV | GAGATATCCTGTGGCTGCGCCACCGCCG | NCgl1480 SP3 cloning to Lsv-PtacAmiEcoRV |
| 1596F-EcoRI | GAGGTACCACTGATGAGCATTTGGGCAA | *cglR1596* whole region fused with myc tag at C-terminus |
| 1596R-myc-SalI | GAGTCGACCTACAGATCCTCTTCAGAGATGAGTTTCTGCTCAATGAGGCGTACCACACTCTC | *cglR1596* whole region fused with myc tag at C-terminus |
| 1596MLF-SacI | GAGAGCTCCAGCCACAGAATCCGGATGA | cglR1596 sequence except signal sequence cloning to pCold-I |
| 1596R-KpnI | GAGGTACCAATGAGGCGTACCACACTCT | *cglR1596* sequence except signal sequence cloning to pCold-I |
| 2070MLF-SacI | GAGAGCTCGCTGAGGTTGTTGTTCCTGG | *cglR2070* sequence except signal sequence cloning to pCold-I |
| 2070R-KpnI | GAGGTACCTTAGAAACGAACTGCAGAGT | *cglR2070* sequence except signal sequence cloning to pCold-I |
| 802FF-SmaI | GACCCGGGTTCGGCTTCTTCCTTGGCGA | *cglR0802* targeting vector |
| 802FR-XbaI | GATCTAGAAGCACTGCCGGCTTCGACTG | *cglR0802* targeting vector |
| 802RF-SalI | GAGTCGACTGCGCATGACTAATCCCGCT | *cglR0802* targeting vector |
| 802RR-SphI | GAGCATGCCTGTGGCAGCGCCTCAAGGT | *cglR0802* targeting vector |
| 2070FF-SmaI | GACCCGGGCAACGTTGAAAGGTAGCTCA | *cglR2070* targeting vector |

Table. 2 *continued*

| Name | Sequence (5'-3') | Application |
|---|---|---|
| 2070FR-XbaI | GATCTAGACAACGTTCCAGGAATCGATC | *cglR2070* targeting vector |
| 2070RF-SalI | GAGTCGACTACCCACTGATGAAACTCCA | *cglR2070* targeting vector |
| 2070RR-SphI | GAGCATGCCAGAAGACATCCGTGCGAAG | *cglR2070* targeting vector |
| SpF-LElox-XbaI | GATCTAGATACCGTTCGTATAGCATACATTATACGAAGTTATGAGGATCGATCTGTATAATA | Spectinomycine resistance cassette containing *lox71* for *cglR0802* and *cglR2070* targeting vector |
| SpR-RElox-XbaI | GATCTAGATACCGTTCGTATAATGTATGCTATACGAAGTTATCATATGGGATTCACCTTTAT | Spectinomycine resistance cassettecontaining *lox66* for *cglR0802* and *cglR2070* targeting vector |
| 20692070F-EcoRI | GAGAATTCAATCCTAGCTTTTTCCTTGT | *cglR2069-2070* targeting vector |
| 20692070R-XbaI | GATCTAGACCAATCCCACTGCCATACGC | *cglR2069-2070* targeting vector |
| SpF-LElox-BamHI | GAGGATCCTACCGTTCGTATAGCATACATTATACGAAGTTATGAGGATCGATCTGTATAATA | Spectinomycine resistance cassette containing *lox71* for *cglR2069-cglR2070* targeting vector |
| SpR-RElox-BAmHI | GAGGATCCTACCGTTCGTATAATGTATGCTATACGAAGTTATCATATGGGATTCACCTTTAT | Spectinomycine resistance cassettecontaining *lox66* for *cglR2069-cglR2070* targeting vector |

**Table 3.** Primers used for verification of deletion

| Primers | Sequence (5' - 3') | Start | End | Primers | Sequence (5' - 3') | Start | End |
|---|---|---|---|---|---|---|---|
| RD1-F | GGAGGACGCGTTTGGTGTCC | 2348890 | 2348909 | RD22-F | TTTGTGTTGCAAAGTAGTAA | 197954 | 197973 |
| RD1-R | CATTCAAGGAATTTTGAAGC | 2352028 | 2352009 | RD22-R | ACTCCCCAGAAATGTAATAG | 248860 | 248841 |
| RD2-F | GAAATCCTGGAAATGTACGC | 1939255 | 1939274 | RD23-F | GACGCCGGCCTCTCACGCCG | 1509421 | 1509440 |
| RD2-R | CGCCGGAATTTTGCTCTCCC | 1955157 | 1955138 | RD23-R | TCCTGGTGAATTCTGCAGTA | 1513202 | 1513183 |
| RD3-F | GGGGCATTAGCTCAGTTGGT | 44355 | 44374 | RD24-F | TTTAAGCAGTAGCGCCGGAA | 1510534 | 1510553 |
| RD3-R | GATCGAGTGGAGATTACTAG | 45994 | 45975 | RD24-R | GGCCCAGAAACCGACCCCAA | 1515191 | 1515172 |
| RD4-F | TTCTTAAGAATCTTTGTAGT | 44177 | 44196 | RD25-F | CGGCGTGTGAACAACTCTTC | 1024545 | 1024564 |
| RD4-R | AAATCCAGTTGCCGATGCGC | 52746 | 52727 | RD25-R | TCAAGCCCCTTAGACATCTT | 1027030 | 1027011 |
| RD5-F | CCATCGGTAACGGAGTAGTA | 3258109 | 3258128 | RD26-F | CCTTTTGAAGAATGAGCCTG | 107880 | 107899 |
| RD5-R | CTCGGGAGTTGCTTCGGCAG | 3274445 | 3274426 | RD26-R | GCACCGTGCGGGTAGACCCG | 162806 | 162787 |
| RD6-F | TGAGCCAAAAACCGACGGAA | 1384075 | 1384094 | RD27-F | TGTGAGGCGTGATGCGGAGC | 108935 | 108954 |
| RD6-R | ACCACTCGATGTTATTGGCG | 1398502 | 1398483 | RD27-R | GAAGGTGACACCAGCTGTTC | 184775 | 184756 |
| RD7-F | GATGTCTTAACCTTGGCCCA | 109614 | 109633 | RD28-F | GTAATTGGCAACACAACGGT | 1509548 | 1509567 |
| RD7-R | GTTGGTAGCAATGCCAGATC | 242792 | 242773 | RD28-R | AAGTAACTTCTGCACATTGG | 1510878 | 1510859 |
| RD8-F | ACAGCACAACCAACCTGTTG | 110385 | 110404 | RD29-F | GGTTTCAGCGTTCGACTTGC | 2703213 | 2703232 |
| RD8-R | AGGTGCTCTCCAGGGAGCCC | 112405 | 112386 | RD29-R | ATAGAGGGAAGCGACTGTTT | 2712986 | 2712967 |
| RD9-F | GGCTGTCCCTAGTACGAGAG | 110068 | 110087 | RD30-F | ACACGGGGAAGTGAAACATC | 107413 | 107432 |
| RD9-R | GGGAATTGTGGATACTTGCT | 237329 | 237310 | RD30-R | AAGCGGGACGAAGCAGGGAT | 117862 | 117843 |
| RD10-F | GCAACCCCATCACAGGCTTT | 3209667 | 3209686 | RD31-F | TGTCGTGAGATGTTGGGTTA | 106360 | 106379 |
| RD10-R | GTGACATAAACAAATGCCGC | 3210334 | 3210315 | RD31-R | CTTATTCCAAAAGCCGCTGG | 110868 | 110849 |
| RD11-F | AGCAGCTCTTCGCATAGATT | 28968 | 28987 | RD32-F | AGTGAGGCGGATCTAGGAAA | 104785 | 104804 |
| RD11-R | ATTCTAGAGCCGTCCCAACG | 29597 | 29578 | RD32-R | AATTTGGAAGATCTCGTCGA | 249240 | 249221 |
| RD12-F | ATCCGTTCATGTACCTTTTG | 38098 | 38117 | RD33-F | AGCGGCCTAGGACGCCGGCC | 1509411 | 1509430 |
| RD12-R | TTTTCCGTCGTGCATTCGAC | 44632 | 44613 | RD33-R | GCTTGCTCGGGGCGGAACGT | 1526415 | 1526396 |
| RD13-F | ATTATGATCAAATAGACCCA | 2694129 | 2694148 | RD34-F | AGAACTCCTCGCACCAGGTA | 2707056 | 2707075 |
| RD13-R | TTTTGATAATGCCAGTAACA | 2703406 | 2703387 | RD34-R | GAGTCAGCGCAACGAACCGA | 2717346 | 2717327 |
| RD14-F | ACCAGGATGACGCGTCGGAT | 2921139 | 2921158 | RD35-F | CTAGCAGCCACGTTCGGCCA | 3161869 | 3161888 |
| RD14-R | ACCAATTCCAAAGCTGGGAT | 2921883 | 2921864 | RD35-R | GCAATGATGTGGACAGTGAG | 3177655 | 3177636 |
| RD15-F | GGACCGGTTAGCCGTAAGGC | 109279 | 109298 | RD36-F | CCCACCAAGCAACTCCTCCC | 3171534 | 3171553 |
| RD15-R | ATGGATCGCCATTACCAACC | 152680 | 152661 | RD36-R | GCAATGATGTGGACAGTGAG | 3177655 | 3177636 |
| RD16-F | GATTGGGACGAAGTCGTAAC | 106772 | 106791 | RD37-F | TAGCCGTAAGCCGACCAATG | 3192934 | 3192953 |
| RD16-R | GGTAGATGAATAAGATCTTC | 292914 | 292895 | RD37-R | GACTAGATCATTTTCTGGAT | 3194682 | 3194663 |
| RD17-F | AGCGGGAGATGTAGCCAAAT | 933378 | 933397 | RD38-F | ATGCCCTCTTGACCTGCGGG | 104682 | 104701 |
| RD17-R | GGTTCTGAATCGTATGTTTA | 935132 | 935113 | RD38-R | ATTTATTCTGAGCTGGTCAT | 209072 | 209053 |
| RD18-F | CCATTAGACTGTGAGACCCG | 2501004 | 2501023 | RD39-F | CAGACGCGCTAGCTTGAGGT | 1752002 | 1752021 |
| RD18-R | TTCAATTCAGTTGTGAATTT | 2535110 | 2535091 | RD39-R | TAGGAGCCATTCCTGGCAGC | 1753178 | 1753159 |
| RD19-F | GCTAGGGAAACAGTGATACC | 3155487 | 3155506 | RD40-F | GGCTGTCCCTAGTACGAGAG | 110068 | 110087 |
| RD19-R | GTAAGGTTGAACAAGCTGCC | 3177547 | 3177528 | RD40-R | ATTTCCCAGATAGAGCACGG | 198368 | 198349 |
| RD20-F | GACGCCGGCCTCTCACGCCG | 1509421 | 1509440 | RD41-F | AGGGCTTGGTTATTCGTATC | 1774934 | 1774953 |
| RD20-R | GACAAAAATGCACTAGTAGT | 1511830 | 1511811 | RD41-R | AACTGCGCGGGTGAACCCCA | 1765012 | 1764993 |
| RD21-F | GCCCACCAGGAACAAAAAAG | 20540 | 20559 | RD42-F | ACTGGTACGCAACAAGGCGC | 2127176 | 2127195 |
| RD21-R | CTACCTCCACCACTTCAGGT | 45241 | 45222 | RD42-R | GAAGGCTCGCGAGGACAAGG | 2128740 | 2128721 |

**Table 4.** Features of deletion strains

| No. | Start | End | Deleted length | Deleted ORF | Number of deleted genes | Growth on M. M. |
|---|---|---|---|---|---|---|
| RD1 | 2,348,990 | 2,351,909 | 2,918 | CglR2128 → CglR2130 | 3 | + |
| RD2 | 1,939,355 | 1,955,038 | 15,682 | CglR1747 → CglR1759 | 13 | + |
| RD3 | 44,455 | 45,875 | 1,419 | CglR0038 → CglR0039 | 2 | + |
| RD4 | 44,277 | 52,627 | 8,349 | CglR0038 → CglR0045 | 8 | + |
| RD5 | 3,258,209 | 3,274,326 | 16,116 | CglR2939 → CglR2950 | 12 | + |
| RD6 | 1,384,175 | 1,398,383 | 14,207 | CglR1261 → CglR1271 | 11 | – |
| RD7 | 109,714 | 242,673 | 132,958 | CglR0095 → CglR0225 | 131 | – |
| RD8 | 110,485 | 112,286 | 1,800 | CglR0095 → CglR0095 | 1 | + |
| RD9 | 110,168 | 237,210 | 127,041 | CglR0095 → CglR0219 | 125 | – |
| RD10 | 3,209,767 | 3,210,215 | 447 | CglR2893 → CglR2893 | 1 | + |
| RD11 | 29,068 | 29,478 | 409 | – | 0 | + |
| RD12 | 38,198 | 44,513 | 6,314 | CglR0033 → CglR0038 | 6 | + |
| RD13 | 2,694,229 | 2,703,306 | 9,076 | CglR2452 → CglR2453 | 2 | + |
| RD14 | 2,921,239 | 2,921,783 | 543 | CglR2638 → CglR2638 | 1 | + |
| RD15 | 109,379 | 152,561 | 43,181 | CglR0095 → CglR0132 | 38 | + |
| RD16 | 106,872 | 292,795 | 185,922 | CglR0095 → CglR0265 | 171 | – |
| RD17 | 933,478 | 935,032 | 1,553 | CglR0846 → CglR0847 | 2 | + |
| RD18 | 2,501,104 | 2,535,010 | 33,905 | CglR2270 → CglR2300 | 31 | + |
| RD19 | 3,155,587 | 3,177,447 | 21,859 | CglR2842 → CglR2858 | 17 | – |
| RD20 | 1,509,521 | 1,511,730 | 2,208 | CglR1378 → CglR1378 | 1 | + |
| RD21 | 20,640 | 45,141 | 24,500 | CglR0019 → CglR0038 | 20 | + |
| RD22 | 198,054 | 248,760 | 50,705 | CglR0179 → CglR0231 | 53 | – |
| RD23 | 1,509,521 | 1,513,102 | 3,580 | CglR1378 → CglR1379 | 2 | + |
| RD24 | 1,510,634 | 1,515,091 | 4,456 | CglR1379 → CglR1381 | 3 | + |
| RD25 | 1,024,645 | 1,026,930 | 2,284 | CglR0928 → CglR0928 | 1 | + |
| RD26 | 107,980 | 162,706 | 54,725 | CglR0095 → CglR0145 | 51 | + |
| RD27 | 109,035 | 184,675 | 75,639 | CglR0095 → CglR0169 | 75 | + |
| RD28 | 1,509,648 | 1,510,778 | 1,129 | CglR1378 → CglR1378 | 1 | + |
| RD29 | 2,703,313 | 2,712,886 | 9,572 | CglR2454 → CglR2461 | 8 | + |
| RD30 | 107,513 | 117,762 | 10,248 | CglR0095 → CglR0096 | 2 | + |
| RD31 | 106,460 | 110,768 | 4,307 | – | 0 | + |
| RD32 | 104,885 | 249,140 | 144,254 | CglR0095 → CglR0231 | 137 | – |
| RD33 | 1,509,511 | 1,526,315 | 16,803 | CglR1378 → CglR1391 | 14 | – |
| RD34 | 2,707,156 | 2,717,246 | 10,089 | CglR2456 → CglR2461 | 6 | + |
| RD35 | 3,161,969 | 3,177,555 | 15,585 | CglR2851 → CglR2858 | 8 | + |
| RD36 | 3,171,634 | 3,177,555 | 5,920 | – | 0 | + |
| RD37 | 3,193,034 | 3,194,582 | 1,547 | CglR2877 → CglR2878 | 2 | + |
| RD38 | 104,782 | 208,972 | 104,189 | CglR0095 → CglR0190 | 96 | + |
| RD39 | 1,752,102 | 1,753,078 | 975 | CglR1584 → CglR1584 | 1 | + |
| RD40 | 110,168 | 198,268 | 88,099 | CglR0095 → CglR0179 | 85 | + |
| RD41 | 1,775,053 | 1,764,912 | 10,140 | CglR1595 → CglR1604 | 10 | + |
| RD42 | 2,127,295 | 2,128,621 | 1,325 | CglR1938 → CglR1940 | 3 | + |

M.M, minimal medium

**Table 5.** Growth rate of deletion strains

| No. | Growth rate | No. | Growth rate | No. | Growth rate |
|---|---|---|---|---|---|
| W.T. | 1.00 | RD15 | 0.38 | RD30 | 0.97 |
| RD1 | 1.09 | RD16 | 0.2* | RD31 | 0.82 |
| RD2 | 0.60 | RD17 | 0.80 | RD32 | 0.59* |
| RD3 | 1.04 | RD18 | 0.92 | RD33 | 0.99* |
| RD4 | 0.65 | RD19 | 0.58* | RD34 | 0.78 |
| RD5 | 0.97 | RD20 | 0.83 | RD35 | 0.70 |
| RD6 | 0.98* | RD21 | 0.85 | RD36 | 0.72 |
| RD7 | 0.78* | RD22 | 0.87* | RD37 | 0.71 |
| RD8 | 0.94 | RD23 | 0.87 | RD38 | 0.25 |
| RD9 | 0.83* | RD24 | 0.67 | RD39 | 0.68 |
| RD10 | 0.96 | RD25 | 0.83 | RD40 | 0.35 |
| RD11 | 0.93 | RD26 | 0.24 | RD41 | 0.88 |
| RD12 | 0.86 | RD27 | 0.30 | RD42 | 0.92 |
| RD13 | 0.82 | RD28 | 0.98 | | |
| RD14 | 0.88 | RD29 | 0.94 | | |

*Strains which did not grow on minimal medium were tested using A medium.

**Table 6.** Deleted genes in deletion strains

| Gene name | Annotation | Ortholog gene | |
|---|---|---|---|
| | | *E. coli* | *B. subtilis* |
| CglR0019 | hypothetical membrane protein | b2158 | |
| CglR0020 | bacterial regulatory protein, LysR family | b2157 | BG10635 |
| CglR0021 | hypothetical protein | | |
| CglR0022 | putative integral membrane cytochrome biogenesis | | |
| CglR0023 | HCCA isomerase, secreted protein | | BG14104 |
| CglR0024 | regulatory protein, MarR family | | |
| CglR0025 | protease with chaperone function | | |
| CglR0026 | 5'-nucleotidase (putative pseudogene) | | |
| CglR0027 | putative membrane protein | | |
| CglR0028 | putative glycosyltransferase | | |
| CglR0029 | putative polysaccharide biosynthesis protein | | BG10778 |
| CglR0030 | putative polysaccharide deacetylase | b0130 | BG12544 |
| CglR0031 | hypothetical protein | | |
| CglR0032 | hypothetical protein | | |
| CglR0033 | glycosyltransferase | | |
| CglR0034 | probable glycosyltransferase | | |
| CglR0035 | glycosyltransferase | b2047 | BG11871 |
| CglR0036 | probable glycosyltransferase | | |
| CglR0037 | probable glycosyl hydrolase | | |
| CglR0038 | organic hydroperoxide resistance protein | b1482 | BG19021 |
| CglR0039 | putative transcriptional regulator | | |
| CglR0040 | putative secreted protein | | |
| CglR0041 | putative solute-binding lipoprotein, signal peptid | b1857 | BG13851 |
| CglR0042 | ABC transporter protein, integral membrane subunit | b1859 | BG12765 |
| CglR0043 | ABC transport protein, ATP-binding subunit | | |
| CglR0044 | probable solute-binding lipoprotein, signal peptide | b2548 | |
| CglR0045 | probable ABC transport protein, membrane component | | |
| CglR0095 | permease of the major facilitator superfamily | | |
| CglR0096 | creatinine deaminase | b0337 | |
| CglR0097 | secreted protein | | |
| CglR0098 | putative SIR2-like regulatory protein | | |
| CglR0099 | triacylglycerol lipase precursor | | BG11951 |
| CglR0100 | triacylglycerol lipase precursor | | |
| CglR0101 | hypothetical protein | | |
| CglR0102 | bacterial regulatory protein, MarR family | | |
| CglR0103 | probable urease gamma subunit | | BG11981 |
| CglR0104 | urease beta subunit | | BG11982 |
| CglR0105 | urease alpha subunit | | BG11983 |
| CglR0106 | urease accessory protein | | |
| CglR0107 | urease accessory protein | | |
| CglR0108 | urease accessory protein | b2727 | |
| CglR0109 | urease accessory protein | | |
| CglR0110 | permease of the major facilitator superfamily | | |
| CglR0111 | putative glycerol 3-phosphate dehydrogenase | | |
| CglR0112 | putative heat shock protein (HSP90-family) | b0473 | BG11359 |
| CglR0113 | AMP nucleosidase | b1982 | |
| CglR0114 | putative glutathione-dependent aldehyde dehydrogenase | b0608 | BG11902 |
| CglR0115 | ribose operon repressor | b3753 | BG13211 |
| CglR0116 | histidinol dehydrogenase | | |
| CglR0117 | H+/gluconate symporter and related permeases | | BG11152 |
| CglR0118 | putative transposase, ISCg15b | | |
| CglR0119 | hypothetical protein | | |
| CglR0120 | secreted multicopper oxidase | | |
| CglR0121 | putative secreted protein | | |
| CglR0122 | putative two component response regulator | | |
| CglR0123 | probable two component sensor kinase | b2078 | |
| CglR0124 | cation transport ATPase | | |
| CglR0125 | probable cation-transporting ATPase transmembrane protein | | |
| CglR0126 | hypothetical protein | | BG12188 |
| CglR0127 | putative bacterial regulatory proteins, AsnC family | | |
| CglR0128 | putative transposase | | |
| CglR0129 | cation transport ATPase | | |
| CglR0130 | hypothetical protein | | |
| CglR0131 | hypothetical protein | | |
| CglR0132 | copper chaperone | | |
| CglR0133 | bacterial regulatory protein, Crp family | | |
| CglR0134 | hypothetical protein | | |
| CglR0135 | putative DNA invertase | | |
| CglR0136 | putative transposaset, ISCg11a | | |
| CglR0137 | hypothetical protein | | |
| CglR0138 | hypothetical protein | | |
| CglR0139 | putative transcriptional regulator | | |
| CglR0140 | hypothetical protein | | |
| CglR0141 | putative transposase | | |
| CglR0142 | putative transposase | | |
| CglR0143 | putative copper resistance protein D / membrane protein | | BG12046 |
| CglR0144 | hypothetical protein | b1841 | |
| CglR0145 | putative transcription regulator | | |
| CglR0146 | putative cation-transporting ATPase | | |
| CglR0147 | hypothetical protein | | |
| CglR0148 | cadmium translocating P-type ATPase | b3469 | BG13325 |
| CglR0149 | transposase-fragment, ISCg11a | | |
| CglR0150 | putative transposase | | |
| CglR0151 | putative transposase | | |
| CglR0152 | putative transposase | | |
| CglR0153 | transposase, ISCg2b | b0256 | |
| CglR0154 | putative transposase | | |
| CglR0155 | cadmium resistance transporter | | |
| CglR0156 | bacterial regulatory protein, ArsR family | | |
| CglR0157 | Iron dependent repressor | b0817 | |
| CglR0158 | probable manganese transport protein | b2392 | BG12065 |
| CglR0159 | probable DNA invertase | | BG10458 |
| CglR0160 | probable DNA invertase | | |
| CglR0161 | bacterial regulatory protein, ArsR family | | |
| CglR0162 | cadmium resistance transporter | | |
| CglR0163 | mercuric reductase | | |
| CglR0164 | transcriptional regulator, MerR family | b3292 | BG13777 |
| CglR0165 | putative transcriptional regulator | | |
| CglR0166 | putative 2,5-diketo-D-gluconic acid reductasetase, fragment | | |
| CglR0167 | putative 2,5-diketo-D-gluconic acid reductasetase, fragment | | |
| CglR0168 | hypothetical protein | | |
| CglR0169 | 5-methylcytosine-specific restriction enzyme B | b4346 | BG12791 |
| CglR0170 | putative protein mcrC | b4345 | |
| CglR0171 | putative transposase, ISCg15b | | |
| CglR0172 | secreted protein, signal peptide | | BG13255 |
| CglR0173 | proline dehydrogenase/delta-1-pyrroline-5-carboxylatedehydrogenase | b1014 | |
| CglR0174 | putative oxidoreductase | b0419 | BG11363 |
| CglR0175 | p-aminobenzoyl-glutamate transporter | b1336 | BG11999 |
| CglR0176 | hydrolase, Ama/HipO/HyuC family | b1337 | |
| CglR0177 | putative inner membrane protein | b3676 | |
| CglR0178 | putative membrane protein | | |
| CglR0179 | hypothetical protein | | |
| CglR0180 | bacterial regulatory proteins, DeoR family | | |
| CglR0181 | hypothetical protein | | |
| CglR0182 | predicted hydrolase of the HAD family | | |
| CglR0183 | glyoxalase/bleomycin resistance protein/dioxygenas | b1651 | BG14006 |
| CglR0184 | phosphohistidine phosphatase | | |
| CglR0185 | mannitol 2-dehydrogenase | b2172 | |
| CglR0186 | putative ribitol transporter | | BG11922 |
| CglR0187 | bacterial regulatory proteins, DeoR family | | |
| CglR0188 | xylulose kinase | b3564 | BG10807 |
| CglR0189 | pantoate--beta-alanine ligase | b0133 | BG11520 |
| CglR0190 | 3-methyl-2-oxobutanoate hydroxymethyltransferase | b0134 | BG11519 |
| CglR0191 | bacterial regulatory protein | | |
| CglR0192 | putative 3-methylpurine DNA glycosylase | | BG12555 |
| CglR0193 | hypothetical protein | | |
| CglR0194 | probable esterase/lipase protein | b0476 | |
| CglR0195 | haloacid dehalogenase-like hydrolase | b3885 | |
| CglR0196 | putative acetyltransferase | b3279 | BG13896 |
| CglR0197 | bacterial regulatory proteins, Crp family | | |
| CglR0198 | putative membrane transport protein | b0591 | BG12895 |
| CglR0199 | hypothetical protein | | |
| CglR0200 | hypothetical protein | | |
| CglR0201 | putative secreted or membrane protein | | |
| CglR0202 | membrane spanning protein | | |
| CglR0203 | N-acetylglucosaminyltransferase | | BG12060 |
| CglR0204 | ABC-2 type transporter | | |
| CglR0205 | repeat containing protein | | |
| CglR0206 | transcriptional regulator, DeoR family | b2735 | |
| CglR0207 | 2-dehydro-3-deoxyphosphogluconate aldolase / 4-hydroxy-2-oxoglutarate aldolase | b1850 | BG11396 |
| CglR0208 | hypothetical transport protein | b4356 | BG11160 |
| CglR0209 | hypothetical protein | | |
| CglR0210 | permease | | |
| CglR0211 | permease | | |
| CglR0212 | probable transmembrane protein | | |

**Table 6.** *Continued*

| Gene name | Annotation | Ortholog gene | |
|---|---|---|---|
| | | *E. coli* | *B. subtilis* |
| CglR0213 | secreted protein | | |
| CglR0214 | aspartate 1-decarboxylase | b0131 | BG11493 |
| CglR0215 | hypothetical protein | | |
| CglR0216 | putative transport protein | b0715 | BG13080 |
| CglR0217 | secreted protein, signal peptide | | |
| CglR0218 | permease | b2327 | |
| CglR0219 | hypothetical protein | | |
| CglR0220 | probable ATP-dependent RNA helicase protein | b0148 | |
| CglR0221 | hypothetical protein | | |
| CglR0222 | maltose O-acetyltransferase | b0459 | BG10043 |
| CglR0223 | putative DNA repair protein | b2212 | |
| CglR0224 | probable DNA-3-methyladenine glycosylase I protein | | |
| CglR0225 | putative LysE type translocator | b0328 | |
| CglR0226 | hypothetical protein | | BG10067 |
| CglR0227 | glyoxalase/bleomycin resistance protein/Dioxygenas | | |
| CglR0228 | methylated-DNA--protein-cysteine methyltransferase | | |
| CglR0229 | hypothetical protein | | |
| CglR0230 | hypothetical protein | | |
| CglR0231 | translation initiation inhibitor | b1010 | |
| CglR0232 | hypothetical protein | | |
| CglR0233 | endopeptidase O | | |
| CglR0234 | hypothetical secreted protein | | |
| CglR0235 | membrane protein | | |
| CglR0236 | bacterial regulatory proteins, GntR family | b2101 | BG14124 |
| CglR0237 | sugar kinase, ribokinase family | | BG11119 |
| CglR0238 | hypothetical protein | | |
| CglR0239 | methylmalonate-semialdehyde dehydrogenase | | BG11117 |
| CglR0240 | enzyme involved in inositol metabolism | | BG11118 |
| CglR0241 | putative acetolactate synthase protein | | BG11120 |
| CglR0242 | phosphate isomerases/epimerase | | BG11121 |
| CglR0243 | putative oxidoreductase myo-inositol 2-dehydrogenase | b1068 | BG10669 |
| CglR0244 | myo-inositol catabolism protein | | BG11123 |
| CglR0245 | probable transporter | b3754 | |
| CglR0246 | hypothetical oxidoreductase | b1624 | BG14057 |
| CglR0247 | hypothetical protein | | |
| CglR0248 | hypothetical protein | | |
| CglR0249 | LacI-family transcriptional regulatory protein | | |
| CglR0250 | putative oxidoreductase | b1315 | BG13076 |
| CglR0251 | phosphate isomerases/epimerase | | BG11855 |
| CglR0252 | cold-shock protein CspA | | BG10824 |
| CglR0253 | putative ABC transporter permease protein | | |
| CglR0254 | hypothetical ABC transporter periplasmic solute-binding protein | | |
| CglR0255 | ABC-type sugar transport systems, ATPase components | | |
| CglR0256 | putative hydrolase | | |
| CglR0257 | transcriptional regulator | | BG10847 |
| CglR0258 | O-methyl transferase | | BG13794 |
| CglR0259 | acetyltransferase, GNAT family | | |
| CglR0260 | probable LacI-family transcriptional regulator | | |
| CglR0261 | metabolite transport protein | b2943 | BG12802 |
| CglR0262 | sensor histidine kinase of two-component system, fragment | | |
| CglR0263 | glutamine 2-oxoglutarate aminotransferase large SU | b3212 | BG10811 |
| CglR0264 | glutamine 2-oxoglutarate aminotransferase | b3213 | BG12594 |
| CglR0265 | hypothetical protein | | |
| CglR0323 | 2-isopropylmalate synthase | b0074 | BG11948 |
| CglR0846 | bacterial extracellular solute-binding protein, family | | BG11910 |
| CglR0847 | ABC-type sugar transport systems, ATPase component | | |
| CglR0928 | probable pyridoxal phosphate aminotransferase | b0600 | BG12362 |
| CglR1261 | hypothetical protein | | |
| CglR1262 | homoserine dehydrogenase | | BG10460 |
| CglR1263 | homoserine kinase | b0003 | BG10462 |
| CglR1264 | hypothetical protein | | |
| CglR1265 | respiratory nitrate reductasetase 2 gamma chain | b1465 | BG11084 |
| CglR1266 | nitrate reductasetase delta chain | b1226 | BG11083 |
| CglR1267 | probable respiratory nitrate reductasetase oxidoreductase | b1467 | BG11082 |
| CglR1268 | nitrate reductasetase 2, alpha subunit | b1224 | BG11081 |
| CglR1269 | putative nitrate/nitrite transporter | b1223 | BG11342 |
| CglR1270 | putative molybdopterin biosynthesis MOG protein | | |
| CglR1271 | secreted phospholipid phosphatase | | |
| CglR1378 | hypothetical protein | | |
| CglR1379 | putative ATP/GTP-binding protein | | |
| CglR1380 | hypothetical protein | | |
| CglR1381 | thiamine biosynthesis protein | b3994 | BG11246 |
| CglR1382 | putative glycogen phosphorylase | | |
| CglR1383 | putative glycogen phosphorylase | b3417 | BG10911 |
| CglR1384 | hypothetical protein | | |
| CglR1385 | Zn-dependent hydrolases, including glyoxylases | | |
| CglR1386 | putative membrane protein | | |
| CglR1387 | putative secreted protein | | |
| CglR1388 | putative metal dependent phosphohydrolase, RelA/SpoT homolog | | |
| CglR1389 | bacterial regulatory proteins, IclR family | | |
| CglR1390 | 3-isopropylmalate dehydratase large subunit | b0072 | BG11949 |
| CglR1391 | 3-isopropylmalate dehydratase (small subunit) | b0071 | BG11950 |
| CglR1584 | hypothetical protein | | |
| CglR1595 | ferrochelatase precursor | b0475 | BG10430 |
| CglR1596 | secreted cell wall-associated hydrolase (invasion-associated protein) | b0739 | BG11023 |
| CglR1597 | putative membrane protein | | |
| CglR1598 | aconitate hydratase | b1276 | BG10478 |
| CglR1599 | transcriptional regulator, TetR family | | BG14194 |
| CglR1600 | glutamine amidotransferase domain | | |
| CglR1601 | putative nucleoside-diphosphate-sugar epimerase | | |
| CglR1602 | putative membrane protein | | BG12551 |
| CglR1603 | ACT domain-containing protein | | |
| CglR1604 | hypothetical protein | | |
| CglR1747 | rRNA or tRNA methylase | b4371 | |
| CglR1748 | putative D-tyrosyl-tRNA(Tyr) deacylase | b3887 | BG13805 |
| CglR1749 | RNA polymerase sigma factor | | |
| CglR1750 | iron dependent regulatory protein-DtxR homolog | | |
| CglR1751 | UDP-glucose 4-epimerase | | |
| CglR1752 | hypothetical protein | | |
| CglR1753 | hypothetical protein | | |
| CglR1754 | superfamily II DNA or RNA helicase | | |
| CglR1755 | hydrogen peroxide sensing regulator | b3961 | |
| CglR1756 | putative membrane protein | | |
| CglR1757 | probable ATP-dependent RNA helicase protein | b1413 | |
| CglR1758 | putative transcriptional regulator | b0413 | BG13833 |
| CglR1759 | hypothetical protein | | |
| CglR1938 | putative secreted lipoprotein | | |
| CglR1939 | RimM protein (16S rRNA processing protein) | b2608 | BG13404 |
| CglR1940 | double-stranded beta-helix domain | | |
| CglR2128 | putative phosphatase in N-acetylglucosamine metabolism | b0675 | BG14042 |
| CglR2129 | ATPase component of ABC transporters with duplicated ATPase domains | | |
| CglR2130 | hypothetical protein | | |
| CglR2270 | hypothetical protein | | |
| CglR2271 | probable succinyl-CoA:3-ketoacid-coenzyme A transferase subunit | b2222 | BG11154 |
| CglR2272 | probable fesuccinyl-CoA:3-ketoacid-coenzyme A transferase subunit | b2221 | BG11153 |
| CglR2273 | bacterial regulatory proteins, IclR family | b4018 | |
| CglR2274 | putative acetyl-CoA:acetyltransferase | | |
| CglR2275 | 3-oxoadipate enol-lactone hydrolase /4-carboxymuconolactonedecarboxylase | | BG11957 |
| CglR2276 | ATP-dependent transcriptional regulator, LuxR family | b3418 | |
| CglR2277 | 4-carboxymuconolactone decarboxylase | | |
| CglR2278 | 3-carboxy-cis,cis-muconate cycloisomerase | | |
| CglR2279 | protocatechuate dioxygenase alpha subunit | | |
| CglR2280 | protocatechuate dioxygenase beta subunit | | |
| CglR2281 | putative restriction endonuclease | | |
| CglR2282 | muconolactone isomerase | | |
| CglR2283 | chloromuconate cycloisomerase | b3692 | BG13232 |
| CglR2284 | catechol 1,2-dioxygenase | | |
| CglR2285 | benzoate 1,2-dioxygenase alpha subunit (aromatic ring hydroxylation dioxygenase A) | | |
| CglR2286 | benzoate dioxygenase small subunit | b2539 | |
| CglR2287 | benzoate dioxygenase reductasetase | b3924 | |
| CglR2288 | benzoate diol dehydrogenase bend | | |
| CglR2289 | bacterial regulatory protein, LuxR family | | |
| CglR2290 | putative benzoate transport protein | | |
| CglR2291 | benzoate membrane transport protein | b1433 | |

**Table 6.** *Continued*

| Gene name | Annotation | Ortholog gene | |
|---|---|---|---|
| | | *E. coli* | *B. subtilis* |
| CglR2292 | putative two-component system sensor kinase | | |
| CglR2293 | hypothetical protein | | |
| CglR2294 | hypothetical protein | | |
| CglR2295 | putative ring-cleavage dioxygenase large subunit | b2538 | |
| CglR2296 | putative ring-cleavage dioxygenase small subunit | | |
| CglR2297 | 3-oxoacyl-(acyl-carrier protein) reductase | b1093 | |
| CglR2298 | putative vanillate O-demethylase oxidoreductase | b1803 | |
| CglR2299 | putative two-component system response regulator | | |
| CglR2300 | hypothetical protein | | |
| CglR2452 | ABC-type transport system, involved in lipoprotein release, permease component | | |
| CglR2453 | ABC-type transport system, involved in lipoprotein release, ATPase component | | |
| CglR2554 | neuraminidase NANP | | |
| CglR2555 | bacterial regulatory proteins, GntR family | | |
| CglR2556 | ABC-type dipeptide/oligopeptide/nickel transport system, secreted component | | |
| CglR2557 | ABC-type dipeptide/oligopeptide/nickel transport system, permease component | b1486 | |
| CglR2558 | ABC-type dipeptide/oligopeptide/nickel transport system, fused permease and ATPase components | b1246 | BG10774 |
| CglR2559 | ATPase components of ABC-type transport system, contain duplicated ATPase domains | | |
| CglR2560 | LysE type translocator | b1798 | BG12304 |
| CglR2561 | bacterial regulatory proteins, AsnC family | | |
| CglR2638 | predicted hydrolase or acyltransferase (alpha/beta hydrolase superfamily) | | |
| CglR2842 | permease of the major facilitator superfamily | | |
| CglR2843 | bacterial regulatory protein, MarR family | | BG11106 |
| CglR2844 | putative two component response regulator | b2193 | BG14133 |
| CglR2845 | probable two component sensor kinase | b1222 | |
| CglR2846 | putative secreted protein | | |
| CglR2847 | putative membrane protein | | |
| CglR2848 | putative sortase (surface protein transpeptidase) | | |
| CglR2849 | putative preprotein translocase subunit YidC, SpoIIIJ homolog | | |
| CglR2850 | TetR-type transcriptional regulator of sulfur metabolism | | |
| CglR2851 | putative membrane protein | b1795 | BG12066 |
| CglR2852 | universal stress protein family | b1376 | BG11134 |
| CglR2853 | alkanal monooxygenase alpha chain | | |
| CglR2854 | hypothetical protein | | |
| CglR2855 | putative ribosomal pseudouridine synthase | | |
| CglR2856 | putative glutamyl-tRNA(Gln) amidotransferase subunit A | | |
| CglR2857 | putative regulatory protein | | |
| CglR2858 | putative transcriptional regulator | | |
| CglR2877 | bacterial regulatory protein, MarR family | | BG14027 |
| CglR2878 | universal stress protein UspA or related nucleotide-binding protein | | |
| CglR2893 | putative ABC-type cobalamin/Fe3+-siderophores transport systems, periplasmic components | | |
| CglR2939 | dipeptide/tripeptide permease | b1634 | BG12027 |
| CglR2940 | bacterial regulatory protein, TetR family | b0464 | |
| CglR2941 | catechol 1,2-dioxygenase | | |
| CglR2942 | maleylacetate reductasetase | b2799 | BG11941 |
| CglR2943 | permease of the major facilitator superfamily | | |
| CglR2944 | bacterial regulatory proteins, IclR family | | BG11214 |
| CglR2945 | predicted dehydrogenase | | |
| CglR2946 | sugar phosphate isomerase/epimerase | | |
| CglR2947 | myo-inositol 2-dehydrogenase | | BG12280 |
| CglR2948 | myo-inositol 2-dehydrogenase | | |
| CglR2949 | putative secreted phosphoesterase | | |
| CglR2950 | hypothetical protein | | BG14123 |