J. Microbiol. Biotechnol. (2003), **13**(5), 783–788

JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY

Genome Diversification by Phage-Derived Genomic Islands in *Pseudomonas aeruginosa*

KIM, SEOL-HEE, KYOUNG-BOON LEE, JI-SUN LEE, AND YOU-HEE CHO*

Department of Life Science, Sogang University, Seoul 121-742, Korea

Received: June 28, 2003
Accepted: July 29, 2003

Abstract A 27 bp $tRNA^{Gly}$ region (*att1*) was identified as the integration site for a 12,384 bp Pf1-derived genomic island containing 15 open reading frames (ORFs) from PA0715 to PA0729 in *P. aeruginosa* strain PAO1. Homologous island was observed in *P. aeruginosa* strain PA14, but not in *P. aeruginosa* strain K (PAK). We isolated the Pf1 island from PA14, and determined its 10,657 bp sequences containing 14 ORFs, with significant sequence variations near the borders. In contrast to the PAO1 Pf1 island, the PA14 Pf1 island was integrated at the 10 bp *att2* site between PA1191 and PA1192. The *att1* site of PA14, however, was still occupied by a third genetic segment, whereas both *att1* and *att2* sites of PAK remained unutilized. These results exemplify an extensive genomic variation of Pf1-related islands involving differential genetic organizations and differential *att* site utilizations.

Key words: Genome plasticity, genomic island, *Pseudomonas aeruginosa*, Pf1 phage, *att* site

Genetic variability is primarily the result of sexual reproduction, which involves chromosomal recombination during meiosis in eukaryotes. In prokaryotes, however, no such shuffling strategy is available, and other factors play important roles in genetic variability, determining the rate of evolution. These include the frequent occurrence of point mutations, high levels of recombination and gene silencing, and the transfer of genetic material between different species, even genera. In particular, the latter process, referred to as horizontal gene transfer, represents a basis of prokaryotic evolution and diversification, and it has led to dramatic changes in the composition of microbial genomes over a relatively short period of time [16].

The significant impact of horizontal gene transfer has become more apparent from recent analyses based on the growing pool of microbial genome sequence information, differentiating the genomic islands from the core genome by marked differences in their G+C content and codon usage [9]. It is also important to take note that many highly expressed genes encoding ribosomal proteins, translation and transcription processing factors, and chaperone and degradation complexes also show significant differences from the average codon usage, probably for the maximal or constitutive expressions [11]. Therefore, if the codon usage of a gene substantially deviates from those of both average genes and highly expressed genes, it is defined as a putative alien (pA), the clusters of which are relevant for *in silico* detection of genomic islands that are flanked by two separate direct repeats [10]. The repeat sequence that can be used as the attachment (*att*) site, in many cases, is a part of a tRNA gene [5].

Pseudomonas aeruginosa is a Gram-negative rod ubiquitously distributed in aquatic and soil habitats. It is a common respiratory tract pathogen that is innocuous in healthy individuals, but can cause serious infection in immunocompromised patients or in patients with cystic fibrosis. The pathophysiology of infections due to *P. aeruginosa* is quite complex, as shown by the clinical diversity of the diseases associated with this phenotypically and genetically diverse species, which is mainly attributed to the genome diversification by horizontal gene transfers [12]. Therefore, the studies conducted on genetic variability involving genomic islands in *P. aeruginosa* isolates may have a considerable impact on the diagnosis and control of this opportunistic pathogen.

Here, we defined the borders of a pA region in *P. aeruginosa* strain PAO1, which is known as a prophage derived from a filamentous single-stranded DNA phage, Pf1 (Pf1 island hereafter). By cloning and sequencing the homologous region from *P. aeruginosa* strain PA14, we found several differences between both Pf1 islands; genetic make-ups, attachment (*att*) site specificity, and chromosomal locations. We propose that the genome

*Corresponding author
Phone: 82-2-705-8793; Fax: 82-2-704-3601;
E-mail: youhee@sogang.ac.kr

diversification exploiting Pf1-related islands in *P. aeruginosa* might be one of the examples of genome evolutions to facilitate its adaptations to the adverse environmental changes that are provided by its natural habitats as well as its hosts.

MATERIALS AND METHODS

Bacterial Strains and Culture Conditions

We used *Escherichia coli* DH5α [18] for general purpose cloning and three *P. aeruginosa* strains, PAO1, PAK, and PA14 [17]. Both *E. coli* and *P. aeruginosa* were grown at 37°C using Luria-Bertani (LB) medium. Antibiotics, rifampin (100 μg/ml) for PA14 and carbenicillin (50 μg/ml) for *E. coli*, were supplemented, if necessary.

DNA Manipulations

Restriction and modifying enzymes were used according to the manufacturer's recommendations (KOSCHEM, Promega, and TaKaRa). Standard methods for DNA extractions, purifications, and elutions were used, as described elsewhere [4, 18].

Preparation of the Pf1 Island Probe

Oligonucleotide primers for the PA14 Pf1 island were designed based on the sequence of the PAO1 Pf1 island. The primer pairs (718D, 5'-CAC GGA ATT CGT GCC CGT GCG C-3' and C-out, 5'-TGG CGT TGC CAC GGG CAC CTT CC -3'; underlines indicate mutation for *Eco*RI site) were used under the following conditions: denaturation at 94°C for 3 min followed by 30 cycles of the amplification (denaturation at 94°C for 30 s, annealing at 44°C for 30 s, and extension at 72°C for 1 min) and the final extension at 72°C for 10 min. The 703 bp PCR fragment (Fig. 1) was cloned in pGEM-T Easy vector (Promega, Madison, WI, U.S.A.) and verified by sequencing. The insert was digested with *Eco*RI, gel-purified using Gel Extraction kit (Qiagen, Hilden, Germany), labeled using Rediprime II random priming labeling kit (Amersham-Pharmacia, Uppsala, Sweden), and used as a probe for Southern analysis and library screening.

Southern Analysis

Chromosomal DNA was prepared from 10 h cultures using a rapid method [3] with slight modifications. Harvested cells were disrupted promptly in 200 μl of lysis buffer [40 mM Tris-acetate (pH 7.8), 20 mM sodium acetate, 1 mM EDTA, 1% SDS]. Cell lysate was mixed with 66 μl of 5 M NaCl and 250 μl of chloroform, and the separated aqueous phase was subjected to ethanol precipitation. The chromosomal DNA (~3 μg) was digested with *Nru*I, and then separated on a 0.8% agarose gel. The gel was transferred to Hybond-N+ membranes (Amersham-Pharmacia) and subjected to hybridization under standard conditions [18], using the radiolabeled PA14 Pf1 probe. Signals were visualized using the phospho-image analyzer (Fuji, Tokyo, Japan).

Screening for the Pf1 Island in PA14 Genomic Libraries

Cosmid and plasmid genomic libraries of PA14 [17] were screened to isolate the Pf1 island, using the radiolabeled PA14 Pf1 probe. Among several primary candidates, the cosmid clone AC4 contained the entire Pf1 island. The 10,657 bp Pf1 island of PA14 and its flanking regions were sequenced and used for further analyses.

Nucleotide Sequencing and Sequence Analysis

Nucleotide sequencing was performed on double-stranded DNA fragments cloned in general cloning vectors, using an automatic sequencer (ALFexpress, Pharmacia). Sequence comparison with the database was carried out by BLAST programs [1], and multiple alignments were by CLUSTAL W [19]. GCG PACKAGE and OMIGA 2.0 (Oxford Molecular, SanDiego, CA, U.S.A.) were used for general analyses and manipulations of sequence data. Putative ORFs were identified with GeneMark [2]. tRNA genes were identified by tRNA-scan-SE [15].

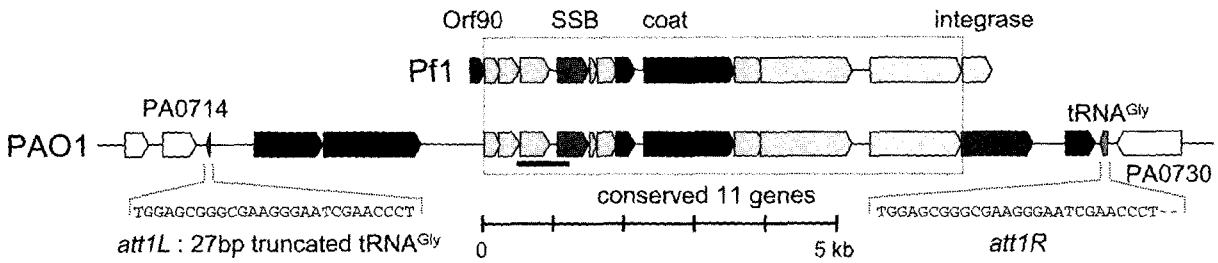

Fig. 1. Representation of a Pf1 phage-derived genomic island in PAO1.

Genetic organization of the PAO1 Pf1 island is shown in comparisons with that of the phage Pf1. Pf1 phage contains 13 ORFs and the 15 ORFs that have been incorporated at the specific location (*att1*) within the $tRNA^{Gly}$ gene between PA0714 and PA0730 are indicated. The 11 common ORFs, including SSB (single-strand binding protein, PA0720) and coat proteins (CoaB, PA0723 and CoaA, PA0724), are indicated by the dotted box. Both borders are shown as *att1L* and *att1R* which is the 27 bp 3'-end of the $tRNA^{Gly}$ gene. The solid line denotes the 703 bp PCR fragment that was designed for Southern hybridization in Fig. 2.

PCR Amplification of the Pf1 Island Attachment (*att*) Sites

The primer pairs (714D, 5'-CAT GAG GTC TAT CTG GAA CC-3' and 730U, 5'-TCG ACG TCG AGC ACA AGC CC-3') were used to amplify the *att1* site contained in the $tRNA^{Gly}$ gene between PA0714 and PA0730, and the primer pairs (1190D, 5'-CAT CGC CAG GTA GGT GAG GAC CG-3' and 1192U, 5'-CTG GGA ATA GGC CTC GAT GTC C-3') were used for the *att2* site between PA1191 and PA1192. PCR amplifications under standard conditions generated a 759 bp (*att1*) and a 1,676 bp (*att2*) fragment, unless the corresponding *att* sites were uninterrupted. The nucleotide sequence of 719 bp *att1* region of PAK was determined.

Nucleotide Sequence Accession Numbers

The nucleotide sequences of the PA14 Pf1 island and the PAK *att1* region have been submitted to the DDBJ/EMBL/GenBank databases under the accession numbers AY324828 and AY324832, respectively.

RESULTS AND DISCUSSION

Identification of a Genomic Island Associated with a Pf1-Like Prophage in *P. aeruginosa* Strain PAO1

An attempt to search for putative aliens (pAs) in *P. aeruginosa* reported in the presence of two cryptic regions of bacteriophage genes, based on the differences of overall G+C content and codon usage [10]. One of them (named Pf1 island) contains phage Pf1-derived genes. Pf1 phage was isolated more than 40 years ago and known to only infect a single host *P. aeruginosa* strain K (PAK) via the polar sex pili that PAK displays [6]. Since some of these genes are highly expressed during biofilm formation [21], we selected this pA region for genome plasticity studies, in the hope that its expressions and biochemical functions might be associated with the extreme survival capabilities of *P. aeruginosa*.

First, we examined whether this pA region might have some canonical features of genomic islands, such as flanked by direct repeats contained in a tRNA gene. As shown in Fig. 1, we identified the nearest tRNA gene ($tRNA^{Gly}$) between two open reading frames (ORFs), PA0729 and PA0730, as the right border (*att1R*) and the 27 bp 5'-terminally truncated $tRNA^{Gly}$ between PA0714 and PA0715 as the left border (*att1L*), by exploiting the genome sequence information of *P. aeruginosa* type strain, PAO1 (http://www.pseudomonas.com). This 27 bp site (*att*) might serve as the attachment site that had been duplicated when this island was incorporated into the genome of the progenitor strain.

The Pf1 island of PAO1 contains 15 open reading frames (ORFs), and the central 11 ORFs are highly conserved (more than 95% identity in nucleotide) in the Pf1 phage genome that contains 13 genes [7]. The striking differences of the PAO1 Pf1 island from the phage Pf1 are i) the absence of Orf90, ii) the presence of three additional ORFs (PA0715, PA0716, and PA0729), and iii) the structural difference between both integrases (PA0728 for Pf1 island). PA0715 and PA0716 encode a putative reverse transcriptase and an ABC (ATP-binding cassette) transporter protein, respectively, whereas PA0729 shows 47% similarity to a hypothetical protein of *Yersinia pestis* plasmid pMT1 (data not shown). The significance of these structural differences remains to be explored in terms of the roles of this island in genome evolutions and/or ecological or pathogenic advantages.

Distribution of Pf1-Derived Genomic Islands in *P. aeruginosa* Strains

To test for the differences in the genomic architectures related with this region, we performed the Southern blot analysis using additional *P. aeruginosa* strains, PA14 and PAK. A single PCR product was observed from either PA14 or PAO1 chromosome (data not shown). The PA14 PCR product was used to hybridize to *Nru*I-digested *P. aeruginosa* chromosomes (Fig. 2). PAO1 exhibited a 2.6 kb (2,642 bp) fragment. PAK showed no hybridizing fragment, probably due to the absence of this island on the genome. In contrast, PA14 showed an 1.8 kb fragment (1,797 bp after sequencing), suggesting the polymorphic natures around this region. The absence of Pf1 island in PAK might be related with the fact that PAK is sensitive (i.e. not immune) to Pf1 infection, unlike PAO1 and PA14. We, therefore, determined the complete nucleotide sequences of the Pf1 island of PA14, anticipating that the

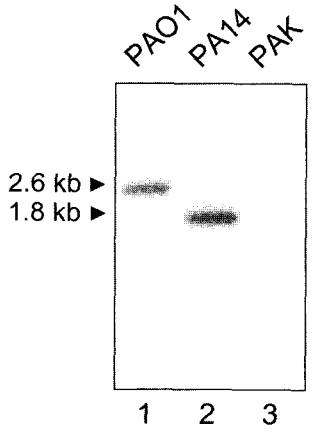

Fig. 2. Distribution of Pf1-derived genomic island in other *P. aeruginosa* strains.
Southern analysis with the *Nru*I-digested chromosomal DNAs from PAO1, PA14, and PAK verified the presence of one Pf1 island in both PAO1 (lane 1) and PA14 (lane 2) and the absence of any Pf1 island in PAK (lane 3). The probe was designed based on the PAO1 genome sequence, as represented by the solid line in Fig. 1. The sizes of the two fragments that hybridized to the probe are indicated.

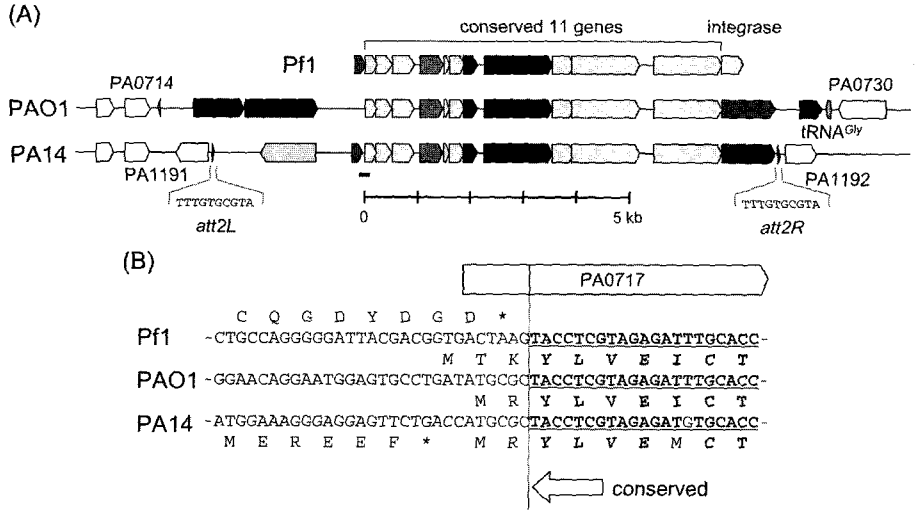

Fig. 3. Representation of the Pf1 island of PA14.
(A) Genetic organization of PA14 Pf1 island is shown in comparisons with PAO1 Pf1 island and the phage Pf1. The borders are the duplicated 10 bp sites (*att2L* and *att2R*) located between PA1191 and PA1192. PA14 Pf1 island contains 14 ORFs with the variable regions around both borders. Two ORFs (Orf317 and Orf71) and the integrase are contained in the variable regions. The solid line denotes the junction between the left variable and the conserved regions, the nucleotide sequences of which are shown in B. (B) Sequence variations in the left region are shown. Strong conservations (bold and underlined) are observed from the third or fourth codon (TAC) of PA0717, whereas highly variable sequences are at the nearest upstream of PA0717. Partially translated Orf90, Orf71, and PA0717 are shown with the conserved amino acid residues marked as bold.

genome diversifications around this region took place in both strains.

Organizations of the Pf1 Island in PA14

We screened for the Pf1 region in PA14 genomic libraries, isolated three cosmid clones, and determined the nucleotide sequences that sufficiently cover the ~10 kb region of the Pf1 island (data not shown). We could not find any tRNA elements within the 12,859 nucleotide sequences that we had determined. Instead, two 10 bp direct repeats of TTTGTGCGTA were separated by a 10,657 bp segment from comparisons made with the corresponding sequences of PAO1 (Fig. 3A), which might function as another *att* site (*att2*). The Pf1 island of PA14 contained 14 predicted ORFs and the conserved 11 ORFs showed strong identities to the corresponding ORFs, in PAO1 and Pf1, constituting the ancient and/or essential set of genes for the Pf1-related functions. The remarkable difference of the PA14 Pf1 island from that of PAO1 was the sequence variations around the left border (Fig. 3B) and the right border (data not shown), besides the *att* site sequences and locations, which results in i) the presence of Orf71 and Orf317 and ii) relatively lower identities (51.4%) in integrases (Fig. 4A). The weak identities in integrases are profound at the C-terminal and likely associated with the *att* site specificity.

Orf71 is a small protein that is, to some extent, similar to phage Pf1 Orf90 (26.1% identity and 59.6% similarity), whose function is not yet known (Fig. 4B). Orf317 is highly homologous to the regulatory protein, cII of retronphage ϕR73 (31.9% identity and 70.2% similarity). Both Orf317 and ϕR73 cII have a ParA (Soj)-like domain responsible

Fig. 4. Sequence comparisons of PA14 Pf1 proteins with related proteins.
(A) Integrases from both Pf1 islands (PA14 and PAO1) share significant homologies, whereas the putative integrase from the Pf1 phage is only of 100 amino acids. The relatively diverged C-terminal amino acids are shown as shaded. Identical residues are marked by asterisks, and similar residues are marked by colons or dots, following the definitions in CLUSTAL W program [19]. (B) Amino acid sequence comparison between Orf71 (PA14 Pf1) and Orf90 (phage Pf1) is shown as described in A.

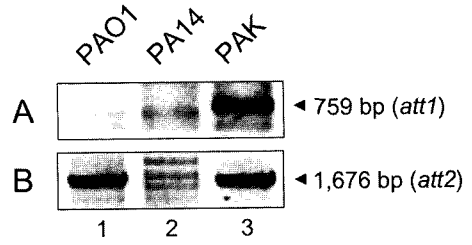

Fig. 5. PCR amplification of both attachment sites.
PCR primer pairs were used to amplify the region around the attachment (*att*) sites. For the $tRNA^{Gly}$ *att* site (*att1*, A), 714D and 730U primers were used as described in Materials and Methods, resulting in the amplification of a 759 bp fragment, unless the *att1* is interrupted. For the 10 bp *att* site (*att2*, B), 1190D and 1192U primers were used to amplify a 1,676 bp fragment, as long as the *att2* remains intact.

for ATP-dependent chromosome partitioning [8], although the physiological roles of Orf317 and its implications have yet to be revealed. By comparing both Pf1 islands with the Pf1 phage, we suggest that both Pf1 islands might be structurally diverged from the original Pf1 phage in a manner to increase their survival potentials in nature.

Attachment (*att*) Sites for Pf1 Islands for Genome Diversification

In PAO1, Pf1 is integrated at the *att1* between PA0714 and PA0730, whereas the Pf1 island of PA14 was integrated at the *att2* between PA1191 and PA1192. Since PAK did not have Pf1 island on its genome, we designed two primer pairs to see whether or not the *att1* and *att2* sites are occupied by other genetic segments in PAK. Thus, both PCR products (1,686 bp for *att2* and 759 bp for *att1*) were amplified from the PAK chromosome (Fig. 5). We isolated the *att1* PCR product from PAK, determined its nucleotide sequences, and compared the sequences with the corresponding region of PAO1 (data not shown). Some nucleotide differences (13 out of 719) were observed, which are most likely attributed to intraspecific variations. We found the presence of the intact $tRNA^{Gly}$ gene, implying the preservation of the unintegrated *att1* from the progenitor strain.. PAO1 gave the PCR product only from *att2*, with the absence of any DNA segment at the *att2* region, as expected. PA14, however, exhibited no PCR product from either region, implying that the PA14 *att1* is not empty either. This observation led us to suggest that the *att1* site has already been occupied by a genomic island that might be unrelated to the Pf1 phage (see Fig. 2).

CONCLUSION

In this study using three different *P. aeruginosa* strains, we have shown one of the intraspecific genome diversification mechanisms that exploit different phage-derived island variants and the different *att* site utilizations. Also, we found that apparently unrelated genomic islands might utilize the same *att* element. The highly-induced transcription of the Pf1 island and the abundance of Pf1-like phages in the fluid over PAO1 biofilms [20, 21] may play an important role in horizontal gene transfers to diversify its genomic architectures. A further understanding of the interrelationships between the gene expressions and functions within these genomic islands may provide a new perspective on the genome evolutions along with their roles in the pathogenic and ecological capabilities of different clinical and environmental isolates of *P. aeruginosa*, and may lead to the development of new detection and enumeration methods [13, 14] to identify the survival potential of this organism in natural samples.

Acknowledgment

This work was supported by the 21C Frontier Microbial Genomics and Applications Center Program from Ministry of Science and Technology, Korea to Y.-H.C. (MG02-0201-004-2-2-0).

REFERENCES

1. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25:** 3389-3402.
2. Besemer, J. and M. Borodovsky. 1999. Heuristic approach to deriving models for gene finding. *Nucleic Acids Res.* **27:** 3911-3920.
3. Chen, W. P. and T. T. Kuo. 1993. A simple and rapid method for the preparation of gram-negative bacterial genomic DNA. *Nucleic Acids Res.* **21:** 2260.
4. Cho, Y.-H., E.-J. Lee, B.-E. Ahn, and J.-H. Roe. 2001. SigB, an RNA polymerase sigma factor required for osmoprotection and proper differentiation of *Streptomyces coelicolor*. *Mol. Microbiol.* **35:** 150-160.
5. Hacker, J. and E. Carniel. 2001. Ecological fitness, genomic islands and bacterial pathogenicity: A Darwinian view of the evolution of microbes. *EMBO Rep.* **2:** 376-381.
6. Hansen, M. R., P. Hanson, and A. Pardi. 2000. Filamentous bacteriophage for aligning RNA, DNA, and proteins for measurement of nuclear magnetic resonance dipolar coupling interactions. *Methods Enzymol.* **317:** 220-240.
7. Hill, D. F., N. J. Short, R. N. Perham, and G. B. Petersen. 1991. DNA sequence of the filamentous bacteriophage Pf1. *J. Mol. Biol.* **218:** 349-364.
8. Ireton, K., N. W. Gunther, and A. D. Grossman. 1994. *spo0J* is required for normal chromosome segregation as well as the initiation of sporulation in *Bacillus subtilis*. *J. Bacteriol.* **176:** 5320-5329.
9. Jain, R., M. C. Rivera, J. E. Moore, and J. A. Lake. 2002. Horizontal gene transfer in microbial genome evolution. *Theor. Popul. Biol.* **61:** 489-495.

10. Karlin, S. 2001. Detecting anomalous gene clusters and pathogenicity islands in diverse bacterial genomes. *Trends Microbiol.* **9:** 335-343.
11. Karlin, S. and J. Mrázek. 2000. Predicted highly expressed genes of diverse prokaryotic genomes. *J. Bacteriol.* **182:** 5238-5250.
12. Larbig, K., C. Kiewitz, and B. Tümmler. 2002. Pathogenicity islands and PAI-like structures in *Pseudomonas* species. *Curr. Top. Microbiol. Immunol.* **264:** 201-211.
13. Lee, J. H., K.-H. Lee, and S. H. Choi. 2001. Enumeration of *Vibrio vulnificus* in natural samples by colony blot hybridization. *J. Microbiol. Biotechnol.* **11:** 302-309.
14. Lee, S.-P. and Y.-S. Lee. 2002. Development of rapid molecular detection marker for *Colletotrichum* spp. in leaf and fruit tissues of sweet persimmon. *J. Microbiol. Biotechnol.* **12:** 989-992.
15. Lowe, T. M. and S. R. Eddy. 1997. tRNA scan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25:** 955-964.
16. Ochman, H., J. G. Lawrence, and E. A. Groisman. 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405:** 299-304.
17. Rahme, L. G., E. J. Stevens, S. F. Wolfort, J. Shao, R. G. Tompkins, and F. M. Ausubel. 1995. Common virulence factors for bacterial pathogenicity in plants and animals. *Science* **268:** 1899-1902.
18. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, U.S.A.
19. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22:** 4673-4680.
20. Webb, J. S., L. S. Thompson, S. James, T. Charlton, T. Tolker-Nielsen, B. Koch, M. Givskov, and S. Kjelleberg. 2003. Cell death in *Pseudomonas aeruginosa* biofilm development. *J. Bacteriol.* **185:** 4585-4592.
21. Whiteley, M., M. G. Bangera, R. E. Bumgarner, M. R. Parsek, G. M. Teitzel, S. Lory, and E. P. Greenberg. 2001. Gene expression in *Pseudomonas aeruginosa* biofilms. *Nature* **413:** 860-864.