Oral Microbiology Immunology 2007: 22: 429–432
Printed in Singapore. All rights reserved

ORAL MICROBIOLOGY
AND IMMUNOLOGY

SHORT COMMUNICATION

Identification of strain-specific DNA of *Actinobacillus actinomycetemcomitans* by representational difference analysis

J. Mena, C. Chen
Divisions of Primary Oral Health Care, School of Dentistry, University of Southern California, Los Angeles, CA, USA

Mena J, Chen C. Identification of strain-specific DNA of Actinobacillus actinomycetemcomitans *by representational difference analysis.*
Oral Microbiol Immunol 2007: 22: 429–432. © 2007 The Authors. Journal compilation © 2007 Blackwell Munksgaard.

A genomic subtraction method, the representational difference analysis (RDA), was tested for its use in identifying strain-specific DNA in *Actinobacillus actinomycetemcomitans.* Subtraction of strain D7S with strain HK1651 yielded D7S-specific 2.3-kilobase (kb) islet-A and 5.3-kb islet-B. Islet-A contains a 1.5-kb region that is homologous to a region found in the *A. actinomycetemcomitans* plasmid pVT745. Islet-B contains a 2.1-kb homolog of *vgr*, a component of a DNA repeat element *rhs.* The distribution of these islets among *A. actinomycetemcomitans* strains was further examined by polymerase chain reaction. Islet-A was found in nine serotype a and two serotype b strains but was missing from 34 strains. Islet-B was found in one serotype a strain, four serotype d strains and two serotype e strains, but was missing from 34 strains.

Key words: genomic island; horizontal gene transfer; periodontal pathogens; representational difference analysis

Casey Chen, DEN 4335, USC School of Dentistry, Division of Primary Oral Health Care, Los Angeles, CA 90089, USA
Tel.: +1 213 740 1075;
fax: +1 213 740 2194;
e-mail: cchen@usc.edu
Accepted for publication January 17, 2007

The gram-negative facultative bacillus *Actinobacillus actinomycetemcomitans* is a major pathogen in human periodontitis (5, 27, 33). *A. actinomycetemcomitans* strains are classified into six distinct serotypes, a to f (14, 25). The extent of genomic differences among *A. actinomycetemcomitans* has not been fully examined.

Comparisons of completed genome sequences of closely related bacterial species or strains of the same species have revealed a surprising fluidity in bacterial genomes (3, 4, 7, 15, 16, 18–20, 23, 28, 32). While two closely related bacteria could share numerous common genes, each bacterium may possess one or more distinct DNA blocks acquired by horizontal gene transfer from distantly related bacteria.

Representational difference analysis (RDA) is a subtractive hybridization method for identifying unique DNA sequences between two samples (17). This method has been applied to examine the genetic differences between virulent and non-virulent strains within bacterial species, or between closely related pathogenic and non-pathogenic bacterial species. The resultant specific DNA fragments often provide new insight into the pathogenic mechanisms of bacteria. Moreover, RDA and similar subtractive hybridization methods have identified genomic-island-like elements in various bacterial species (9, 26, 34). The horizontally acquired genomic islands presumably carry genes that confer virulence mechanisms on the host bacteria.

The aim of this study was to test the utility of RDA to identify strain-specific DNA fragments and examine their distribution among diverse *A. actinomycetemcomitans* strains. Forty-seven *A. actinomycetemcomitans* strains were examined, including HK1651 (serotype b), Y4 (serotype b), JP2 (serotype b), ATCC29523 (serotype a), ATCC29524 (serotype b), ATCC33384 (serotype c), and 41 clinical isolates of serotypes a ($n = 9$), b ($n = 9$), c ($n = 10$), d ($n = 8$), e ($n = 4$), and f ($n = 1$). The *A. actinomycetemcomitans* strains were routinely grown in the sTSB medium (3% trypticase soy broth, 0.3% yeast extract, 1.5% agar, and 5% horse serum), or mTSB broth (3% trypticase soy broth, and 0.6% yeast extract) at 37°C in 5% CO_2.

The protocols for RDA were as described previously (17, 29) for strains D7S (tester) and HK1651 (driver). Two different libraries were constructed. In the first

library, the tester DNA was digested with *Sau*3AI, extracted with phenol–chloroform and precipitated with ice-cold 95% ethanol. The driver DNA was sheared by a sonifier (Brandson Sonifier 450, VWR Scientific) to produce DNA fragments of 0.4 to 3.0 kilobases (kb). Oligonucleotide adapters were ligated to the ends of the digested tester DNA. Three sets of complementary oligonucleotide pairs were designed to produce adapters for *Sau*3AI-restricted DNA. Each of these pairs consisted of a 24-mer and a 12-mer which, when annealed, generate a *Sau*3AI-compatible overhang. These three sets were FBam12 (5′-GATCCTCGGTGA-3′) and FBam24 (5′-AGCACTCTCCAGCCTCTCACCGAG-3′); SBam12 (5′-GATCCGTTCATG-3′) and SBam24 (5′ACCGACGTCGACTATCCATGAACG-3′); TBam12 (5′-GATCCTCCCTCG-3′) and TBam24 (5′-AGGCAACTGTGCTATCCGAGGGAG-3′). The second library was constructed using *Tsp*509I for digestion of tester DNA (D7S) and the corresponding three sets of adaptors were Feco12 (5′-AATTCTCGGTGA-3′) and Feco24 (5′-AGCACTCTCCAGCCTCTCACCGAG-3′), Seco12 (5′-AATTCGTTCATG-3′) and Seco24 (5′-ACCGACGTCGACTATCCATGAACG-3′), and Teco12 (5′-AATTCAGTCGGT-3′) and Teco24 (5′-GCGACTTCCAGGTTCAACCGACTG-3′). The details for amplification and hybridization were as described previously (17, 29). Fragments generated after three successive rounds of RDA were cloned into the *Bam*HI site of pBluescript (Stratagene, La Jolla, CA) then transformed into *Escherichia coli* DH5α using standard protocols. Transformants containing the desirable clones were submitted for sequencing at the USC School of Medicine Microchemical Core Facility.

A modified inverted polymerase chain reaction (PCR) sequencing protocol (30) was employed to determine the sequences flanking the RDA-detected strain D7S-specific DNA fragments. Briefly, strain D7S genomic DNA was digested with a selected restriction enzyme. The digested DNA was then self-ligated with T4 ligase at 15°C overnight to generate a circularized template. PCR primers were designed to amplify outwards from both ends of either islet-A or islet-B (see Table 1 for primer sequences). PCR amplification was performed to generate an amplicon that might include the flanking regions of the islets. The amplicons were purified and submitted for sequence determination. The inverted PCR sequencing process was repeated with a different restriction enzyme and with or without a different set of primers until a common DNA region between strain D7S and HK1651 was identified. Enzymes *Hin*dIII and *Bfa*I were used for sequencing of islet-A, and enzymes *Hin*dIII and *Sph*I were used for islet-B.

A standard PCR amplification protocol was used for sequence verification and to determine the distributions of islet-A and islet-B. The sequence of the primers and their usages are given in Table 1. Briefly, 30 μl PCR mixture contained 50–100 ng template, 0.5 μM of each primer, 3 μl dNTP stock (2 mM each of dATP, dCTP, dDTP and dTTP), 1.5 mM $MgCl_2$ and 1 unit *Taq* DNA polymerase in 1x *Taq* DNA polymerase buffer. The amplification performed had the following profile: 5 min at 94°C for denaturation followed by 30 cycles at 94°C for 30 s, an annealing step at 50–60°C for 1 min, an extension step at 72°C for 1–5 min and then a final extension of 7 min at 72°C. The annealing temperatures were calculated based on the GC compositions of the primers by the following formula: 4°C per G or C and 2°C per A or T (10). When indicated, the PCR amplicons were purified and submitted for sequencing at the USC School of Medicine Microchemical Core Facility. The genome sequence of strain HK1651 (University of Oklahoma, http://www.uoklahoma.edu) was used to search for homologs of the identified DNA fragments. The annotations of the HK1651 genome are provided at the Los Alamos National Laboratory Oral Pathogen Sequence Databases (http://www.oralgen.lanl.gov/). BLAST search was used to identify homologs in the National Center for Biotechnology Information (NCBI) database.

Twenty randomly selected clones from two RDA-constructed libraries were screened for strain D7S-specific DNA fragments. The results led to the identification of a 191-base-pair (bp) fragment and a 153-bp fragment that were not found in the genome sequence of strain HK1651. The flanking regions of the 191-bp fragment in strain D7S were identified by inverted PCR sequencing, which revealed an upstream region of the 191-bp fragment that was nearly identical to the 3′ end of

Table 1. Oligonucleotide primers and their uses

Primer	Sequence (5′–3′)	Usage
T3 promoter	AATTAACCCTCACTAAAGGG	T3/T7 for PCR amplification of the cloned inserts in pBluescript KS
T7 promoter	GTAATACGACTCACTATAGGGC	
191 W	CTA CGA GCA CGC CCA CAG CC	First-round inverted PCR sequencing for the flanking regions of islet-A.
191 CW	GTA CTC TCG GCC AAC GAC GTG	
W1-F	CTC TGG ATT TGG AGC GAC G	Second-round inverted PCR sequencing for the flanking regions of islet-A.
CW3-R	GAT TTC GTG GTA TTG TTG AAT G	
153 W	ATC AGA AAG AGC CTG TTC TCA	First- and second-round inverted PCR sequencing for the flanking regions of islet-B.
153 CW	CAA TGG GGC AGC GTG AAA C	
HisC-F (forward)	GAT CAG CCA GGC TTT GAA CGC	PCR sequencing of islet-A and islet-B and detection of these islets among *A. actinomycetemcomitans* strains. HisC-F, AroA-R, PonC, and Sgbh-B anneal to regions either upstream or downstream of the islets.
191 CW (reverse)	GTA CTC TCG GCC AAC GAC GTG	
191W (forward)	CTA CGA GCA CGC CCA CAG CC	
Vul-R (reverse)	GCT CGT GCT GTT TGA TAG CG	
Somnus-F (forward)	CTT GCA CAG TGT CGC CAG AC	
AroA-R (reverse)	GAT TTT CCA CGC TCA CGC CG	
PonC (forward)	GCT CCG AAA GCA GCT ATA TGC	
153 CW (reverse)	CAA TGG GGC AGC GTG AAA C	
VGR C (forward)	GGG ATT ACC GTT ATA AGT AAG CG	
Mena-R (reverse)	CAC CGG ACT TGC ATC TGC CG	
VgrG-S (forward)	CAG GTG GAA CTC AGT GGC C	
Sgbh-D (reverse)	AGA AGC TGA CAA CCA CCC AA	
Mena-F (forward)	GAG TGA GTG GTC GGT GGG AC	
Sgbh-B (reverse)	GAC GCA AGC CAT TTA CCA CCG	

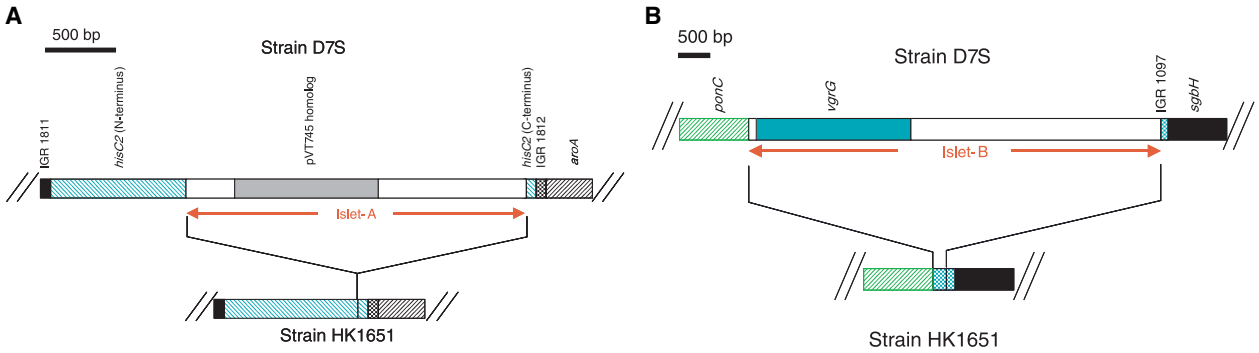

Fig 1. Comparison of islet-A (A) and islet-B (B) and their insertion sites in strain D7S relative to the corresponding regions in strain HK1651. The boundaries of the islets are indicated with red arrows. Islet-A was inserted within *hisC2* of strain D7S. Islet-B was inserted between *ponC* and *sgbH* in strain D7S. These islets were not found in strain HK1651.

the *hisC2* gene in strain HK1651 (see Fig. 1). Since *hisC2* was located upstream of *aroA* in the genome of strain HK1651, we made an assumption that a large DNA fragment (that contains the 191-bp fragment) was inserted within the region of *hisC2* and *aroA* in the genome of strain D7S. As expected, PCR amplification with primers HisC–F to AroA–R (primers located in *hisC2* and *aroA* respectively) resulted in a 3.6-kb amplicon. This 3.6-kb amplicon was then sequenced from both ends, and the results were used to design new primers to generate overlapping PCR amplicons for sequence confirmation. The results identified a 2.3-kb D7S-specific DNA fragment designated islet-A (GenBank accession no. EF187432). Similarly, the inverted PCR sequencing of the 153-bp fragment identified an upstream region that was identical to *ponC* of strain HK1651. The *ponC* gene was located upstream of *sgbH* in strain HK1651. A 5.6-kb amplicon was generated by PCR using primers annealing to *ponC* and *sgbH*. The PCR amplicon was sequenced to reveal a 5.3-kb D7S-specific DNA fragment, designated as islet-B (GenBank accession no. EF187433). The genetic maps of islet-A and islet-B in strain D7S and their corresponding regions in strain HK1651 are shown in Fig. 1.

The 2.3-kb islet-A was found within *hisC2* of strain D7S, splitting the gene into a larger 1014-bp fragment and a smaller 84-bp fragment, with a duplication of three nucleotides. The insertion site was mapped to the nucleotide coordinate 1,931,787 of strain HK1651 (which has an intact *hisC2*). The *hisC2* gene is presumably involved in the biosynthesis of aromatic amino acids in *A. actinomycetemcomitans.* Within this 2.3-kb islet is a 1.5-kb region that is homologous to a region in the *A. actinomycetemcomitans* plasmid pVT745 nucleotide coordinates 9241–10,732 (http://www.oralgen.lanl.gov) but missing the two small regions in the nucleotide coordinates 9788–9804 and 9874–9923. This 1.5-kb region contains three partial ORFs of pVT745: the last 80 nucleotides of AA11, a truncated AA13 and a truncated AA12 transcribing in the opposite direction. The functions of AA11, AA12, and AA13 are unknown.

The 5.3-kb islet-B was flanked by an intact *ponC* and an intact *sgbH* in strain D7S. The *ponC* and *sgbH* genes may function in cell wall synthesis. The insertion site of islet-B was mapped to the nucleotide coordinates 1,149,202–1,149,282 of strain HK1651 and resulted in a small deletion of the intergenic region IGR1097 (http://www.oralgen.lanl.gov/). Islet-B contains a 2166-bp homolog of *vgr* (named for its frequent Val-Gly repeats, see below for more information). The remaining DNA on islet-B did not show any significant homology with anything in the NCBI database.

The distributions of islet-A and islet-B were determined among 47 *A. actinomycetemcomitans* strains by PCR analysis. Primers that flanked the insertion site of islet-A (primers HisC-F/AroA-R) or islet-B (primers PonC-F/SgbH-B) were used to amplify the genomic DNA of the test strains. The insertion of the islets, or the lack of the islets, was determined based on the sizes of the amplicons. Strains with an insertion of the islet were subsequently confirmed by PCR with primers that amplified the corresponding joint regions. Islet-A was found in nine of the 10 serotype a strains and in two of the 13 serotype b strains. For the remaining 36 strains, 34 were negative for islet-A and two strains showed no reactivities in PCR assays. Islet-B was found in one of the 10 serotype a strains, four of the eight serotype d strains and two of the four serotype e strains, but were missing from 34 strains. The remaining six strains showed no reactivities by PCR assays.

An important consideration in the application of RDA is its efficiency in identifying target DNA fragments. While RDA application between closely related species has resulted in a significant number of target DNA (24, 29), the application of this method to strains of the same species generally did not result in a large number of desirable DNA fragments (1, 2). The limitation here may be the genetic similarity between the strains selected for RDA. Other methods are available for amplification and identification of strain-specific DNA fragments (6, 9, 26). It is likely that an extensive cataloging of strain-specific DNA fragments between a tester and a driver *A. actinomycetemcomitans* strain would require combinations of different methods. Importantly, the RDA and other PCR-based genomic subtraction methods have the advantage of not requiring previous sequence information of the strains.

The islet-A and islet-B contained apparent mutations that may have rendered these elements functionless. These two islets may represent ancient remnants of horizontal gene transfer-acquired genetic elements that underwent significant modifications such as deletions, insertions, and point mutations. The significance of the disruption of *hisC2* by the insertion of islet-A remains to be determined. The insertion of homologs of the pVT745 (found in islet-A) in the *A. actinomycetemcomitans* genome has been reported previously by Novak et al. (11, 21, 22). However, the regions and the insertion sites of the pVT745 homologs were different between their studies and this study. The *vgr* locus (found in islet-B) is commonly found to be part of a larger repetitive genetic element called Rhs in many bacteria (13, 31). The complete Rhs element consists of *vgr*, an Rhs core, and

the distal core extension (13, 31). The *vgr* locus and the Rhs core have different origins and can be disassociated in some bacteria. The Rhs elements are not essential for bacterial growth but may function in some specific ecological niches.

A question may arise concerning whether the ancestor of *A. actinomycetemcomitans* may have acquired these genetic elements, which were then selectively lost in some descendants. However, this alternative hypothesis seems unlikely, in particular to explain the presence of islet-A within the housekeeping gene *hicC2* in the ancestral strains of *A. actinomycetemcomitans*.

Genetic differences, either as the result of point mutations or horizontal gene transfer, could result in phenotypic differences among *A. actinomycetemcomitans* strains. For example, a deletion of 153-bp fragment in the promoter of the leukotoxin operon results in a highly leukotoxic phenotype in the JP-2 clone of *A. actinomycetemcomitans* (8). Hayashida et al. (12) showed that the inability of the JP-2 clone of *A. actinomycetemcomitans* to utilize heme was the result of mutations in the *hgpA* gene. In this study, the gene content and the insertion sites of islet-A and islet-B did not provide obvious clues to their functions. We did not detect any association of these islets to the clinical sources of the strains (data not shown). Additional studies are needed to understand the impact of horizontal gene transfer to the virulence/fitness of *A. actinomycetemcomitans* virulence.

Acknowledgments

This study was supported by NIDCR grant R01 DE12212. We are grateful to Bruce Roe, Fares Najar, Sandy Clifton, Tom Ducey, Lisa Lewis, and Dave Dyer for the use of unpublished nucleotide sequence data from the *A. actinomycetemcomitans* Genome Sequencing Project at The University of Oklahoma.

References

1. Allen NL, Hilton AC, Betts R, Penn CW. Use of representational difference analysis to identify *Escherichia coli* O157-specific DNA sequences. FEMS Microbiol Lett 2001: **197**: 195–201.
2. Allen NL, Penn CW, Hilton AC. Representational difference analysis: critical appraisal and method development for the identification of unique DNA sequences from prokaryotes. J Microbiol Methods 2003: **55**: 73–81.
3. Andersson JO, Andersson SG. Insights into the evolutionary process of genome degradation. Curr Opin Genet Dev 1999: **9**: 664–671.
4. Andersson SG, Kurland CG. Reductive evolution of resident genomes. Trends Microbiol 1998: **6**: 263–268.
5. Asikainen S, Chen C. Oral ecology and person-to-person transmission of *Actinobacillus actinomycetemcomitans* and *Porphyromonas gingivalis*. Periodontol 2000 1999: **20**: 65–81.
6. Bogush ML, Velikodvorskaya TV, Lebedev YB et al. Identification and localization of differences between *Escherichia coli* and *Salmonella typhimurium* genomes by suppressive subtractive hybridization. Mol Gen Genet 1999: **262**: 721–729.
7. Boucher Y, Nesbo CL, Doolittle WF. Microbial genomes: dealing with diversity. Curr Opin Microbiol 2001: **4**: 285–289.
8. Brogan JM, Lally ET, Poulsen K, Kilian M, Demuth DR. Regulation of *Actinobacillus actinomycetemcomitans* leukotoxin expression: analysis of the promoter regions of leukotoxic and minimally leukotoxic strains. Infect Immun 1994: **62**: 501–508.
9. Brown PK, Curtiss R, 3rd. Unique chromosomal regions associated with virulence of an avian pathogenic *Escherichia coli* strain. Proc Natl Acad Sci USA 1996: **93**: 11149–11154.
10. Dieffenbach CW, Lowe TM, Dveksler GS. General concepts for PCR primer design. PCR Methods Appl 1993: **3**: S30–37.
11. Galli DM, Chen J, Novak KF, Leblanc DJ. Nucleotide sequence and analysis of conjugative plasmid pVT745. J Bacteriol 2001: **183**: 1585–1594.
12. Hayashida H, Poulsen K, Kilian M. Differences in iron acquisition from human haemoglobin among strains of *Actinobacillus actinomycetemcomitans*. Microbiology 2002: **148**: 3993–4001.
13. Hill CW. Large genomic sequence repetitions in bacteria: lessons from rRNA operons and Rhs elements. Res Microbiol 1999: **150**: 665–674.
14. Kaplan JB, Perry MB, MacLean LL, Furgang D, Wilson ME, Fine DH. Structural and genetic analyses of O polysaccharide from *Actinobacillus actinomycetemcomitans* serotype f. Infect Immun 2001: **69**: 5375–5384.
15. Lan R, Reeves PR. Intraspecies variation in bacterial genomes: the need for a species genome concept. Trends Microbiol 2000: **8**: 396–401.
16. Lawrence JG, Ochman H. Molecular archaeology of the *Escherichia coli* genome. Proc Natl Acad Sci USA 1998: **95**: 9413–9417.
17. Lisitsyn N, Wigler M. Cloning the differences between two complex genomes. Science 1993: **259**: 946–951.
18. Martin W. Mosaic bacterial chromosomes: a challenge en route to a tree of genomes. Bioessays 1999: **21**: 99–104.
19. Maurelli AT, Fernandez RE, Bloch CA, Rode CK, Fasano A. "Black holes" and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. Proc Natl Acad Sci USA 1998: **95**: 3943–3948.
20. Nakata N, Tobe T, Fukuda I et al. The absence of a surface protease, OmpT, determines the intercellular spreading ability of *Shigella*: the relationship between the ompT and kcpA loci. Mol Microbiol 1993: **9**: 459–468.
21. Novak KF, Dougherty B, Pelaez M. *Actinobacillus actinomycetemcomitans* harbours type IV secretion system genes on a plasmid and in the chromosome. Microbiology 2001: **147**: 3027–3035.
22. Novak KF, LeBlanc DJ. Characterization of plasmid pVT745 isolated from *Actinobacillus actinomycetemcomitans*. Plasmid 1994: **31**: 31–39.
23. Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. Nature 2000: **405**: 299–304.
24. Perrin A, Nassif X, Tinsley C. Identification of regions of the chromosome of *Neisseria meningitidis* and *Neisseria gonorrhoeae* which are specific to the pathogenic *Neisseria* species. Infect Immun 1999: **67**: 6119–6129.
25. Saarela M, Asikainen S, Alaluusua S, Pyhalea L, Lai CH, Jousimies Somer H. Frequency and stability of mono- or polyinfection by *Actinobacillus actinomycetemcomitans* serotypes a, b, c, d or e. Oral Microbiol Immunol 1992: **7**: 277–279.
26. Schouler C, Koffmann F, Amory C, Leroy-Setrin S, Moulin-Schouleur M. Genomic subtraction for the identification of putative new virulence factors of an avian pathogenic *Escherichia coli* strain of O2 serogroup. Microbiology 2004: **150**: 2973–2984.
27. Slots J, Ting M. *Actinobacillus actinomycetemcomitans* and *Porphyromonas gingivalis* in human periodontal disease: occurrence and treatment. Periodontol 2000 1999: **20**: 82–121.
28. Sokurenko EV, Hasty DL, Dykhuizen DE. Pathoadaptive mutations: gene loss and variation in bacterial pathogens. Trends Microbiol 1999: **7**: 191–195.
29. Tinsley CR, Nassif X. Analysis of the genetic differences between *Neisseria meningitidis* and *Neisseria gonorrhoeae*: two closely related bacteria expressing two different pathogenicities. Proc Natl Acad Sci USA 1996: **93**: 11109–11114.
30. Triglia T, Peterson MG, Kemp DJ. A procedure for *in vitro* amplification of DNA segments that lie outside the boundaries of known sequences. Nucleic Acids Res 1988: **16**: 8186.
31. Wang YD, Zhao S, Hill CW. Rhs elements comprise three subfamilies which diverged prior to acquisition by *Escherichia coli*. J Bacteriol 1998: **180**: 4102–4110.
32. Whittam TS, Bumbaugh AC. Inferences from whole-genome sequences of bacterial pathogens. Curr Opin Genet Dev 2002: **12**: 719–725.
33. Zambon JJ. *Actinobacillus actinomycetemcomitans* in human periodontal disease. J Clin Periodontol 1985: **12**: 1–20.
34. Zhang YL, Ong CT, Leung KY. Molecular analysis of genetic differences between virulent and avirulent strains of *Aeromonas hydrophila* isolated from diseased fish. Microbiology 2000: **146**: 999–1009.