Mapping transposon insertion sites by touchdown PCR and hybrid degenerate primers

Julio Levano-Garcia, Sergio Verjovski-Almeida, and Ana C.R. da Silva
Universidade de São Paulo, São Paulo, SP, Brazil

BioTechniques 38:225-229 (February 2005)

A novel mapping method based on touchdown PCR was developed for identifying a transposon insertion site in genomic DNA using a hybrid consensus-degenerate primer in combination with a specific primer that anneals to the transposon. The method was tested using Xanthomonas citri *transposon mutants. PCR products contained adjacent DNA regions that belonged to both* X. citri *genomic DNA and the transposon. Products were directly sequenced from PCRs using only the specific primer. Different PCR conditions were tested, and the optimized reaction parameters that increased product yields and specificity are described. Best results were obtained with the HIB17 hybrid primer, which is a 25-mer oligonucleotide having degenerate bases at 6 different positions within the last 12 bases at the 3′ end. An X. citri mutants library was produced by random transposition using the EZ::TN <KAN-2> transposon, and we identified the insertion sites within the genome of 90 mutants. Insertions were found within both the chromosomal and the plasmid DNA in these* X. citri *mutants. Restriction mapping and Southern blot analysis confirmed the insertion sites for eight randomly chosen mutants. This method is a very useful tool for large-scale characterization of mutants in functional genomics studies.*

INTRODUCTION

Genomic technology popularized whole genome sequencing. To date, more than a hundred microbial and eukaryotic genomes have been sequenced (1,2). This expansion increases the number of genes with unknown function or genes with just putative functions that have been deposited at GenBank® (http://www.ncbi.nlm.nih.gov). Usually, a second phase in genome projects involves the functional characterization of annotated genes. This task is achieved using several methodologies, and one of the most popular is gene disruption by marker insertion (3). This methodology has the advantage of producing a large number of mutants in a few rounds of insertion (4,5). However, finding the exact site of insertion is time-consuming because it requires mapping and isolation of flanking DNA followed by sequencing. The most frequently used approach to map these insertion mutants involves whole genome subcloning into cosmids/bacterial artificial chromosomes (BACs) devoid of marker selection genes, which are screened by the presence of the marker that had been previously inserted into the genome for gene disruption. The selected cosmid/BACs are sequenced in parts to determine the insertion position (6). This methodology precludes large-scale screening of mutants. To overcome these difficulties, most laboratories first perform phenotype screening and then map the insertion region (7). However, the latter approach has the bias of only mapping disrupted genes that cause detectable phenotype-positive mutants.

In addition, two other methods are used for mapping random insertions within genomic DNA, namely thermal asymmetric interlaced PCR (tail PCR; Reference 8) and inverse PCR (9,10). They have been demonstrated as good tools but involve multiple steps, which become very expensive for large-scale analysis of mutants. Moreover, inverse PCR is difficult to apply to genomes containing over 10^9 bp (11).

Here we report a new methodology for mapping a randomly inserted marker into the genome of a target cell, which allows rapid large-scale screening and identification of an insertion site. It is based on touchdown PCR (12,13) using a pair of primers where one is a hybrid consensus-degenerate oligonucleotide and the other is a sequence-specific primer that only anneals to one of the strands of the inserted marker gene, with its 3′ end pointing toward the unknown genomic sequence. The resulting PCR products are sequenced using solely the specific primer. The initial portion of the obtained sequence serves as confirmation that the PCR product has started at the desired inserted marker, and then the remaining sequence corresponds to genomic sequence and reveals where in the genome of the target cell the marker has been inserted. This methodology also has other applications such as genome walking/gap closure in a genomic sequencing project. In this case, the specific primer represents the end of known sequence at the border of a gap. To test the methodology, we mapped several mutants generated by random insertion of a transposon containing a kanamycin resistance marker gene into the *Xanthomonas axonopodis* pv. *citri* (*X. citri*) genome.

MATERIALS AND METHODS

Bacterial Strains, Vectors, and Culture Conditions

The *X. citri* strain 306 (14) used in this study is a virulent citrus canker type strain, which was routinely cultured in liquid or on solid Klebs-Loeffler Bacillus (KLB) medium at 28°C. For the selection of recombinants, the KLB media were supplemented with 50 µg/mL kanamycin. *Escherichia coli* strain DH10B (Invitrogen, Carlsbad, CA, USA) was used as a host for the plasmids and was cultured in liquid or on solid Luria-Bertani (LB) medium at 37°C. *E. coli* was transformed by electroporation (15), and culture media were supplemented with 50 µg/mL kanamycin.

In Vivo Transposition

X. citri randomly inserted mutants have been produced with a transposition methodology (16) using the EZ::TN™<KAN-2> Transposome™ Kit (EPICENTRE, Madison, WI, USA). Transposons were introduced into *X.*

citri cells by electroporation (17) using a Gene Pulser® (Bio-Rad Laboratories, Hercules, CA, USA). *X. citri* mutants were selected on solid KLB media supplemented with 50 µg/mL kanamycin after 48 h of incubation at 28°C. The mutants obtained were cultured in 96-well microplates with liquid 2× trypsin-yeast (TY) extract media supplemented with 8% glycerol and stored at -80°C.

Table 1. Primers Used in This Work

Primer Name	Sequence	Melting Temperature (°C)
KAN-2 FP-1	5′-ACCTACAACAAAGCTCTCATCAACC-3′	61.3
KAN-2 RP-1	5′-GCAATGTAACATCAGAGATTTTGAG-3′	58.1
APYBAF1	5′-CGGAATTCCGTGTTAAATATGGTATTGTGA TNGAYKSNGGNTC-3′	71.6 (average)
HIB17	5′-CGGAATTCCGGATNGAYKSNGGNTC-3′	67.1 (average)

DNA Isolation

Genomic DNA from mutants was isolated with DNAzol® (Invitrogen). Mid-scale plasmid DNA isolation from *E. coli* was performed using 500 mL of culture and the Plasmid Maxi Kit (Qiagen, Valencia, CA, USA). Genomic and plasmid DNA samples were analyzed on a 1% agarose gel [1× TAE (Tris-acetate-EDTA)] and stained with ethidium bromide.

Specific Primer Design Strategy

Specific primer can be any 20- to 25-mer oligonucleotide, preferably a 23- to 25-mer oligonucleotide having a sequence composition that exactly matches the known sequence of one of the strands of the inserted marker gene. Sequence is chosen so that annealing at the known site occurs with the 3′ end of the primer pointing toward the end of the inserted marker gene, near the target genomic sequence. This site of annealing is preferably at 40–60 bp from the end of the inserted marker. Table 1 shows the two specific primers used in this work. These primers originally belonged to the EZ::TN <KAN-2> Primers Transposon Insertion Kit and were chosen because of the absence of similarity to the *X. citri* genome sequence (14).

Consensus-Degenerate Hybrid Primers Design Strategy

Hybrid primers used in this study were designed according to the CODEHOP method strategy (18) and can be any 25- to 43-mer oligonucleotide having a nondegenerate consensus sequence composition at the 5′ end (13–31 bp), followed by a sequence comprising 10 bp with degenerate bases at different positions within this segment, and followed by 2 bp at the 3′ end, which are nondegenerate. These primers were originally used in a touchdown PCR method for the construction of cDNA libraries using *Schistosoma mansoni* mRNA (13). Examples of consensus-degenerate hybrid primers APYBAF1 and HIB17 used in this work are presented in Table 1.

Low-Stringency PCR Conditions Design

Touchdown PCR methodology (12) was used, and the conditions employed were modified from those previously described for the generation of cDNA mini-libraries by low-stringency touchdown reverse transcription PCR (RT-PCR; Reference 13). A GeneAmp® PCR System 9700 (Applied Biosystems, Foster City, CA, USA) was used at maximal ramp speed. The PCRs were performed in a 50-µL mixture containing 1× PCR buffer (60 mM Tris-SO_4, 18 mM NH_4SO_4), 2.5 mM $MgSO_4$, 0.2 mM each dNTP, and 1.5 U of either Platinum® *Taq* DNA Polymerase High Fidelity or Recombinant *Taq* DNA Polymerase (Invitrogen). Touchdown PCR protocol #1 consisted of two phases: phase 1 included an initial step of 94°C for 8 min, followed by 30 cycles of denaturation at 94°C for 1 min, annealing at variable temperatures for 1 min, and extension at 72°C for 1 min. In the first cycle, the annealing temperature was set to 60°C and, at each of the 29 subsequent cycles, the annealing temperature was decreased by 0.5°C (i.e., it varied from 60° to 45°C at 0.5°C decrements along the 30 cycles). Phase 2 consisted of 10 cycles of 94°C for 1 min, 45°C for 1 min, and 72°C for 1 min.

Protocol #2 consisted of phase 1, with an initial step of 95°C for 5 min, followed by 25 cycles of denaturation at 95°C for 45 s, annealing at variable temperatures for 45 s, and extension at 72°C for 2 min. The annealing temperature was set at 60°C in the first cycle and, at each of the 24 subsequent cycles, it was decreased by 0.5°C per cycle down to 47.5°C. Phase 2 consisted of 25 cycles of 95°C for 45 s, 50°C for 45 s, and 72°C for 2 min. After the last PCR cycle, the samples were cooled to 4°C, and a 6-µL aliquot of the amplification products was electrophoresed on a 1% agarose gel (1× TAE), stained with ethidium bromide, and visualized under ultraviolet (UV) light. Other conditions were tested, modifying both annealing and extension times and the number of cycles, without further improvement in the yield of PCR amplifications.

DNA Sequencing

PCR products obtained were purified using a Concert™ PCR Purification Kit (Invitrogen) essentially to remove unused primers. Sequencing reactions were performed by dideoxynucleotide chain termination (19) with the BigDye™ Terminator Kit (Applied Biosystems) using 5 pmol of sequencing primer (which is the specific primer used in the touchdown PCR; that is, KAN-2 RP-1 or FP-1) and analyzed with an ABI PRISM® 3700 Automated Sequencer (Applied Biosystems).

Analysis of Transposon-Insertion Sites

For mapping the insertion position within genome, the sequence reads obtained were aligned (BlastN) against the *X. citri* whole genome sequence (14). For each analyzed sequence, the 5′-AGATGTGTATAGAGACAG-3′

consensus region, which flanks the transposon, was localized and the location where the transposon was inserted within the genome was determined by identifying the start and end positions of the matching sequence with respect to the genome sequence. The orientation of each transposon insertion was also established.

Southern Blot Analysis

Restriction analysis using the Southern blot method was performed (20) on eight *X. citri* mutants to confirm the accuracy of the present mapping methodology. Genomic DNA of mutants was digested with the enzymes *Pst*I, *Pvu*I, and *Ava*I (New England Biolabs, Beverly, MA, USA) for this analysis.

RESULTS AND DISCUSSION

A library of random mutants was constructed by the introduction of mini-transposon EZ::TN <KAN-2> into a kanamycin-sensitive *X. citri* strain. We reproducibly obtained approximately 1×10^5 kanamycin-resistant colonies per 0.1 pmol of transposon. To optimize the mapping methodology, we initially identified the transposon insertion site within a single *X. citri* mutant (IA3). A combination of either 10 or 20 pmol of KAN-2 RP-1-specific primer with 10, 50, and 100 pmol of each of two hybrid consensus-degenerate primers was evaluated using the Touchdown PCR protocol #1, as described in Materials and Methods (Figure 1A). Both examples of hybrid primers, namely APYBAF1 or HIB17, worked well at 100 pmol per reaction with 20 pmol of specific KAN-2 RP-1 primer (Figure 1A). Genomic DNA template concentrations in the range of 100–400 ng were used, with no significant differences (data not shown). The critical control, done in the presence of specific primer KAN-2 RP-1 alone (Figure 1A, lane C1) shows no nonspecific amplification. Nonspecific amplification products were distinguished (a faint broad band with 1.1–1.4 kb and a defined 1.8 kb band) with primer HIB17 control (Figure 1A, lane C3). However, it should be noted that these nonspecific amplifications originating from priming with hybrid primer alone do not interfere with the method because the resulting products contain sequence of nonspecific hybrid primers that are not sequenced in the final step (sequencing is done with the nondegenerate specific primer).

Subsequently, touchdown PCR protocol #2 was tested with six different *X. citri* mutants that were assayed using the optimized primer concentrations from the experiments of Figure 1A; namely, 100 pmol hybrid primer (APYBAF1 or HIB17) plus 20 pmol KAN-2 RP1-specific primer (Figure 1B). Essentially, an increment in the annealing temperature was introduced (from 45°–50°C) in phase 2 (increase of stringency). Also, an increase in the number of cycles in phase 2 of PCR (from 10–25) and an increase in the elongation time (from 1–2 min) were introduced. Good amplifications were obtained using HIB17 hybrid primer (Figure 1B, even numbered lanes); however, APYBAF1 yielded poor or no amplifications (Figure 1B). The more diverse annealing capacity of HIB17 hybrid primer might be related to its shorter length (25-mer) as compared to APYBAF1 (43-mer). Again, two clear bands (<0.5 kb) of amplification by HIB17 are present in all even lanes of Figure 1B, suggesting nonspecific products that did not interfere with the final step of sequencing.

Protocol #2 was considered to be the most robust one, and it was subsequently used with all mutants tested. Figure 2 shows the amplification obtained for 14 different mutants using the optimized conditions described above. All mutants generated products, except for mutant IE1 (Figure 2, lane 13). In addition, we found that freshly prepared genomic DNA templates were essential for good amplifications. Thus, when using genomic DNA from IG1 mutant that was either freshly prepared (lane 1) or a 4-month-old isolated DNA (lane 14), we observed a remarkable loss in performance with the latter template. Thus, all mutants described in Figure 2 and Table 2 were characterized with freshly isolated DNA.

Using protocol #2 and the optimized conditions described above, we found that specific primer KAN-2 FP-1 could be used in place of RP-1 with no detectable effect on the amount of product yield (data not shown). Therefore, we used either one of the

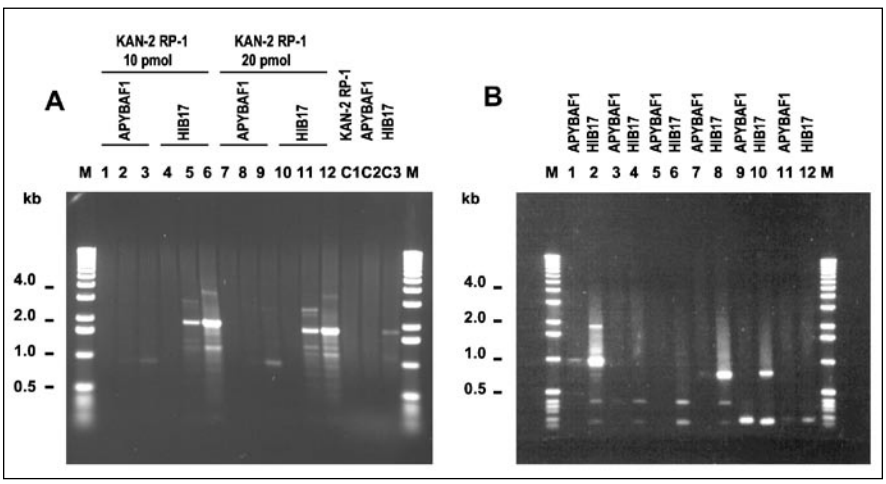

Figure 1. **Touchdown PCR products obtained using hybrid consensus-degenerate primers in *Xanthomonas citri* insertion mutants.** (A) Genomic DNA from a single *X. citri* mutant (IA2) was used. Touchdown PCRs were performed with a specific KAN-2 RP-1 primer and one of the two hybrid primers APYBAF1 or HIB17. Different concentrations of the specific KAN-2 RP-1 primer were tested: 10 pmol (lanes 1–6) or 20 pmol (lanes 7–12). Either APYBAF1 (lanes 1–3 and 7–9) or HIB17 (lanes 4–6 and 10–12) hybrid consensus-degenerate primers were tested at the concentration range of 10, 50, and 100 pmol. Controls were performed with 20 pmol of a single primer: KAN-2 RP-1-specific primer (lane C1), APYBAF1 hybrid primer (lane C2), or HIB17 hybrid primer (lane C3). Nonspecific amplification was only detected in control C3. Molecular weight marker (1 kb DNA ladder; Invitrogen) is shown (lane M). (B) Different *X. citri* mutants were tested using a single optimized touchdown PCR condition, namely 100 pmol hybrid primer (APYBAF1 or HIB17) plus 20 pmol KAN-2 RP-1-specific primer. Mutants tested were IC1 (lanes 1 and 2), IE1 (lanes 3 and 4), IG1 (lanes 5 and 6), IH1 (lanes 7 and 8), IH3 (lanes 9 and 10), and IH5 (lanes 11 and 12). Ethidium bromide stained 1% agarose gels are shown.

Table 2. Summary of Touchdown Amplification Success Rates

Touchdown PCR Samples	Total	Sequences Yielded[c]	No Sequence[d]	Yield[e] (%)
Strong Amplification[a]	73	70	3	96
Weak Amplification[b]	39	20	19	51
No Amplification	32	N.A.	32	0
Total Tested	144	90	54	63

N.A., not applicable.
[a]Strong amplification was defined as having PCR products similar to those in Figure 2, lanes 9–12.
[b]Weak amplification was defined as having PCR products similar to those in Figure 2, lanes 6–8.
[c]Sequencing of touchdown PCR product was possible.
[d]Sequencing of touchdown PCR product was not possible because of more than one insertion site of transposon in the cell or very weak amplifications.
[e]Percent of successful sequencing with respect to total number of clones studied.

Table 3. Summary of Sequenced Transposon Insertion Sites

	gDNA	Plasmid DNA[e]	Total
Mapped Clones[a]	81	9	90
Distinct Sites[b]	77	9	86
Intergenic Sites	14	1	15
Sites Within Genes[c]	67	8	75
Redundant Sites[d]	4	N.A.	4

gDNA, genomic DNA; N.A., not applicable.
[a]Total number of insertions located within the *Xanthomonas citri* genome.
[b]Number of unambiguously different positions in the genome sequence where a transposon was observed.
[c]Annotated *X. citri* open reading frames.
[d]Clones with possible hot spot sites of transposase or cells originated from the duplication of the same clone after the transposition reaction.
[e]*X. citri* plasmids pXac 33 and pXac 66 (14).

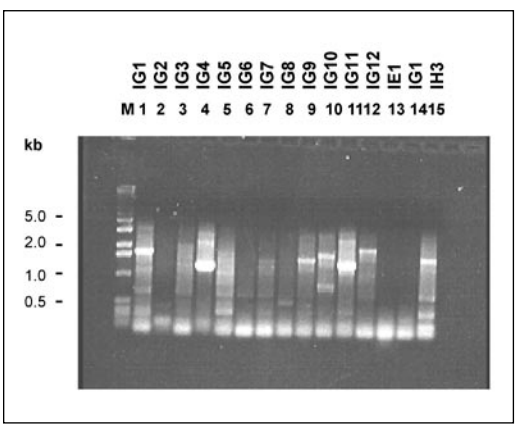

Figure 2. Robust amplification obtained with optimized touchdown PCR conditions. Fourteen different *Xanthomonas citri* transposon insertion mutants were screened using optimized condition two, which is described in the text. Note that robust amplifications were obtained for the majority of mutants tested using HIB17 hybrid primer (lanes 1–13 and 15). In lane 14, genomic DNA isolated from mutant IG1 and stored for three months with apparently no degradation was used as template; however, no amplification products were obtained (lane 14), in contrast to a robust amplification obtained with a fresh preparation of the same DNA (lane 1).

specific primers for characterization of the sites of insertion for 144 mutants. Sequencing the PCR products with the respective KAN-2 RP-1 (or FP-1)-specific primer and comparing the obtained sequence with the whole genome sequence of *X. citri* (14) resulted in the determination of sites of insertion for 90 mutants, which represented 63% of 144 mutants analyzed (Table 2). Sequences obtained from both strong and weak amplifications represented a 96% and 51% success rate, respectively (Table 2). Moreover, in the mapped mutants, 90% of insertions were found in gDNA (chromosomal), and 10% of insertions were located within plasmids (pXac 33 and pXac 66; Table 3). These results show that mutations were random (Table 3), and their insertion events occurred with a unique transposon per cell.

The high GC content of *Xanthomonas* might be responsible for the slightly low percentage of success (63%) that we obtained in the screening of 144 mutants (Table 2) because it increases the amount of secondary structure present in the genome and makes PCR amplification and sequencing a more difficult task (3). In fact, during the *Xanthomonas* genome sequencing project, we faced difficulties when sequencing some specific stretches of the genome because of very stable hairpin loops (14).

Southern blot analysis was performed on eight different *X. citri* mutants to confirm the insertion site (data not shown). Restriction fragments detected for each clone matched with the predicted in silico fragments.

In conclusion, a new method based on touchdown PCR with consensus-degenerate hybrid primers was developed for the mapping of a transposon insertion site in genomic DNA. In general, low-stringency touchdown PCR conditions can be found to ensure that the consensus-degenerate hybrid oligonucleotide primer will anneal to the target genomic sequence in the vicinity of the insertion site and generate a PCR product. Thus, this method should be useful for the rapid large-scale screening and identification of the insertion site of a known inserted marker gene within the target genome of a number of different cells. However, one limitation is that the present strategy cannot be applied for characterizing mutants with two or more transposon insertions in the genome. The method may have other applications such as genome walking/gap closure in a genomic sequencing project. In this case, the specific primer represents the end of known sequence at the border of a gap.

ACKNOWLEDGMENTS

We thank Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil, for supporting this work. We also thank Drs. Julio C.F. Oliveira and Luiz Roberto Furlan for helpful discussion.

REFERENCES

1. **Benson, D.A., I. Karsch-Mizrachi, D.J. Lipman, J. Ostell, and D.L. Wheeler.** 2003. GenBank. Nucleic Acids Res. *31*:23-27.

2. **Bernal, A., U. Ear, and N. Kyrpides.** 2001. Genomes OnLine Database (GOLD): a monitor of genome projects world-wide. Nucleic Acids Res. *29*:126-127.
3. **Dhandayuthapani, S., W.G. Rasmussen, and J.B. Baseman.** 1999. Disruption of gene mg218 of *Mycoplasma genitalium* through homologous recombination leads to an adherence-deficient phenotype. Proc. Natl. Acad. Sci. USA *96*:5227-5232.
4. **Goryshin, I.Y., J. Jendrisak, L.M. Hoffman, R. Meis, and W.S. Reznikoff.** 2000. Insertional transposon mutagenesis by electroporation of released Tn5 transposition complexes. Nat. Biotechnol. *18*:97-100.
5. **Mangan, M.W. and W.G. Meijer.** 2001. Random insertion mutagenesis of the intracellular pathogen *Rhodococcus equi* using transposomes. FEMS Microbiol. Lett. *205*:243-246.
6. **Bishop, A., S. Fielding, P. Dyson, and P. Herron.** 2004. Systematic insertional mutagenesis of a streptomycete genome: a link between osmoadaptation and antibiotic production. Genome Res. *14*:893-900.
7. **Gallagher, L.A., S.L. McKnight, M.S. Kuznetsova, E.C. Pesci, and C. Manoil.** 2002. Functions required for extracellular quinolone signaling by *Pseudomonas aeruginosa*. J. Bacteriol. *184*:6472-6480.
8. **Liu, Y.G., N. Mitsukawa, T. Oosumi, and R.F. Whittier.** 1995. Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant J. *8*:457-463.
9. **Rich, J.J. and D.K. Willis.** 1990. A single oligonucleotide can be used to rapidly isolate DNA sequences flanking a transposon Tn5 insertion by the polymerase chain reaction. Nucleic Acids Res. *18*:6673-6676.
10. **Huang, G., L. Zhang, and R.G. Birch.** 2000. Rapid amplification and cloning of Tn5 flanking fragments by inverse PCR. Lett. Appl. Microbiol. *31*:149-153.
11. **Ochman, H., F.J. Ayala, and D.L. Hartl.** 1993. Use of polymerase chain reaction to amplify segments outside boundaries of known sequences. Methods Enzymol. *218*:309-321.
12. **Don, R.H., P.T. Cox, B.J. Wainwright, K. Baker, and J.S. Mattick.** 1991. "Touchdown" PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res. *19*:4008.
13. **Fietto, J.L., R. DeMarco, and S. Verjovski-Almeida.** 2002. Use of degenerate primers and touchdown PCR for construction of cDNA libraries. BioTechniques *32*:1404-1411.
14. **da Silva, A.C., J.A. Ferro, F.C. Reinach, C.S. Farah, L.R. Furlan, R.B. Quaggio, C.B. Monteiro-Vitorello, M.A. Van Sluys, et al.** 2002. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. Nature *417*:459-463.
15. **Dower, W.J., J.F. Miller, and C.W. Ragsdale.** 1988. High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Res. *16*:6127-6145.
16. **Goryshin, I.Y. and W.S. Reznikoff.** 1998. Tn5 in vitro transposition. J. Biol. Chem. *273*:7367-7374.
17. **Yang, Y.C., M.K. Yang, T.T. Kuo, and J. Tu.** 2001. Structural and functional characterization of the lexA gene of *Xanthomonas campestris pathovar citri*. Mol. Genet. Genomics *265*:316-326.
18. **Rose, T.M., E.R. Schultz, J.G. Henikoff, S. Pietrokovski, C.M. McCallum, and S. Henikoff.** 1998. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. Nucleic Acids Res. *26*:1628-1635.
19. **Sanger, F., S. Nicklen, and A.R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*:5463-5467.
20. **Sambrook, J., E.F. Fritch, and T.E. Maniatis.** 1989. Molecular Cloning: A Laboratory Manual. CSH Laboratory Press, Cold Spring Harbor, NY.

Received 6 July 2004; accepted 23 August 2004.

Address correspondence to Ana C.R. da Silva, Alellyx Applied Genomics, Rodovia Anhanguera Km 104 (Techno Park), Rua James C. Maxwell 320, 13067-850 Campinas, SP, Brazil. e-mail: ana.rasera@alellyx.com.br