

Full Length Research Paper

# Genetic sequences derived from suppression subtractive hybridization analysis provides insight into their possible roles in *Xanthomonas albilineans*

Rehana Shaik<sup>1</sup>, Balakrishna Pillay<sup>1</sup> and Dorsamy Pillay<sup>1,2\*</sup>

<sup>1</sup>Discipline of Microbiology, Faculty of Science and Agriculture, University of KwaZulu-Natal (Westville Campus), Private Bag X54001, Durban 4000, Republic of South Africa.

<sup>2</sup>Department of Biotechnology, Faculty of Engineering, Science and the Built Environment, Durban University of Technology, P. O. Box 1334, Durban 4000, Republic of South Africa.

Accepted 4 April, 2008

Leaf scald disease (LSD) is caused by the Gram-negative bacterium, *Xanthomonas albilineans*. Genomic DNA from *X. albilineans* and *Xanthomonas hyacinthi* were analyzed by suppression subtractive hybridization (SSH) using *X. albilineans* as the tester from which unique sequences were sought and *X. hyacinthi* as the driver. Following the SSH procedure, amplification products within the size range of 100 - 600 bp were generated, purified, directly cloned with the Promega pGEM-T vector cloning kit, and transformed into ultracompetent *Escherichia coli* X L2-blue MRF' cells (Stratagene, La Jolla, CA). Clones selected were sequenced (using a Perkin Elmer ABI PRISM Dye terminator cycle sequencing kit and ABI Model 377 DNA sequencer) in one direction with SP6 and T7 primers (Promega). Clone Xa 6 revealed very close homology with a probable bacterioferritin from *Pseudomonas aeruginosa*. Clone X. *albilineans* 12 showed 92% homology to the acetate repressor proteins and clone X. *albilineans* 18 displayed 85% homology to the plasmid pTOM9 from *Alcaligenes xylosoxidans*. Sequencing data also revealed homology to various hypothetical proteins.

**Key words:** suppression subtractive hybridization, *Xanthomonas albilineans*, *Xanthomonas hyacinthi*, sequencing.

## INTRODUCTION

Leaf scald an insidious disease caused by the bacterium *Xanthomonas albilineans* colonizes the vascular system of sugarcane (*Saccharum* spp. hybrids) in either a chronic or acute phase. The disease can also be manifested in an asymptomatic latent phase. Numerous outbreaks have occurred throughout all regions of the world giving rise to concern that leaf scald could become a limiting economic factor in sugarcane production. This investigation focuses on the identification of sequences unique to *X. albilineans* using the SSH procedure.

The availability of new genomic sequences in the last

decade has created novel opportunities to analyze the organization of a genome's regulatory machinery, the function of particular genes or gene clusters and the evolutionary relationships between different bacterial strains and species. Genes that are present in certain isolates of a given bacterial species and absent or substantially different in others can be of great interest biologically (Akopyants et al., 1998; Janke et al., 2001). Some may determine strain-specific traits such as drug-resistance (Davies, 1994), bacterial surface structure (Stroeher and Manning, 1997; Zhang et al., 2005), or restriction-modification systems (King and Murray, 1994). The "pathogenicity islands" (PAIs) or multigene segments of virulent strains that tend to be absent from avirulent members of the same species and that help determine the nature and severity of disease are of special signific-

\*Corresponding author. E-mail: pillayb1@ukzn.ac.za. Tel: +27 312607404. Fax: +27 312607803.

ance (Groisman and Ochman, 1996; Hacker et al., 1997). Many of the genes or DNA segments specific to individual strains were found by the special phenotypes they confer, detailed physical mapping, or comparisons of sequence data from different isolates or taxa (Janke et al., 2001).

In light of the vast amount of sequence information available, molecular genetic techniques such as SSH (Diatchenko et al., 1996) have become an important tool in analyzing pathogenic bacteria. Use of the SSH approach facilitates identification of genomic differences between related bacterial strains. Pools of genomic DNA fragments from a bacterial strain of interest (tester) are, in effect, depleted, by hybridization and PCR, of sequences that are also present in a reference strain (driver). The remaining DNA fragments, highly enriched for tester-specific sequences, are then cloned for further analysis. SSH, with the advantage of requiring only small amounts of genomic DNA, is applicable to any characterized or uncharacterized genome identifies large genomic differences typically found between bacterial genomes (Radnedge et al., 2001). The aim of this study was to identify sequences that were unique to *X. albilineans* in order to further understand the biological profile of this bacterial phytopathogen.

## MATERIALS AND METHODS

DNA from *X. albilineans* and *Xanthomonas hyacinthi* was extracted using the QIAamp DNA mini kit (Qiagen-supplied by Southern Cross Biotechnology). Genomic subtraction was performed by using the PCR-Select Bacterial Genomic Subtraction Kit (Clontech), with *X. hyacinthi* as the driver and *X. albilineans* as the tester.

In order to create tester and driver fragments with blunt ends, purified genomic DNA was restricted with *Rsa*1, a four-base-cutting restriction enzyme (GT/AC). The reaction mixture contained approximately 2 µg DNA, 2 µl of 10x *Rsa*1 buffer and 1 U *Rsa*1 restriction enzyme, in a 40 µl total reaction volume. This mixture was vortexed, centrifuged briefly and incubated for 4 h at 37°C. The MinElute PCR Purification Kit (Qiagen) was used to clean up *Rsa*1 digested DNA. Fragments ranging from 70 to 4 kb were purified from primers, nucleotides, enzymes, and salts using MinElute spin columns in a microcentrifuge. DNA was stored at -20°C in deionised water.

Ligation of adaptors to the tester DNA was necessary for PCR-based amplification of subtracted material in the first and second rounds of PCR following subtraction. In order to ligate adaptors, tester DNA has to be divided into two portions, and each ligated to a different adaptor, either adaptor1 or 2R. Adaptors were not ligated to the driver DNA.

Purified *Rsa*1-digested tester DNA was diluted by adding 5 µl of sterile water to 1 µl of DNA (2 µg). Two microlitres of this solution was placed in a 0.5 ml Eppendorf tube and mixed with 2 µl adaptor 1 (10 µM), 1 µl of 10x ligation buffer (660 mM Tris-HCL, 50 mM MgCl<sub>2</sub>, 10 mM dithiothreitol, 10 mM ATP – pH 7.5), 1 µl of 1U/µl T<sub>4</sub> DNA ligase (Roche) and 4 µl sterile water. This ligation reaction is referred to as 1-1. An equivalent reaction was prepared in order to ligate adaptor 2 R to tester DNA, (referred to as 1-2). From each ligation 2 µl was removed and mixed in a third tube (referred to as to sample 1-3). Sample 1-3 contained tester DNA and a mixture of adaptor 1 and adaptor 2 R and was used to amplify unsubtracted material, and also to test that ligation had occurred. Ligations were

incubated at 14°C overnight and thereafter heated to 72°C for 5 min to inactivate the DNA ligase.

Sample 1-3, generated above, was used to assess the efficiency of the adaptor ligation. As 1-3 contains both adaptors 1 and 2 R, PCR amplification with SSH Primer 1 resulted in a smear of products representing different *Rsa*1 restriction enzyme products. One microlitre of sample 1-3 was diluted with 99 µl of water. From this diluted sample 1 µl was placed in a 0.5 ml Eppendorf and mixed with 1 µl each of 10 µM nested primer 1 and nested primer 2, 0.5 µl dNTP's (10 mM), 2.5 µl of 10x *Taq*-buffer and 1U *Taq*-polymerase enzyme (Roche). The final volume of this reaction mixture was 25 µl. This reaction mixture was thereafter subjected to PCR conditions of 1 cycle at 94°C for 1 min, 35 cycles denaturation at 94°C for 10 s, annealing at 68°C for 10 s and extension at 72°C for 1.5 min. The final extension step included 1 cycle at 72°C for 5 min. Seven microlitres of the PCR-products were separated by electrophoresis in a 2% (w/v) agarose gel.

In the first hybridization, an excess of driver DNA was added to each tester DNA (1-1 and 1-2), and the samples were heat denatured and allowed to anneal. After this hybridization, single-stranded DNA will be enriched for tester-specific DNA, as DNA fragments that are not tester specific will form hybrid molecules with the driver DNA. Prior to subtraction, tester DNA was diluted relative to driver DNA. In this study, the tester was previously diluted 1:6 after adaptor ligation and then followed by a 1:5 dilution (hence 30-fold dilution in total). A further 10-fold dilution with water (300-fold dilution in total) was performed prior to hybridization with driver material. However, dilution level can be varied to increase or decrease the stringency of the subtraction.

Two master mixtures contained 1.5 µl *Rsa*1 digested DNA (driver), 1 µl of 4 x hybridization buffer (Clontech) and 1.5 µl of 1:10 diluted tester 1-1 and 1.5 µl of tester 1-2, respectively, in a total volume of 4 µl. Once hybridizations were set-up, samples were overlaid with 10 µl of mineral oil, incubated at 98°C for 1.5 min and at 68°C for more than 6 h. The incubation did not exceed 12 h.

In the second hybridization, both samples from the first hybridization containing, adaptor1 and adaptor 2 R-ligated tester hybridized with an excess of driver, were mixed together. Freshly denatured *Rsa*1-digested driver DNA was added to further enrich for tester sequences. Single stranded DNAs, not annealed in the hybridization, formed new hybrid molecules, carrying different adaptors on each end. The reaction mixture comprised 1 µl of *Rsa*1 digested driver DNA, 1 µl of 4 x hybridization buffer and 2 µl of sterile distilled water. This mixture was overlaid with 4 µl of mineral oil and incubated at 98°C for 1.5 min (to denature the sample) and thereafter held at 68°C. The sample from hybridization 2 was removed into the pipette and then some air was drawn into the tip. Thereafter, the entire volume of freshly denatured driver was drawn into the tip. All of this was transferred into hybridization 1 and mixed with a pipette. It was important to keep all samples at 68°C while proceeding with the above steps. The mixture was thereafter incubated at 68°C overnight. One hundred microlitres of dilution buffer were added to the hybridization, heated to 68°C for 7 min and stored at -20°C. Unique sequences of tester DNAs, enriched through hybridization, were amplified using PCR. In the first suppression PCR only double stranded (ds) cDNA with different adaptors at the 5' and 3' ends were exponentially amplified, whereas those with the same adaptors form secondary structures and amplification is suppressed.

The reaction mixtures containing 1 µl template DNA, 2.5 µl of 10 x PCR buffer (Clontech), 0.5 µl of 10 mM dNTP mix, 1 µl of 10 mM PCR primer 1 (Clontech), 0.5 µl of 50 x Advantage DNA polymerase mix (Clontech) and 19.5 µl sterile distilled water were prepared in triplicate. This reaction mixture was briefly centrifuged. To this mixture, 1 µl of experimental sample (2<sup>nd</sup> hybridization, sterile distilled water and 1 µl of 1-3 ligation) respectively, was added. These mixtures were heated at 75°C for 5 min to extend the adaptors. Reactions were subjected to PCR conditions of 1 cycle at

94°C for 25 s, thereafter, 25 cycles of denaturation at 94°C for 10 s, annealing at 66°C for 10 s, and extension at 72°C for 1.5 min. Thereafter, the reaction was subjected to 1 cycle at 72°C for 5 min. The second nested PCR amplification was used to further reduce background and to enrich for unique tester sequences. Three microlitres of primary PCR product was diluted with 27 µl of sterile distilled water. The master mix for the secondary PCR reaction comprised the following: 2.5 µl of 10x PCR buffer, 1 µl of 10 µM nested PCR primer1 (Clontech), 1 µl of 10 µM nested PCR primer 2R (Clontech), 0.5 µl of 10 mM dNTP mix, 0.5 µl of 50 x Advantage DNA polymerase mix and 18.5 µl of sterile distilled water. This was briefly mixed and centrifuged. Twenty microlitres were added to 1 µl of hybridization/primary PCR product and 1 µl of sterile distilled water. The reaction mixture was subjected to the following PCR conditions: 1 cycle at 94°C for 1 min, followed by 30 cycles denaturation at 94°C for 10 s, annealing at 68°C for 10 s and extension at 72°C for 1.5 min and a final extension at 72°C for 5 min. Nine microlitres of secondary PCR product were separated after electrophoresis in a 2% agarose gel. The secondary PCR amplification was repeated. Reactions were then set-up for hybridized- and unextracted- material (1-3).

Purification of secondary PCR products is required for probe generation for Southern analysis and for the cloning of PCR products into the pGEMT-easy vector (Clontech). Glycerol and high salt concentration from PCR reactions are liable to inhibit subsequent applications and have to be removed. PCR products were purified using the MinElute PCR Purification Kit (Qiagen). To generate complex probes from hybridized- and subtracted- material (1-3) common sequences such as adaptors were removed by restriction digestion with *Rsa*I. The restriction reactions comprised the following components: 10 µl of purified PCR products, 2 µl of 10 x *Rsa*I buffer, 2 µl of *Rsa*I (Roche 10 U/ µl) and 6 µl of sterile distilled water. This reaction mixture was incubated for 4 - 5 h at 37°C. Digested material was separated by electrophoresis on a 1.5% (w/v) agarose gel and desired products minus the adaptors were extracted and purified using the QIAquick Gel Extraction Kit (Qiagen-supplied by Southern Cross Biotechnology). This protocol is designed to extract and purify DNA from standard or low-melting agarose gels in TAE or TBE buffer.

DNA labelling was performed using High Prime (Roche). Twenty five nanograms of template DNA was diluted with 13 µl of sterile distilled water. The probe was denatured at 95°C for 5 min and then placed on ice. The denatured DNA was transferred to a tube containing 4 µl High Prime (Roche) and 25 µCi radioactively labelled [ $\alpha$ -<sup>32</sup>P] dCTP. This reaction was incubated at 37°C for 1 h. Samples were incubated at 100°C for 5 min followed by incubation on ice. Thereafter, the contents of the tube were centrifuged and used immediately for hybridization. Southern analysis is used to assess the level of enrichment following the SSH procedure. PCR products from subtracted- and unextracted material were separated by gel electrophoresis, transferred to a positively charged nylon membrane and hybridized with complex probes derived from PCR amplified subtracted or unextracted probes.

Pre-hybridization and hybridization of blots were carried out in Denhardt's hybridization buffer (5 x SSC, 5 x Denhardt's buffer [2% (w/v) polyvinylpyrrolidin, 2% (w/v) ficoll 400, 2% (w/v) BSA, 0.5% (w/v) SDS] and incubated at 65°C in a hybridization oven. Pre-hybridization of filters was carried out for at least an hour prior to adding radioactive-labelled denatured DNA. The hybridization was then carried out for 16 h at 65°C.

Following hybridization, membranes were washed twice in 100 ml of 1 x SSC, 0.1% (w/v) SDS, followed by two washing steps in 0.1 x SSC, 0.1% (w/v) SDS. The first low stringency wash was carried out for 5 min at room temperature and the second for 15 min at 65°C for 30 min. Washed membranes were blotted dry using 3 MM paper, wrapped in cling film to prevent dehydration and exposed to X-ray film at -80°C utilizing intensifying screens. Blots were stored at -20°C.

In order to remove hybridized probes from the blots, membranes were incubated with 50 ml stripping solution I [0.4 M NaOH for 30 min at 45°C followed by washing with stripping solution 2 (0.2 M Tris-HCl-pH 7.6, 1x SSC, 0.1% (w/v) SDS)] for 30-60 min at 45°C. Filters were rinsed in 2 x SSC, blotted dry and stored at -20°C.

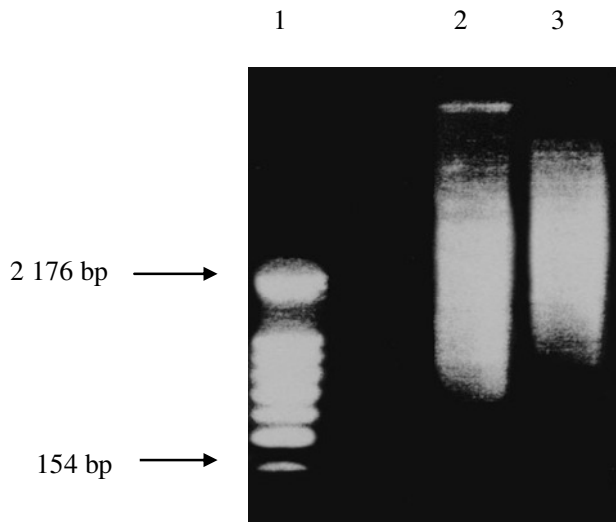
Insert DNA was ligated with pGEM-T Easy vector (Promega) which contains T7 and SP6 RNA polymerase promoters, flanking a multiple cloning region within the  $\alpha$ -peptide coding region of the enzyme  $\beta$ -galactosidase. Insertion of DNA leads to inactivation of the  $\alpha$ -peptide and allows identification of recombinant clones by colour screening on indicator plates. The pGEM-T Easy vector contains a 3'-terminal deoxythymidine supporting ligation of PCR products with a 3'-terminal deoxyadenosine and added during PCR amplification by non-proofreading enzymes such as *Taq* DNA polymerase.

The pGEM-T Easy vector system was optimized using a 1:1 molar ratio of insert DNA to vector. Reactions were carried out in 10 µl reaction volumes containing 10-50 ng insert DNA, 50 ng pGEM-T Easy vector (50 ng/µl), 1 x rapid ligation buffer (30 mM Tris-HCl - pH 7.8, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 1 mM ATP, 5% (w/v) polyethylene glycol MW 8000) and 3 Weiss units *T*<sub>4</sub>-DNA-Ligase (Promega). Samples were incubated at 4°C for 16 h and stored at -20°C.

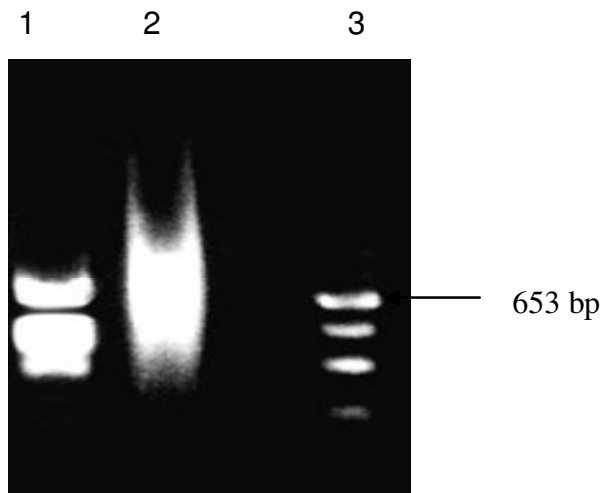
XL2-blue MRF<sup>+</sup> ultracompetent cells (Stratagene, La Jolla, CA) were used. For every 100 µl of ultracompetent cells, 1.7 µl of  $\beta$ -mercaptoethanol was added. This was placed on ice for 10 min and swirled every 2 min. Five microlitres of the ligation reaction were added to the test tube, mixed and placed on ice for 1 h.

Forty-five microlitres of 10 x freezing medium (62.7 g K<sub>2</sub>HPO<sub>4</sub>; 17, 96 g KH<sub>2</sub>PO<sub>4</sub>; 5.0 g TrisNaCitrate; 8. 98 g NH<sub>4</sub>SO<sub>4</sub>; 440 ml glycerol and 900 ml sterile distilled water), 5 ml MgSO<sub>4</sub> and 250 µl of Ampicillin (100 mg/ml) was prepared using LB broth to make 500 ml. Sixty microlitres of this medium were aliquoted into each microtitre well. White colonies were selected using toothpicks and were inoculated into the microtitre well. The microtitre plates were incubated at 37°C overnight. Clones that grew well and appeared very turbid were chosen for further analysis. Terrific broth medium (Sigma) was used for the inoculation of clones. For every 9.25 g of Terrific broth used, 1.6 ml of glycerol was added, brought to 200 ml volume and autoclaved. Thereafter, 100 µl of ampicillin (200 mg/ml) was added to give a concentration of 100 µg/ml. One microlitre of the appropriate clones was removed from the freezing medium and inoculated into a well containing Terrific broth. The plate was then sealed with sticky gas permeable tape and incubated for 24 h in a shaking incubator (300 rpm) at 37°C.

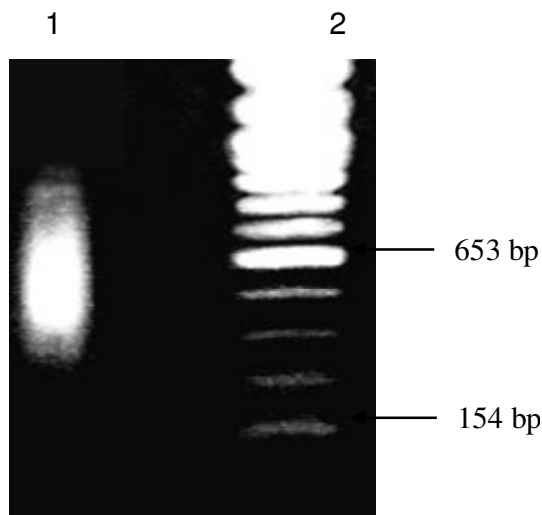
The plate was placed on special adaptors supplied by Beckman and centrifuged at 4 000 rpm for 7 min. Supernatant was decanted and tapped on tissue paper. Two hundred microlitres of solution 1 [TE-*Rnase*A +T1 (stored at 4°C), 50 mM Tris.HCl pH 7.6 (diluted from 1 M stock), 10 mM EDTA pH 8 (diluted from stock), 40 µg/ml *Rnase*, 40 µg/ml *Rnase* T1] were added, sealed and inverted 10 times. This was followed by the addition of 200 µl freshly prepared solution 2 [0.2 M NaOH, 1% SDS], sealed and inverted 10 times. Thereafter, 200 µl solution 3 (NaOAc.3H<sub>2</sub>O) were added, sealed and mixed on vortex for 2 min while moving the plate around. Plates were incubated on ice for 30 min or overnight at -20°C. The plates were centrifuged for 45 min at 4 000 rpm. Two hundred microlitres of the supernatant were removed and placed in a clean 96-well plate. To this 500 µl of 95% ethanol were added, sealed and inverted 10 times. The plate was incubated at -20°C for 30 min and thereafter centrifuged at 4 000 rpm for 30 min. The supernatant was discarded. Five hundred microlitres of 70% ethanol were added to wash the pellet and was centrifuged at 4 000 rpm for 15 min and the supernatant was discarded. The pellet was then air-dried. Fifty microlitres water (HPLC-*Rnase* A) were added to the pellet and incubated at 37°C for 1 h. High fidelity *Taq* and buffer (Roche) was used in these reactions. The reaction mixture comprised 2 µl template DNA, 10 µl of 10 x buffer, 20 µl of 1 mM DNTP's, 0.5 U



**Figure 1.** Restriction of genomic DNA from *Xanthomonas albilineans* (lane 2) and *Xanthomonas hyacinthi* (lane 3). Lane 1 – Marker VI (Roche).



**Figure 3.** Subtracted (lane 1) and unsubtracted (1-3) (lane 2) *X. albilineans* DNA after PCR amplification. Lane 3 – Marker VI (Roche).



**Figure 2.** Products of PCR confirming 1-3 ligations of subtracted *X. albilineans* DNA after 1 hour. Lane 1, *X. albilineans* DNA; Lane 2, Marker VI (Roche).

*Taq* polymerase (Roche), 1 µl each pGEM forward and reverse primers and 65 µl of sterile distilled water. This reaction mixture was subjected to PCR conditions of denaturation at 94°C for 5 min followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 64°C for 30 s and extension at 72°C for 1 min with a final extension at 72°C for 10 min.

Wizard<sup>®</sup> Plus Miniprep DNA Purification System (Promega) was used to purify insert DNA. The sequencing reaction consisted of 2 µl Big Dye (Roche), 0.8 µl of 5 x dilution buffer, 2 µl SP6 primer, 2 µl of DNA and 3.2 µl of sterile distilled water. The following PCR conditions were used: 25 cycles of denaturation at 96°C for 10 s, annealing at 50°C for 5 s and extension at 60°C for 4 min. Thereafter, the reaction was maintained at 4°C.

Sequence data were edited to remove vector sequences using DNAMAN Version 5.0 DNA analysis programme (Lynnon Biosoft).

Similarity searches in DNA and protein databases were performed using BLAST analysis (Pubmed).

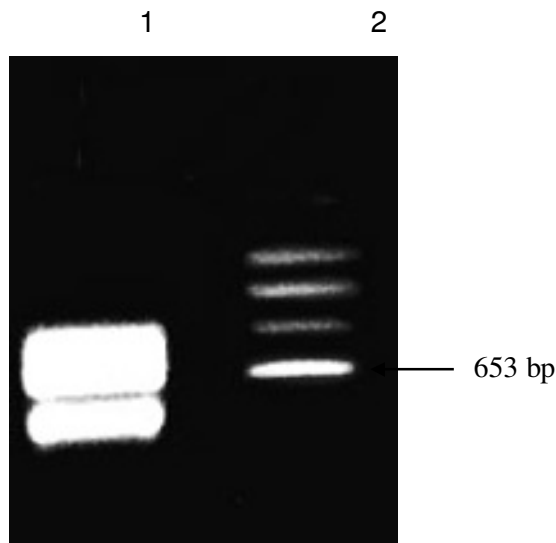
## RESULTS

Genomic DNA from *X. albilineans* and *X. hyacinthi* were successfully digested by the restriction endonuclease *Rsa*I (Figure 1). *X. albilineans* as the tester from which the unique sequences were being sought can be visualized in lane 2 (Figure 1) and *X. hyacinthi* as the driver in lane 3. Figure 2 confirms the 1-3 ligations and visualization of the products that were electrophoresed on an agarose gel after an hour. Figure 3 shows the differences between the subtracted and unsubtracted (1-3) DNA. Electrophoresis following PCR of the unsubtracted DNA resulted in a streak of DNA (lane 2) whereas electrophoresis following PCR of the subtracted DNA resulted in the formation of unique fragments (lane 1).

A second round of PCR was conducted to allow the analysis of the subtracted DNA (Figure 4). PCR products were successfully purified and electrophoresed as indicated in Figure 4.

Following the SSH procedure, amplification products from 100 to over 500 bp were generated, purified with the Wizard<sup>®</sup> Purification Kit (Promega), directly cloned with the Promega pGEM-T vector cloning kit, and transformed into ultracompetent *Escherichia coli* XL2-blue MRF' cells (Stratagene, La Jolla, CA). The library obtained was enriched for *X. albilineans* genomic sequences. Clones were randomly selected from the library for plasmid isolations and were successfully isolated from 22 clones (Figure 5).

Following plasmid isolations, insert lengths of the products were determined by PCR and subsequent agarose



**Figure 4.** Purified secondary PCR product of subtracted DNA. Lane 1 – *X. albilineans* (tester) DNA; lane 2 – Marker VI (Roche).

gel electrophoresis (Figure 6). Sizes of the subtracted DNA fragments varied from 100 – 600 bp (Figure 6). Clones were randomly selected and sequenced (Perkin Elmer ABI PRISM Dye terminator cycle sequencing kit and ABI Model 377 DNA sequencer) in one direction with SP6 and T7 primers (Promega).

Figure 7 indicates confirmation of the subtractive hybridization procedure were the restricted *X. albilineans* and *X. hyacinthi* genomic DNA's as well as other *Xanthomonas* spp. was blotted onto filters and were thereafter hybridized with the PCR amplified inserts (Figure 7). The products hybridized to *X. albilineans*.

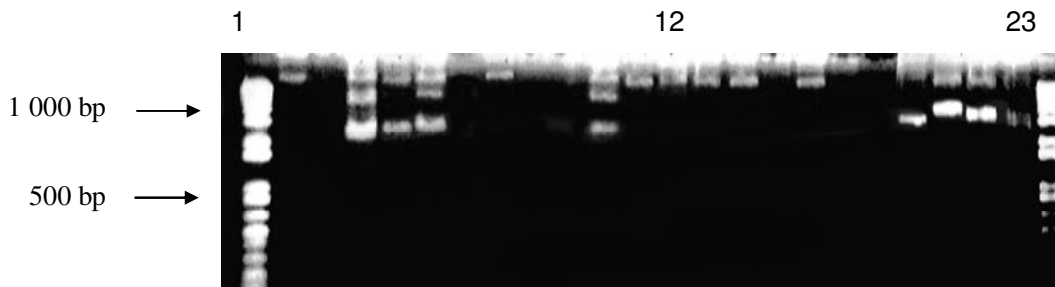
The insert sizes were determined after sequencing and ranged from 402 – 688 bp. These inserts were sequenced and compared to sequences deposited in Genbank using the BLASTX search (Table 1). Of the 15 clones selected for PCR amplification, 6 were found to be redundant.

## DISCUSSION

It is now appreciated that genomic variation within closely related groups of bacteria can be substantial, suggesting a need to define and study these differences (Sauerbaum and Achtman, 1999). High throughput subtractive hybridization is an approach that allows comprehensive genomic surveys of strains by directing sequencing to regions that differ among strains. Genomic variation is often associated with the acquisition and deletion of large (10-50 kb) regions of DNA (Lawrence and Ochman, 1997; Perna et al., 2001). SSH relies on the isolation of restriction fragments, which are contained within such regions.

The present study exploited the use of the SSH technique to isolate sequences that were unique to *X. albilineans*. Large fragments are not efficiently amplified and therefore not well represented in the product pool. Furthermore, small fragments below about 200 bp, which would otherwise preferentially amplify, are rarely amplified because of sequence complementarity in the adaptors, which promotes the formation of panhandle structures that are stable enough below this size range to greatly reduce amplification. Therefore, it is important to have restriction fragments between 0.2 and 2.0 kb. The SSH experiments were performed with the genomic DNA of *X. albilineans* as tester and *X. hyacinthi* as driver using the restriction endonuclease *Rsa*I. Southern blot analysis verified the SSH technique, as the tester specific sequences hybridized only to *X. albilineans* DNA. In some cases excess driver DNA containing complementary sequences may remain single stranded and hybridize with a complementary strand from the tester DNA to form an amplifiable product. This was also observed by Agron et al. (2002). Identifying false positives requires semi-quantitative hybridization experiments (Bogush et al., 1999; Emmerth et al., 1999), or evaluation by PCR amplification (Agron et al., 2002; Radnedge et al., 2001) and these steps can be eliminated when a complete sequence of the driver genome is available, greatly improving the power of the approach. In the case of SSH, the mirror orientation selection (MOS) is a method that has been reported to allow the reduction of false positives, which could reduce the amount of sequencing necessary for genomic surveys (Rebrikov et al., 2000).

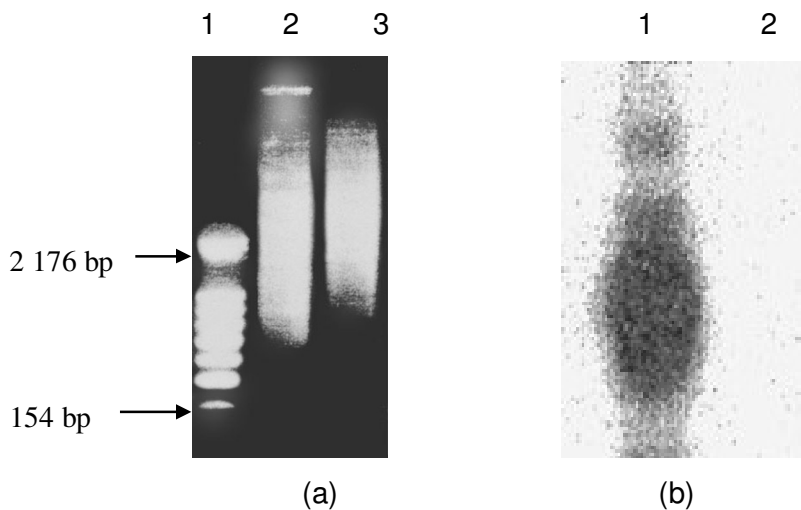
The SSH fragments were further investigated by cloning and determination of their DNA sequences (Table 1). Clones were directly sequenced after PCR amplification of the cloned inserts. Reports by Janke et al. (2001) suggest that although clones were analyzed for further study by dot-blot hybridization and the potential importance of DNA segments (containing sequences that matched and some that did not match to reference DNA), it will often make it worthwhile to move directly to sequencing of subtracted clones without such pre-screening. The selected clones had insert sizes that ranged from 402 – 688 bp and some showed significant similarity to sequences deposited in GenBank. Many of the clones that were initially selected were redundant. This was not surprising as *X. albilineans* and *X. hyacinthi* were very closely related and after subtraction only a few tester specific fragments were identified. Several of the cloned fragments also showed no homology to known DNA or protein sequences. The G+C content of these fragments differs from the usual G+C content of the *X. albilineans* chromosome indicating a possible acquisition by horizontal gene transfer. Similar studies, in which fragments showed no homology to DNA or protein sequences, were also reported by Janke et al. (2001). Akopyants et al. (1998) also suggested that unknown genes might encode completely new classes of proteins,



**Figure 5.** Agarose gel electrophoresis of plasmid products of 22 randomly selected clones. Lanes 1 and 24: 100 bp marker (Roche).



**Figure 6.** Products of 15 inserts after PCR amplification. Lanes 1 and 16: Marker VI (Roche).



**Figure 7.** (a) Restriction of *X. albilineans* (lane 2) and *X. hyacinthi* (lane 3); (b) Hybridization of products to *X. albilineans* verifying the SSH procedure (Lane 1: *X. albilineans*, lane 2: *X. hyacinthi*).

contribute to remarkable specificity of individual strains for particular host individuals, or might help determine the nature and severity of disease and may be of great usefulness for evolutionary studies.

Clone *X. albilineans* 8 showed distinct homology to a probable bacterioferritin from *P. aeruginosa*. Storage and buffering of iron is achieved by this class of proteins (Laulhere et al., 1992). The iron-containing bacterioferri-

**Table 1.** Similarities at the DNA level between cloned DNA sequences from subtracted *X. albilineans* DNA and sequences in the databases.

Clone no.	Clone length (bp)	Matching sequence from data base	Origin of matching sequence	DNA % match BLASTX
Xa2	602	Hypothetical protein	<i>Arabidopsis thaliana</i>	69% (5e-30)
Xa6	489	Probable bacterioferritin	<i>Pseudomonas aeruginosa</i>	90% (8e-09)
Xa8	687	Hypothetical protein	<i>Photobacterium damsalae</i>	65% (1e-05)
Xa9	463	Hypothetical protein	<i>Pseudomonas sp.</i>	73% (8e-25)
Xa12	402	Acetate operon repressor	<i>Salmonella typhimurium</i>	92% (4e-51)
Xa18	472	Plasmid pTOM9	<i>Alcaligenes xylosoxidans</i>	85% (2e-25)
Xa19	609	Hypothetical protein	<i>Bradyrhizobium</i>	54% (6e-08)
Xa22	688	Hypothetical protein	<i>Mycobacterium</i>	45% (0.41)

tins contain the protoporphyrin haem group. The optical spectra of the haem group show it to be predominantly low-spin in both the ferrous and ferric states (Cheesman et al., 1990). The nature of the axial ligands binding the haem group to the polypeptide has, however, remained unknown. Low-spin, bis-coordinate haem centres in proteins, typically have a role in rapid electron transfer as redox changes in the metal ion lead to structural rearrangement. Cheesman et al. (1990) identified the *P. aeruginosa* axial ligands as the thioether side chains of two methionine residues, a ligation scheme not previously reported for haem in any protein. Clone *X. albilineans* 8 showed 90% homology to the bacterioferritin from *P. aeruginosa* and may therefore contain similar axial ligands although future work will have to entail studying the bacterioferritin in *X. albilineans* in more detail with techniques such as electron paramagnetic resonance and near infrared magnetic circular dichroism spectra.

Clone *X. albilineans* 18 showed 85% homology to the plasmid pTOM9 from *Alcaligenes xylosoxidans*. The nickel-cobalt-cadmium resistance genes carried by plasmid pTOM9 of *A. xylosoxidans* are located on a 14.5 kb *Bam*HI fragment and contain two distinct nickel resistance loci, *ncc* and *nre* (Schmidt and Schlegel, 1994). The *ncc* locus causes a high-level nickel, cobalt, and cadmium resistance in strain AE104, which is a cured derivative of the metal-resistant bacterium *Alcaligenes eutrophus* and *E. coli*. Cells expressing *nreB* showed reduced accumulation of Ni (2+), suggesting that NreB mediated nickel efflux. From the function of the pTOM9 it may be possible that *X. albilineans* also possesses or may have a similar gene encoding for nickel efflux.

Clone *X. albilineans* 12 showed significant homology to the acetate repressor protein of *Salmonella typhimurium*. Studies on the acetate operon of *E. coli* are extensive (Cortay et al., 1991) and comparison of the deduced IclR aa sequence to that of *S. typhimurium* revealed that the two IclR repressors exhibit 89% identity. The *iclR* gene encodes a regulatory protein (repressor) for the *aceBK* operon. A search in protein data banks revealed that IclR has a score of 43.7% with GylR, a transcriptional regulator of the glycerol operon of *Streptomyces coelicolor*. It

may be possible that a similar operon is present in *X. albilineans* which may encode a regulatory protein with a similar function.

Many gene products of completely sequenced organisms are 'hypothetical'. They cannot be related to any previously characterized proteins. Therefore, their function is unknown. Hypothetical proteins are proteins that would appear to be encoded by genes that have been identified through analysis of DNA sequences emerging from genome sequencing efforts. Several clones showed similarities to hypothetical proteins. The elucidation of the function of these proteins will be intriguing and provides sufficient interest for speculation that these proteins may play a crucial role in pathogenicity of *X. albilineans*. Hence, this and similar studies demonstrate tremendous potential for laying the foundation for functional genomic studies.

SSH is a valuable technique in assessing genomic differences between individuals and for extracting sequences that are unique to the organism of interest. Correlation between phenotypic differences and gene differences is clear in many cases, such as pathogenic versus non-pathogenic strains. This information will provide a better view of diversity within a closely related group. The unique products can be used for complete sequencing of novel regions when used as hybridization probes to screen libraries, or sequence information can be used for the amplification and sequencing of flanking sequences (Arnold and Hodgson, 1991). Also, such surveys will increase the likelihood of finding species- or strain-specific regions useful for diagnostics.

Valuable and immediate benefit of strain-specific sequences would be to augment information contained within microarrays, thus greatly expanding the scope of either gene content or gene expression analyses. One or more near term benefits would be the development of strain-specific DNA-based diagnostic tools for rapid strain detection and identification. In the long term, strain-specific surface structures could be identified that would provide attractive targets for rapid antibody-based identification assays (Agron et al., 2002), and better knowledge of a common core set of genes within a species could

help in the development of new antibiotics and vaccines (Janke et al., 2001). Strain variation also provides valuable insights into evolutionary processes, finding sequences that are more common among strains enable more precise and reliable taxonomy. Moreover, studies of novel genes may help elucidate the basic biology of unique strain differences, leading to a more fundamental understanding of microbial diversity.

## ACKNOWLEDGEMENT

The authors acknowledge the National Research Foundation of South Africa for financial assistance.

## REFERENCES

- Agron PG, Macht M, Radnedge L, Skowronski EW, Miller W, Andersen GL (2002). Use of subtractive hybridization for comprehensive surveys of prokaryotic genome differences. *FEMS Microbiol. Lett.* 211: 175-182.
- Akopyants NS, Fradkov A, Diatchenko L (1998). PCR-based subtractive hybridization and differences in gene content among strains of *Helicobacter pylori*. *Proc. Natl. Acad. Sci USA* 95: 13108-13113.
- Arnold C, Hodgson IJ (1991). Vectors for PCR: a novel approach to genomic walking. *PCR Method Appl.* 1: 39-42.
- Bogush ML, Velikodvorskaya TV, Lebedev YB (1999). Identification and localization of differences between *Escherichia coli* and *Salmonella typhimurium* genomes by suppression subtractive hybridization. *Mol. Gen. Genet.* 262: 721-729.
- Cheesman MR, Graham GN, Watt GD (1990). Direct observation of bis-sulfur ligation to the heme of bacterioferritin. *Nature* 346: 771-773.
- Cortay JC, Negre D, Galinier A, Ducos B, Cozzone AJ (1991). Regulation of the acetate operon in *Escherichia coli*: purification and functional characterization of the IclR repressor. *EMBO J.* 10: 675-679.
- Davies J (1994). Inactivation of antibiotics and the dissemination of resistance genes. *Sci* 264: 375-382.
- Diatchenko L, Lau YF-C, Campbell AP (1996). Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc. Natl. Acad. Sci. USA* 93: 6025-6030.
- Emmerth M, Goebel W, Miller SI, Hueck CJ (1999). Genomic subtraction identifies *Salmonella typhimurium* prophages, F related plasmid sequences, and a novel fimbrial operon, *stf*, which are absent in *Salmonella typhi*. *J. Bact.* 181: 5652-5661.
- Groisman EA, Ochman H (1996). Pathogenicity islands: bacterial evolution in quantum leaps. *Cell* 87: 791-794.
- Hacker J, Blum-Oehler G, Muhldorfer I, Tschape H (1997). Pathogenicity islands of virulent bacteria: structure, function, and impact on microbial evolution. *Mol. Microbiol.* 23: 1089-1097.
- Janke B, Dobrindt U, Hacker J, Blum-Oehler GA (2001). Subtractive hybridization analysis of genomic differences between the uropathogenic *E. coli* strain 536 and the *E. coli* K-12 strain MG1655. *FEMS Microbiol. Lett.* 199: 61-66.
- King G, Murray NE (1994). Restriction enzymes in cells, not Eppendorfs. *Trend. Microbiol.* 12: 465-469.
- Laulhere JP, Neema C, Expert D (1992). Iron deficiency induced by chrysobactin in *Saintpaulia* leaves inoculated with *Erwinia chrysanthemi*. *Plant Physiol.* 102: 967-973.
- Lawrence J, Ochman H (1997). Amelioration of bacterial genomes: rates of change and exchange. *J. Mol. Evol.* 44: 383-397.
- Perna NT (2001). Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature*, 409: 529-533.
- Radnedge L, Gamez-Chin S, McCready PM, Worsham PL, Andersen GL (2001). Identification of nucleotide sequences for the specific and rapid detection of *Yersinia pestis*. *Appl. Environ. Microbiol.* 67: 3759-3762.
- Rebrikov DV, Britanova OV, Gurskaya NG, Lukyanov KA (2000). Mirror orientation selection (MOS): a method for eliminating false positive clones from libraries generated by suppression subtractive hybridization. *Nucleic Acids Res.* 28: 90.
- Sauerbaum S, Achtman M (1999). Evolution of *Helicobacter pylori*: the role of recombination. *Trends Microbiol.* 7: 182-184.
- Schmidt T, Schlegel HG (1994). Combined nickel-cobalt-cadmium resistance encoded by the ncc locus of *Alcaligenes xylosoxidans* 31A. *J. Bact.* 176: 7045-7054.
- Stroeher UH, Manning PA (1997). Novel *Vibrio cholerae* O139 genes involved in lipopolysaccharide biosynthesis. *J. Bact.* 179: 2740-2747.
- Zhang Q, Melcher U, Zhou L, Najjar FZ, Roe BA, Fletcher J (2005). Genomic comparison of plant pathogenic and non-pathogenic *Serratia marcescens* strains by suppression subtractive hybridization. *Appl. Environ. Microbiol.* 71: 7716-7723.