Published online January 12, 2004

Nucleic Acids Research, 2004, Vol. 32, No. 1 ***e7***
DOI: 10.1093/nar/gnh012

PCR-based positive hybridization to detect genomic diversity associated with bacterial secondary metabolism

Francesco Pomati and Brett A. Neilan*

Cyanobacteria and Astrobiology Research Laboratory, School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney 2052, NSW, Australia

Received October 27, 2003; Revised November 14, 2003; Accepted November 27, 2003 DDBJ/EMBL/GenBank accession nos+

ABSTRACT

A PCR-based positive hybridization (PPH) method was developed to explore toxic-specific genes in common between toxigenic strains of *Anabaena circinalis*, a cyanobacterium able to produce saxitoxin (STX). The PPH technique is based on the same principles of suppression subtractive hybridization (SSH), although with the former no driver DNA is required and two tester genomic DNAs are hybridized at high stringency. The aim was to obtain genes associated with cyanobacterial STX production. The genetic diversity within phylogenetically similar strains of *A.circinalis* was investigated by comparing the results of the standard SSH protocol to the PPH approach by DNA-microarray analysis. SSH allowed the recovery of DNA libraries that were mainly specific for each of the two STX-producing strains used. Several candidate sequences were found by PPH to be in common between both the STX-producing testers. The PPH technique performed using unsubtracted genomic libraries proved to be a powerful tool to identify DNA sequences possibly transferred laterally between two cyanobacterial strains that may be candidate(s) in STX biosynthesis. The approach presented in this study represents a novel and valid tool to study the genetic basis for secondary metabolite production in microorganisms.

INTRODUCTION

One of the most striking aspects of bacterial secondary metabolites is the unforseen ways in which some of these compounds are biosynthesized. These natural products, besides attracting attention because of their potential or actual toxicity, also represent a potentially rich source of new drug candidates. Many bacterial metabolites have chemical structures unprecedented in other aquatic or terrestrial organisms, and display new and interesting biological activity. In addition, the synthetic capability of bioactive metabolite production is not universal to a species, but is exceptional and often limited to certain strains.

Here we have utilized a PCR-based subtractive hybridization approach to explore the genetic diversity associated with the production of saxitoxin (STX) within phylogenetically closely related strains of toxic and non-toxic *Anabaena circinalis*. *Anabaena circinalis* is a common, filamentous, bloom-forming cyanobacterium, found to produce paralytic shellfish poisoning (PSP) toxins in Australian freshwaters (1,2). PSP toxins, of which the most potent representative is STX, are a class of neurotoxic alkaloid that selectively block voltage-gated Na^+ channels in excitable cells, thereby affecting neural impulse generation in animals (3). They occur naturally in red tides composed of marine dinoflagellates (4–6), in filamentous cyanobacteria (7–10), and in certain heterotrophic bacteria (11). Although STX has been studied for more than 30 years, the molecular basis for the synthesis of this peculiar alkaloid is currently unknown.

PCR-based genomic subtraction, also called suppression subtractive hybridization (SSH), has been developed for the rapid identification of differences among pathogenic bacteria and was first applied to strains of *Helicobacter pylori* (12). SSH has also been employed to compare the genomes of virulent and avirulent strains of the aquatic pathogen *Aeromonas hydrophila* (13), and to identify genomic differences between uropathogenic and non-pathogenic *Escherichia coli* strains (14). The SSH technique has been successfully utilized to detect pathogenicity islands in infectious bacteria (15). It has also being used to derive information, via genomic diversity, regarding divergence in ecologically relevant genetic adaptations in closely related strains of *Pseudomonas fluorescens* and *Thermotoga maritima* (16,17). In this study, SSH was used to identify potential differences between the genomes of two STX-producing and two non-toxic strains of *A.circinalis*. The suppression technique was modified to include a third hybridization step, termed PCR-based positive hybridization (PPH), and designed to recover toxic strain-specific genome fragments in common between the two STX-producing strains (Fig. 1). The aim was to obtain genes associated with STX production in *A.circinalis*. In concordance with previous work (15), the toxic specificity of tester DNA fragments from the SSH and PPH libraries was analyzed by quantitative DNA microarray hybridization with

*To whom correspondence should be addressed. Tel: +61 2 9385 3235; Fax: +61 2 9385 1591; Email: b.neilan@unsw.edu.au
+AY445143–AY445188

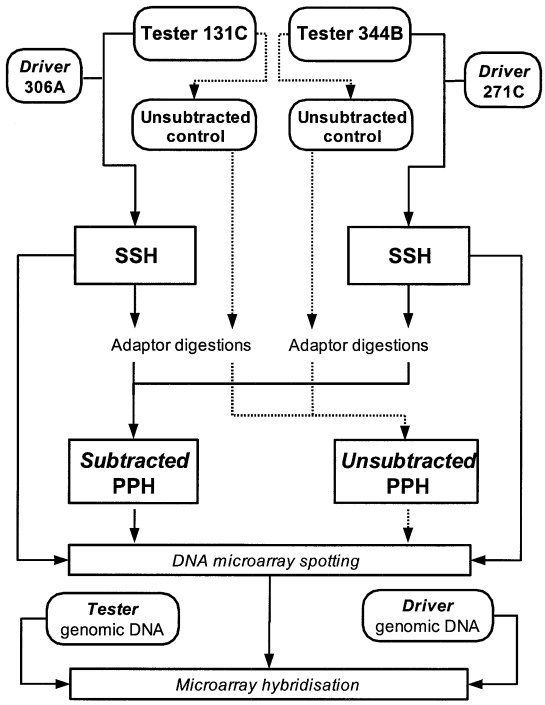

Figure 1. Schematic diagram of the procedure applied in this study. Toxic *A.circinalis* strains were used as testers, while non-toxic strains were used as drivers. Dashed lines indicate unsubtracted tester DNA.

labeled toxic/non-toxic *A.circinalis* genomic DNA, and verified by PCR amplification.

MATERIALS AND METHODS

Cyanobacterial strains and growth conditions

The cyanobacterial strains investigated in this study are listed in Table 1. PSP toxin-producing and non-toxic isolates of *A.circinalis* were obtained from the Australian Water Quality Centre (Adelaide) and maintained in Jaworski's Medium (1). Cyanobacterial cultures were grown without agitation or aeration in glass 250-ml flasks at a constant temperature of 26°C, under continuous irradiance of cool white light at an intensity of 15 μmol photon/m^2/s. Cultures in mid-exponential growth phase were used for DNA extraction.

DNA extraction

Cyanobacterial cultures were filtered through a 3.0 μm pore size filter (Millipore, Billerica, MA) and cells were washed twice with sterile water. This method has been previously demonstrated to be effective in the removal of possible contaminating heterotrophic bacteria (18–20). Genomic DNA was extracted from filtered and washed cyanobacterial cells using the XS procedure, as described previously (21), and resuspended in TE buffer (10 mM Tris–HCl pH 7.4, 1 mM EDTA pH 8).

Suppression subtractive hybridization

Subtraction of cyanobacterial genomes was achieved by means of the modified PCR-based subtractive hybridization protocol described by Akopyants *et al.* (12). One microgram of genomic *A.circinalis* DNA from strains 306A and 271C (non-toxic), and 131C and 344B (toxic) was digested with RsaI, extracted with phenol, and precipitated with ethanol. Subtraction experiments were carried out using *A.circinalis* strains 306A (non-toxic) as the driver and 131C (toxic) as the tester in one experiment, and *A.circinalis* strains 271C (non-toxic) as the driver and 344B (toxic) as the tester in another experiment. Briefly, for each experiment, two different PCR adaptors were ligated to two different aliquots of the restriction-digested tester DNA. Two hybridizations were then performed (Fig. 2A). In the first hybridization, an excess of driver DNA was added to each of the adaptor-ligated tester DNAs. Each sample mixture was then denatured at 98°C for 2 min and allowed to reanneal at 63°C for 90 min. This hybridization enables single-stranded DNA to be enriched for tester-specific DNA, as DNA fragments that are not tester specific will form hybrid molecules with the driver DNA. In the second hybridization, the two primary hybridization reaction mixtures were combined without denaturing and allowed to anneal at 63°C overnight. Only the subtracted, single-stranded, tester-specific DNA should reassociate to make hybrids with the two different terminal adaptors. Molecules with different adaptors at each end were amplified exponentially using PCR primers to the two adaptor sequences. In this study, PCR amplifications were performed essentially as reported elsewhere (12), with the following modification: both primary and secondary PCRs were carried out in a Perkin-Elmer (Shalton, CT) GeneAmp system 2400 for 25 cycles using 1 U *Taq* F1 polymerase (Promega, Madison, WI) per reaction. The putative tester-specific libraries were amplified using, as a positive control, aliquots of unsubtracted tester DNA ligated with both terminal adaptors (unsubtracted tester control; Fig. 2A). PCR products obtained after SSH were cloned into the pGEM-TE vector (Promega). Clones were amplified using the pGEM-TE vector-specific primers (mpF and mpR) and sequenced using PRISM Automated BigDye terminator sequencing and an ABI 373 sequencer (PE Applied Biosystems, Foster City, CA), with reactions performed using 3 μl (~150 ng) of each PCR product and 10 pmol of each appropriate primer in a half-scale reaction as specified by the manufacturer. Sequences were analyzed using the BLAST program (http://www.ncbi.nlm.nih.gov/BLAST/).

Table 1. Cyanobacterial strains used in this study

Strain	PSP-toxin production	Reference
A.circinalis 131C	Yes	(2)
A.circinalis 344B	Yes	P. Baker, personal communication
A.circinalis 134C	Yes	(2)
A.circinalis 279B	Yes	(2)
A.circinalis 306A	No	(2)
A.circinalis 271C	No	P. Baker, personal communication

PCR-based positive hybridization

The PPH step was performed on subtracted and unsubtracted libraries from *A.circinalis* 131C and 344B (Fig. 1). PPH was achieved using the same principles as the SSH protocol (12).

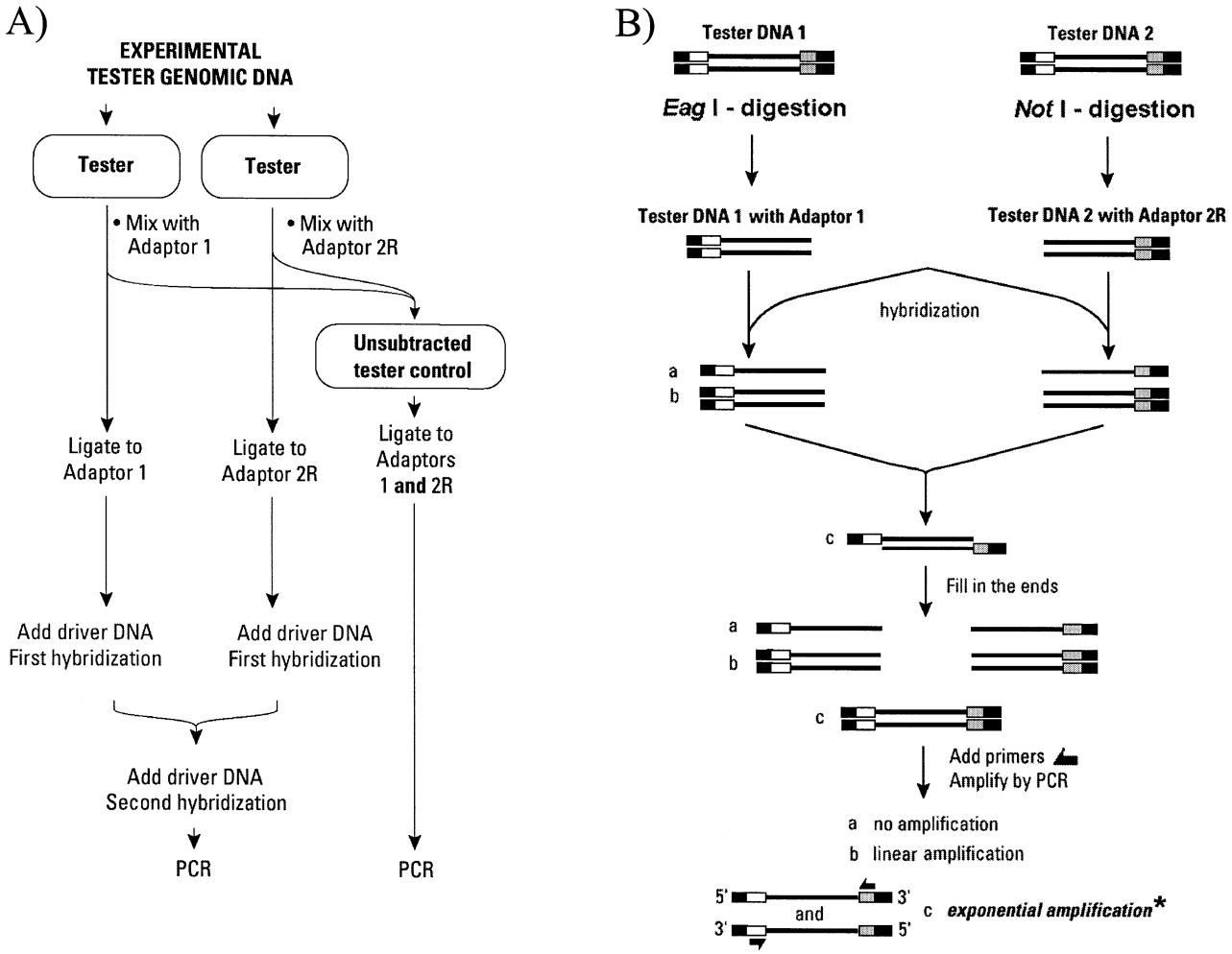

Figure 2. (**A**) Standard SSH procedure, as described previously by Akopyants *et al.* (12). (**B**) PPH method, based on adaptor-mediated PCR. Solid lines represent tester DNA fragments, boxes represent terminal adaptors 1 and 2R. Adaptor 2R was excised from tester DNA 1 (strain 131C) by digestion with EagI, and adaptor 1 was removed from tester DNA 2 (strain 344B) with NotI. *Although there is an identical primer binding sequence on both ends of the type c molecules, the short overall homology at the two ends negates the suppression PCR effect (12).

With this novel technique, however, no driver DNA was required and two tester DNAs were hybridized at high stringency, with each tester DNA ligated to a different terminal adaptor (Fig. 2B). After hybridization, only common single-stranded sequences can reassociate to make hybrids with different adaptors at each terminus. Only DNA fragments with both of the terminal adaptors can be amplified exponentially using PCR primers to the two adaptor sequences (Fig. 2B). The procedure was carried out as follows. The second PCR amplification of the standard SSH protocol (a total of 50 µl reaction volume) was performed and DNA ethanol precipitated. The DNA libraries thereby obtained (subtracted 131C, unsubtracted 131C, subtracted 344B, unsubtracted 344B) were resuspended in 40 µl of water with 10 µl aliquots (~1.5 µg) utilized for restriction enzyme digestions (Fig. 2B). In strain 131C, the adaptor 2R was excised from subtracted DNA and unsubtracted tester control by digestion with EagI (New England Biolabs, Beverly, MA), while adaptor 1 was removed from strain 344B subtracted and unsubtracted DNA using NotI (New England Biolabs). Together with the tester DNA, 50 µl reactions contained 15 U of restriction enzyme, 5 µl 10× NEB3 buffer (New England Biolabs) and 5 µg of bovine serum albumin. Reactions were incubated at 37°C overnight, heat inactivated at 65°C for 10 min, and the DNA was extracted with phenol, precipitated with ethanol, and resuspended in 13 µl of water. The positive hybridization was performed as shown in Fig. 2B, combining the two subtracted libraries in one experiment and the unsubtracted tester controls in another (Fig. 1). In a final 5 µl mixture, 1.5 µl aliquots of both 131C DNA (tester DNA 1, adaptor 1) and 344B DNA (tester DNA 2, adaptor 2R) were combined with 1 µl of 5× hybridization buffer (250 mM HEPES–Cl pH 8.1, 2.5 M NaCl, 1 mM EDTA pH 8.0) and 1 µl water. The DNA was denatured at 98°C for 2 min and the hybridization reactions were incubated at 65°C for 1 h. Subsequently, 100 µl of dilution buffer (20 mM HEPES–Cl pH 8.1, 50 mM NaCl, 0.2 mM EDTA pH 8.0) were added to each mixture and the hybridization was completed by incubation at 65°C overnight. One microliter of each sample was diluted in 100 µl of water, and 1 µl of this diluted hybridization mixture was used as template for the successive round of PCR amplification. PCR was performed essentially

as the secondary amplification of the SSH protocol, with the addition of an initial step at 72°C for 5 min to allow the *Taq* polymerase to fill any overhanging ends of hybrid DNA fragments. PCR products obtained after PPH for both subtracted and unsubtracted libraries were cloned and sequenced as described above.

PCR amplifications

Amplification of sequences 179 and 109 was performed using 20 pmol each of the specific primers 179F (5′-AATAC-CAATGCCTCCACTCC-3′) and 179R (5′-AAAGACGGT-GAAACACCTGC-3′), or 109F (5′-ACAGGTGCGATGC-GACCATT-3′) and 109R (5′-TTGGTTGAGTGCGCT-CCAAC-3′), in 20 µl reactions containing 200 µM dNTP, 2.5 mM $MgCl_2$, *Taq* polymerase buffer, 100 ng of genomic DNA template and 0.25 U *Taq* polymerase. PCRs were cycled using a temperature profile of 94°C, 3 min, followed by 30 cycles of 94°C for 10 s, 55°C for 20 s, and 72°C for 30 s, and concluded with one cycle of 72°C for 5 min. All results were visualized by 2% agarose gel electrophoresis in Tris-acetate-EDTA buffer according to standard protocols (22).

Microarray design and production

Candidate SSH and PPH clones were amplified, purified using 96-well multi-screening membrane plates (Millipore), and resuspended in 70 µl water. Purified PCR products were diluted to a final concentration of ~200 ng/µl in 150 mM $NaPO_4$ (pH 8.0). Six microliters of each sample was then transferred to a 384-well microplate for printing. Slide preparation and printing were performed at the Ramaciotti Centre for Gene Function Analysis (University of New South Wales, Sydney, Australia). DNA samples were arrayed with four pins (SMP-3 Stealth; ArrayIt, Sunny-vale, CA) at a spacing distance of 180 µm on polysine-coated 25 × 75 mm glass slides (Menzel-Glaser, Braunscheig, Germany) using a ChipWriterPro robotic printer (Bio-Rad, Hemel Hempstead, UK) under conditions of 65% relative humidity. The PCR clones obtained in this study were spotted together with >200 other DNA fragments, including 16S rDNA of the investigated strains as housekeeping genes (BGGM[1] microarray). The BGGM[1] array list of genes is available online (http://149.171.168.73/microarray/). Each gene probe was spotted in duplicate at a volume of 600 pl (spot diameter of 100 µm), and each glass slide contained two copies of the BGGM[1] array. Following printing, slides were allowed to age for 3 days before post-processing. DNA microarrays were then baked at 80°C for 30 min and kept in the dark for long-term storage.

Before use, gene arrays were rehydrated over a 40°C water bath for 15 s and dried on a heating block at 90°C for 5 s. DNA on the microarrays was fixed by ultra violet (UV) cross-linking at 60 mJ in a Hoefer UVC UV Crosslinker 115 VAC (Amersham Pharmacia Biotech, Piscataway, NJ), and washed once in 0.2% SDS and twice with water. The glass slides were then treated with 6 g succinic anhydride (Sigma) dissolved in 350 ml of 1-methyl-2-pyrrolidinone (Aldrich) and 15 ml of 1 M boric acid (pH 8.0) (Aldrich). Immediately following blocking, the DNA was denatured by immersing the slides in Milli-Q water at 95°C for 2 min. The microarrays were then rinsed briefly in 95% ethanol at 4°C for 1 min, spun dry at 500 *g* for 5 min and stored dry in the dark.

Labeling of genomic DNA

Fluorescently labeled DNA was prepared indirectly by incorporating amino-allyl (aa)-dUTP followed by coupling with fluorescent dyes. One microgram of genomic *A.circinalis* DNA from strains 271C and 306A (non-toxic), and 131C and 344B (toxic) was prepared by digestion with RsaI, extracted with phenol and precipitated with ethanol. aa-dUTP labeling of cyanobacterial genomic DNA was achieved as follows. DNA was diluted in 38 µl of water, denatured at 99°C for 10 min and chilled on ice. Five microliters of 10× NEB labeling buffer (New England Biolabs), 3 µl of amino-allyl dNTP mix (3 mM dGTP, dATP and dCTP, 1.8 mM aa-dUTP, 1.2 mM dTTP) and 2 µl (10 U) of the large Klenow fragment of DNA polymerase I (Promega) were combined in a final reaction volume of 50 µl. The reaction mixture was incubated at 37°C for 2 h and unincorporated amines were removed by QIAquick purification (Qiagen). DNA samples were dried in a speed-vac, resuspended in 9 µl of 0.1 M $NaHCO_3$, pH 9.0, and added to 2 µl of prepared Cy3 or Cy5 dye aliquots (Amersham Pharmacia Biotech). Reactions were incubated for 60 min at room temperature in the dark and purified (QIAquick PCR Kit). Labeled DNA samples were dried to ~20 µl, combined according to the different experiments, and evaporated to dryness.

Microarray hybridization

Each microarray hybridization was performed in duplicate. Cy3-labeled genomic DNA (green) from the non-toxic strains 271C and 306A was hybridized with Cy5-labeled genomic DNA (red) from the toxic strains 344B and 131C, respectively. For each single hybridization, fluorescently labelled DNA was resuspended in 20 µl of hybridization solution containing DIG Easy buffer (Roche Applied Science, Penzberg, Germany) and 500 µg of yeast tRNA (Sigma), denatured for 2 min at 99°C, cooled to ambient temperature and applied to the microarray slide. Glass coverslips (22 × 22 mm) were placed over the solution and hybridization was performed overnight at 37°C in a water-tight humidified hybridization chamber. Array slides were washed in two stages: three washes at 50°C for 15 min with pre-heated 1× SSC buffer 0.1% SDS, followed by three rinses at room temperature with decreasing concentrations of SSC buffer (0.5×, 0.25× and 0.1× in Milli-Q water). Slides were spun dry at 500 *g* for 5 min, and kept in the dark prior to scanning.

Microarray scanning, data acquisition and statistical analyses

Clean slides were scanned with the ArrayWorx 'e' Microarray Scanner from Applied Precision, Inc. (Issaquah, WA). Scanned slide images were generated automatically with ArrayWorx software provided with the ArrayWorx Scanner. Images were quantified using the GenePix Pro software from Axon Instruments (Foster City, CA). Erroneous spots were flagged manually and removed from the final data set. All microarray data were filtered to remove any spots in which <60% of the signal pixels exceeded the local background value for both lasers (595 and 685 nm). The median Cy5:Cy3 ratio automatically generated by the GenePix Pro software for the filtered data for each spot was used for subsequent analysis. The ratio of medians was normalized to give a ratio

Table 2. *Anabaena circinalis* 131C putative specific sequences with protein matches in The National Center for Biotechnology Information (NCBI) protein database

ID	Best BLASTX hit	Organism	% Identity	% Similarity	Microarray hybridization	Specificity[a]
SSH						
65	Hypothetical protein	*Cytophaga hutchinsonii*	57	81	5.3 ± 0	+
70	TonB, ferric-siderophore uptake	*B.thetaiotaomicron* VPI-5482	42	66	4.9 ± 0.3	+
73	Alpha-amylase precursor	*C.acetobutylicum*	47	58	4.6 ± 2.6	+
71	$NADH_2$ dehydrogenase	*C.elongatum*	70	75	4.4 ± 0.7	+
175	GLP_291_11778_8566	*G.lamblia* ATCC 50803	29	43	4.5 ± 0.8	+
75	No similarity				3.9 ± 1.8	+
74	Hypothetical protein	*Neurosporra crassa*	46	66	3.6 ± 1.9	+
69	Hypothetical protein	*B.japonicum* USDA 110	62	83	2.9 ± 0	+
77	Sulfatase family protein	*N.aromaticivorans*	42	57	2.8 ± 0.5	+
68	Hypothetical protein	*Naromaticivorans*	91	95	2.7 ± 0.2	+
72	Putative integral membrane protein	*S.coelicolor* A3(2)	46	61	2.6 ± 0.5	+
Subtracted PPH						
104	Hypothetical protein, transcriptase	*Caenorhabditis elegans*	32	48	6.8 ± 0.3	+
95	Putative hydrolase	*B.thetaiotaomicron* VPI-5482	53	67	4.1 ± 0.3	+
185	Novel antigenic to ORF2	*T.orientalis*	94	94	1.8 ± 0	≈
120	Hypothetical protein	*Nostoc* sp. PCC7120	87	90	1.7 ± 0.3	≈
119	50S ribosomal protein L33	*P.lunula*	80	96	1.6 ± 0.1	≈
99	Hypothetical protein	*N.punctiforme*	63	80	1.4 ± 0.5	≈
186	Thiamin-phosphate pyrophosphorylase	*Nostoc* sp. PCC7120	70	81	1.2 ± 0.4	≈
98	RNA polymerase ECF-type sigma factor	*B.thetaiotaomicron* VPI-5482	38	61	1.1 ± 0	≈
102	Unknown protein	*Nostoc* sp. PCC7120	80	100	0.9 ± 0.2	≈
97	Putative gluconate aldolase	*Microscilla* sp. PRE1	33	56	ND	ND
106	Hypothetical protein	*C.hutchinsonii*	67	72	ND	ND
187	ABC transporter ATP-binding	*B.thetaiotaomicron* VPI-5482	60	70	ND	ND
189	Phosphoglucomutase–phosphomannomutase	*C.hutchinsonii*	58	69	ND	ND
Unsubtracted PPH						
108	Hypothetical protein	*C.hutchinsonii*	41	61	6.6 ± 1.9	+
179	Putative transposase	*Synechocystis* sp. BO8402	35	60	5.6 ± 1.1	+
109	Hypothetical protein	*B.thetaiotaomicron* VPI-5482	23	43	4.9 ± 0.3	+
112	Hypothetical protein	*T.erythraeum* IMS101	44	65	3.5 ± 1.8	+
115	No similarity				3.4 ± 0.2	+
111	Cell division GTPase	*C.hutchinsonii*	56	72	2.4 ± 0.8	+
107	No similarity				2.4 ± 0	+
183	Acetyltransferase	*B.thetaiotaomicron* VPI-5482	60	82	2.4 ± 0.7	+
178	Hypothetical protein	*T.fusca*	35	57	2.1 ± 0.4	+
118	60 kDa chaperonin GroEL	*C.hutchinsonii*	91	100	2 ± 0.4	+
110	Hypothetical protein, putative protease	*C.hutchinsonii*	35	65	ND	ND
180	DNA polymerase III alpha subunit	*B.thetaiotaomicron* VPI-5482	78	84	ND	ND
182	Transport protein	*X.axonopodis* pv. citri 306	28	46	ND	ND

[a]Toxic-strain specificity.
+, Toxic-strain specific; ≈, unspecific; ND, not detected.
Microarray experiments represent hybridizations between strains 131C (Cy5) and 306A (Cy3) using genomic DNA. Microarray values are expressed as average ± standard error of Cy5:Cy3 normalized ratios. Toxic- and non-toxic-strain specificity levels are defined by Cy5:Cy3 ratios >2 and <0.5, respectively.

measurement of 1 for the control sequences corresponding to the tested strains 16S rRNA gene.

Nucleotide sequence accession numbers

The sequences determined in this study have been submitted to the DDBJ/EMBL/GenBank nucleotide sequence database (accession nos AY445143–AY445188).

RESULTS

Suppression subtractive hybridization

Putative DNA sequences specific to the toxic *A.circinalis* strains 131C and 344B were cloned and the insert size of 50 randomly selected clones per library was estimated by PCR amplification. Insert sizes varied from 0.15 to 0.5 kb and, for each experiment, a total of 20 cloned DNA fragments of different sizes were purified and sequenced. Of the 40 clones analyzed comprising toxic-specific fragments from strains 131C and 344B, seven sequences (17.5%) were encountered more than once (17.5%). General features of these SSH DNA fragments are reported in Tables 2 and 3 for *A.circinalis* 131C and 344B, respectively. Each library was found to be composed of 11 putative toxic-specific sequences, with strain 131C fragments mainly characterized by hypothetical proteins of unknown function (Table 2) and an average mol% GC of 58.4. Toxic sequences specific to strain 344B, however, comprised hypothetical proteins of unknown function together with defined enzymes and membrane proteins (Table 3), with an average mol% GC of 47.5.

Table 3. *Anabaena circinalis* 344B putative specific sequences with protein matches (Id: identity; Sim: similarity) in the NCBI protein database

ID	Best BLASTX hit	Organism	% Identity	% Similarity	Microarray hybridization	Specificity[a]
SSH						
91	Magnesium chelatase, subunit I	*B.thetaiotaomicron* VPI-5482	55	84	8.7 ± 0.2	+
89	Excinuclease ABC subunit A	*C.hutchinsonii*	47	66	7.6 ± 1.3	+
88	50S ribosomal protein L33	*P.lunula*	78	93	6.8 ± 0.6	+
173	Hypothetical protein	*N.aromaticivorans*	61	73	6.2 ± 0.2	+
84	Penicillin-binding protein	*Bacillus subtilis*	51	59	4.9 ± 0.1	+
90	Hypothetical protein	*B.thetaiotaomicron* VPI-5482	36	59	4.1 ± 0	+
85	Hypothetical protein	*Arabidopsis thaliana*	56	77	4 ± 0.1	+
86	Similar to chloride channel	*N.punctiforme*	73	85	2.3 ± 0	+
87	Thiamin-phosphate pyrophosphorylase	*Nostoc* sp. PCC7120	77	89	0.8 ± 0.3	≈
170	Unknown protein	*Oriza sativa*	28	41	ND	ND
174	Putative hydrolase	*B.thetaiotaomicron* VPI-5482	48	67	ND	ND
Subtracted PPH						
95	Putative hydrolase	*B.thetaiotaomicron* VPI-5482	53	67	8.2 ± 0.7	+
119	50S ribosomal protein L33	*P.lunula*	80	96	6.2 ± 0.6	+
98	RNA polymerase ECF-type sigma factor	*B.thetaiotaomicron* VPI-5482	38	61	5.9 ± 0.4	+
104	Hypothetical protein, transcriptase	*C.elegans*	32	48	4.1 ± 0.3	+
106	Hypothetical protein	*C.hutchinsonii*	67	72	4.1 ± 0.1	+
185	Novel antigenic ToORF2	*T.orientalis*	94	94	2.2 ± 0.2	+
99	Hypothetical protein	*N.punctiforme*	63	80	1.3 ± 0.1	≈
186	Thiamin-phosphate pyrophosphorylase	*Nostoc* sp. PCC7120	70	81	0.8 ± 0.2	≈
102	Unknown protein	*Nostoc* sp. PCC7120	80	100	0.7 ± 0	≈
97	Putative gluconate aldolase	*Microscilla* sp. PRE1	33	56	ND	ND
120	Hypothetical protein	*Nostoc* sp. PCC7120	87	90	ND	ND
187	ABC transporter ATP-binding	*B.thetaiotaomicron* VPI-5482	60	70	ND	ND
189	Phosphoglucomutase–phosphomannomutase	*C.hutchinsonii*	58	69	ND	ND
Unsubtracted PPH						
179	Putative transposase	*Synechocystis* sp. BO8402	35	60	22.6 ± 2	+
108	Hypothetical protein	*C.hutchinsonii*	41	61	18.7 ± 0.4	+
107	No similarity				13.8 ± 0.7	+
111	Cell division GTPase	*C.hutchinsonii*	56	72	11.6 ± 1.7	+
178	Hypothetical protein	*T.fusca*	35	57	11 ± 1.9	+
109	Hypothetical protein	*B.thetaiotaomicron* VPI-5482	23	43	8.6 ± 0.1	+
115	No similarity				6.6 ± 1.6	+
182	Transport protein	*X.axonopodis* pv. citri 306	28	46	6.5 ± 1.2	+
183	Acetyltransferase	*B.thetaiotaomicron* VPI-5482	60	82	1.6 ± 0.3	≈
110	Hypothetical protein	*C.hutchinsonii*	35	65	ND	ND
112	Hypothetical protein	*T.erythraeum* IMS101	44	65	ND	ND
118	60 kDa chaperonin GroEL	*C.hutchinsonii*	91	100	ND	ND
180	DNA polymerase III alpha subunit	*B.thetaiotaomicron* VPI-5482	78	84	ND	ND

[a]Toxic-strain specificity.
+, Toxic-strain specific; ≈, unspecific; ND, not detected.
Genomic DNA experiments represent microarray hybridizations between strains 344B (Cy5) and 271C (Cy3). Microarray values are expressed as in Table 2.

PCR-based positive hybridization

Toxic-specific DNA sequences common to both *A.circinalis* strains 131C and 344B were identified by the PPH of subtracted SSH libraries (subtracted PPH) and unsubtracted tester controls (unsubtracted PPH). The insert size of 50 randomly selected clones per library was estimated and, for each experiment, a total of 20 polymorphic DNA fragments were analyzed. These sequences showed a mol% GC of 46.1 for the PPH of subtracted libraries, and 51.4 for the positive hybridization of unsubtracted tester DNA. Among the 40 clones analyzed, 13 distinct fragments were found to characterize each experiment. Two sequences were encountered more than once in the subtracted-PPH library, as were three in the unsubtracted PPH. No sequence was found to be in common between the subtracted- and unsubtracted-PPH libraries. Features of the putative toxic-specific sequences in common between *A.circinalis* 131C and 344B are reported in Tables 2 and 3, with the two groups of clones comprising predicted enzymes, regulatory proteins and hypothetical proteins.

DNA microarray analysis of SSH and PPH libraries

All fragments recovered from *A.circinalis* 131C by SSH were also shown to be toxic strain-specific by microarray hybridization using genomic DNA from strains 131C and 306A (Table 2). Among the library of putative toxic strain-specific sequences obtained from strain 344B, one (clone 87) was found to be a false positive, while another two fragments (clones 170 and 174) could not be detected by genomic DNA-microarray hybridization of strains 344B and 271C (Table 3). Two SSH fragments from strain 344B (clones 88 and 89) demonstrated specificity also for strain 131C.

In the same microarray experiments, only two DNA fragments belonging to the subtracted-PPH library were shown to be toxic strain-specific and common to strains 131C and 344B (clones 95 and 104). Most of the other

subtracted-PPH sequences were specific only for strain 344B, while clones 97, 187 and 189 could not be detected by microarray hybridization of the cyanobacterial digested genomes. Two subtracted-PPH genes from strain 344B, a 50S ribosomal protein L33 (clone 119) and a thiamin-phosphate pyrophosphorylase (clone 186) also recovered by SSH, demonstrated to be false positives, together with fragments 99 and 102 (Tables 2 and 3).

Out of 10 distinct fragments, seven sequences obtained by unsubtracted PPH were found to be toxic strain-specific and in common to both *A.circinalis* 131C and 344B (clones 107, 108, 109, 111, 115, 178 and 179). Two unsubtracted-PPH fragments (clones 110–180) could not be detected (Tables 2 and 3).

PCR amplification of putative, unsubtracted-PPH, toxic-specific sequences

The toxic-strain specificity of two unsubtracted-PPH fragments among the seven toxic-specific sequences in common between strains 131C and 344B was verified by designing primers to clones 179 and 109. Amplicons of the expected size (164 and 298 bp for clones 179 and 109, respectively) were observed only for the two toxic strains 131C and 344B, and not for either of the other toxic (134C and 279B) or non-toxic (306A and 271C) *A.circinalis* isolates (Fig. 3). Similar results were also obtained for other unsubtracted-PPH fragments (data not shown), including clones 111 (cell division GTPase) and 107 (no similarity), found by microarray hybridization to be toxic-strain specific.

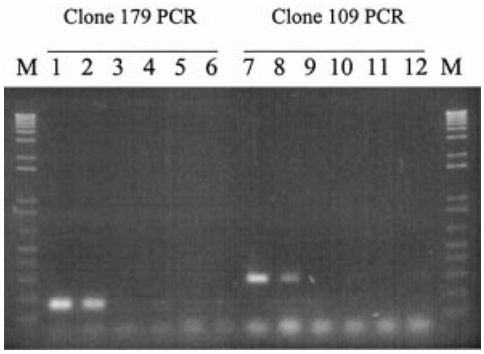

Figure 3. Electrophoretic comparison of fragments 179 (lanes 1–6) and 109 (lanes 7–12) amplified by PCR from toxic and non-toxic strains of *A.circinalis*. Gel lanes 1–12 as follows: 131C (1 and 7), 344B (2 and 8), 134C (3 and 9), 279B (4 and 10), 306A (5 and 11) and 271C (6 and 12). PCR products were loaded as a total of 4 µl per sample and run with 1 kb Plus DNA Ladder (Invitrogen) as standard (lanes M).

DISCUSSION

In this study we used a combined approach of subtractive and positive hybridizations to investigate the molecular basis for STX production in *A.circinalis*. The two groups of clones obtained by SSH, though not exhaustive of the entire pool, were sufficiently representative of the SSH library when compared with previous studies (12). The majority of DNA fragments recovered by SSH were toxic strain-specific but unique to each tester strain, either 131C or 344B, and not common to both isolates (Tables 2 and 3) or to other STX-producing strains of *A.circinalis* (data not shown). Therefore, a novel molecular tool was developed in this study, the PPH technique, with the aim of recovering toxin biosynthetic genes, hypothesized to be in common between the two tester *A.circinalis* strains. Subtracted-PPH libraries identified two toxic-specific DNA fragments common to both 131C and 344B (clones 95 and 104). Most of the subtracted-PPH clones had a low toxic-strain specificity for *A.circinalis* 131C (Table 3). Since SSH genes were found to be mainly characteristic of each distinct strain, it is possible that the percentage of common sequences in the final positive hybridization reaction was relatively low. The subsequent single round of PCR may not, therefore, have sufficiently enriched sequences common to the two SSH libraries, thus resulting in the high percentage of false positives among the characterized libraries of subtracted-PPH clones. On the other hand, PPH of unsubtracted tester controls resulted in seven toxic strain-specific clones, with no fragment found to be in common to both subtracted and unsubtracted PPH, indicating no cross contamination between the two libraries (Tables 2 and 3). The subtracted-PPH library, compared with the unsubtracted PPH, showed a higher percentage of false positives and the presence of SSH sequences, suggesting that this technique was less effective than the PPH of unsubtracted tester DNA. Moreover, PCR amplification of unsubtracted-PPH fragments, including clones 109 and 179 (Fig. 3), indicated that similar genomic loci between two *A.circinalis* strains may not be in common with otherwise phylogenetically closely related (18) toxic and non-toxic strains of the same species. This consideration suggests that the unsubtracted-PPH method may have detected lateral gene transfer (LGT) events that have occurred between strains 131C and 344B of *A.circinalis*. Alternatively, other phenomena may explain these results rather than LGT, such as unusual rates of evolution or gene loss (23) in *A.circinalis* isolates.

However, during the course of the present investigation, we found an average mol% GC of 50 among all the fragments recovered by SSH and PPH. We estimated an average mol% GC of 43 in the genome of *A.circinalis* (F.Pomati and R.Kellmann, unpublished results), which is consistent with what has been reported for other cyanobacteria of the order *Nostocales*, including *Nostoc* PCC7120 (41%) and *Nostoc punctiforme* PCC73102 (42%) (24). This evidence could suggest that the SSH-based techniques preferentially selected for genomic regions of *A.circinalis* with high mol% GC, although this phenomenon has not been observed in previous studies (17,25). Alternatively these methods, and in particular PPH, may have detected DNA fragments belonging to genomic regions of *A.circinalis* with an anomalous G+C content. This finding resembles the discrepancies in mol% GC found in bacterial pathogenicity islands (Pai) (26). These are mobile genomic regions, generally characterized by different mol% GC compared with the DNA of the host bacterium and by the presence of mobility genes, such as integrases and transposases (26). Here we isolated, by unsubtracted PPH, a toxic-strain-specific putative transposase (clone 179). This suggests the possibility of the presence of Pai in *A.circinalis* and supports the hypothesis that LGT may have occurred between toxic strains of this cyanobacterium. Recently, the presence of Pai-like DNA regions has been documented in the genome of the marine cyanobacterium *Synechococcus* sp.

WH8102, and associated with the lateral acquisition of specific ecological and metabolic strategies (27). Transposases have also been associated with toxin biosynthesis gene clusters in other cyanobacterial species, including the cyclic peptide toxin-producing *Microcystis aeruginosa* and *Nodularia spumigena* (28) [M.C.Moffitt and B.A.Neilan, unpublished (DDBJ/EMBL/GenBank accession number AY210783); CDS: 50150..50665, *ORF1*].

It is possible that the molecular basis for STX production is encoded in a relatively large and mobile region of the *A.circinalis* genome. The SSH procedure is designed to select highly divergent genes, while the PPH method would instead select for highly conserved genomic sequences. These loci can be characterized by either a very low rate of mutation or, possibly, by their lateral acquisition in recent evolutionary times. STX, which is unique in its distribution as a bioactive secondary metabolite of both prokaryotes and eukaryotes, may have conversely been acquired early in the evolution of these microorganisms.

Given the pathway proposed by Shimizu (29), based on feeding experiments carried out with labeled precursors in dinoflagellates and cyanobacteria, one gene recovered by unsubtracted PPH appears to be a possible candidate in STX biosynthesis. Clone 183 encodes a putative acetyltransferase, which could be implicated in the first and one of the more peculiar reactions proposed in STX synthesis: the Claisen-type condensation between the C2-carbon of arginine and C1 of acetate (30). This gene had, however, low specificity for the toxic *A.circinalis* strain 344B (Table 3). Since no biosynthetic pathway known to date is similar to the one proposed for STX, it could also be expected that some genes involved in STX biosynthesis would show very limited or no homology to other genes listed in the databases. Several hypothetical proteins and fragments with no similarity have been recovered by unsubtracted PPH, which had high levels of toxic-strain specificity (Tables 2 and 3). The hypothetical proteins investigated in this study by PCR were mainly specific for strains 131C and 344B, rather then being in common with other STX-producing *A.circinalis*. Additionally, one other gene identified by the unsubtracted PPH has the potential of being involved in STX regulation: a 60 kDa chaperonin GroEL (clone 118). Recently, a 70 kDa chaperonin has been correlated with the regulation of STX biosynthesis in the heterotrophic bacterium *Pseudomonas diminuta*, a symbiont of the PSP-toxin producing dinoflagellate *Alexandrium catenella* (31). Although toxic-strain specificity for strain 131C was observed (Table 2), this gene could not be detected by microarray hybridization using digested strain 344B genomic DNA (Table 3).

In conclusion, the PPH technique was employed here for the first time, with the aim of highlighting genomic DNA sequences associated with STX production in *A.circinalis.* The approach presented in this study allowed the identification of genes potentially involved in STX biosynthesis and regulation. This method could also be applied to phylogenetically distant taxa by ligating one terminal adaptor directly to each digested genomic DNA (unpublished results), in order to recover genes with similar primary structure. PPH of two genomes could therefore be utilized as a new tool for the identification of genes that have been laterally transferred. The application of this technique suggested the presence of Pai-like regions in the genome of *A.circinalis*. The combined SSH-PPH/DNA-microarray approach revealed a high degree of genomic diversity among phylogenetically closely related strains of *A.circinalis* and could be used for identifying molecular probes to detect STX-producing genotypes in the environment. PPH may also be useful for acquiring information on secondary metabolism and its regulation in other toxic microalgae.

ACKNOWLEDGEMENTS

The authors are grateful to B. Robertson and M. Moffitt for their assistance in developing the BGGM[1] DNA-microarray, and to D. Vigetti and B. P. Burns for support and for reviewing the manuscript. F.P. was supported by research scholarships from the University of New South Wales and the School of Biotechnology and Biomolecular Sciences.

REFERENCES

1. Humpage,A.R., Rositano,J., Bretag,A., Brown,R., Baker,P., Nicholson,B.C. and Steffensen,D.A. (1994) Paralytic shellfish poisons from Australian cyanobacterial blooms. *Aust. J. Mar. Freshwater Res.*, **45**, 761–771.
2. Llewellyn,L.E., Negri,A.P., Doyle,J., Baker,P.D., Beltran,E.C. and Neilan,B.A. (2001) Radioreceptor assays for sensitive detection and quantitation of saxitoxin and its analogues from strains of the freshwater cyanobacterium, *Anabaena circinalis*. *Environ. Sci. Technol.*, **35**, 1445–1451.
3. Catterall,W.A. (1980) Neurotoxins that act on voltage-sensitive sodium channels in excitable membranes. *Annu. Rev. Pharmacol. Toxicol.*, **20**, 15–43.
4. Shimizu,Y. (1977) Chemistry and distribution of deleterious dinoflagellate toxins. In Faulkner,D.J. and Fenical,W.H. (eds), *Marine Natural Products Chemistry*. Plenum, New York, NY, pp. 261–269.
5. Harada,T., Oshima,Y. and Yasumoto,T. (1982) Structure of two paralytic shellfish toxins, Gonyautoxins V and VI, isolated from a tropical dinoflagellate, *Pyrodinium bahamense* var. *compressa*. *Agric. Biol. Chem.*, **46**, 1861–1864.
6. Oshima,Y., Hasegawa,M., Yasumoto,T., Hallegaeff,G. and Blackburn,S. (1987) Dinoflagellate *Gimnodium catenatum* as the source of paralytic shellfish toxins in Tasmanian shellfish. *Toxicon*, **25**, 1105–1111.
7. Alam,M., Ikawa,M., Sasner,J.J.,Jr and Sawyer,P.J. (1973) Purification of *Aphanizomenon flos-aquae* toxin and its chemical and physiological properties. *Toxicon*, **11**, 65–72.
8. Carmichael,W.W., Evans,W.R., Yin,Q.Q., Bell,P. and Moczydlowsky,E. (1997) Evidence of paralytic shellfish poisons in the freshwater cyanobacterium *Lyngbya wollei* (Farlow ex Gomont) comb. nov. *Appl. Env. Microbiol.*, **63**, 3104–3110.
9. Lagos,N., Onodera,H., Zagatto,P.A., Andrinolo,D., Azevedo,S.M.F.Q. and Oshima,Y. (1999) The first evidence of paralytic shellfish toxins in the freshwater cyanobacterium *Cylindrospermopsis raciborskii*, isolated from Brazil. *Toxicon*, **37**, 1359–1373.
10. Pomati,F., Sacchi,S., Rossetti,C., Giovannardi,S., Onodera,H., Oshima,Y. and Neilan,B.A. (2000) The freshwater cyanobacterium *Planktothrix* sp. FP1: molecular identification and detection of paralytic shellfish poisoning toxins. *J. Phycol.*, **36**, 553–562.
11. Gallacher,S. and Smith,E.A. (1999) Bacteria and paralytic shellfish toxins. *Protist*, **150**, 245–255.
12. Akopyants,N.S., Fradkov,A., Diatchenko,L., Hill,J.E., Siebert,P.D., Lukyanov,S.A., Sverdlov,E.D. and Berg,D.E. (1998) PCR-based subtractive hybridization and differences in gene content among strains of *Helicobacter pylori*. *Proc. Natl Acad. Sci. USA*, **95**, 13108–13113.
13. Zhang,Y.L., Ong,C.T. and Leung,K.Y. (2000) Molecular analysis of genetic differences between virulent and avirulent strains of *Aeromonas hydrophila* isolated from deseased fish. *Microbiology*, **146**, 999–1009.
14. Janke,B., Dobrindt,U., Hacker,J. and Blum-Oehler,G. (2001) A 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.

15. Agron,P.G., Macht,M., Radnedge,L., Skowronski,E.W., Miller,W. and Andersen,G.L. (2002) Use of subtractive hybridization for comprehensive surveys of prokaryotic genome differences. *FEMS Microbiol. Lett.*, **211**, 175–182.
16. Mavrodi,D.V., Mavrodi,O.V., McSpadden-Gardener,B.B., Landa,B.B., Weller,D.M. and Thomashow,L.S. (2002) Identification of differences in genome content among phlD-positive *Pseudomonas fluorescens* strains by using PCR-based subtractive hybridization. *Appl. Environ. Microbiol.*, **68**, 5170–5176.
17. Nesbø,C.L., Nelson,K.E. and Doolittle,W.F. (2002) Suppressive subtractive hybridization detects extensive genomic diversity in *Thermotoga maritima*. *J. Bacteriol.*, **184**, 4475–4488.
18. Beltran,E.C. and Neilan,B.A. (2000) Geographical segregation of the neurotoxin-producing cyanobacterium *Anabaena circinalis*. *Appl. Environ. Microbiol.*, **66**, 4468–4474.
19. Beltran,E.C. (2001) *Taxonomy and Molecular Genetics of the Saxitoxin Producing Cyanobacterium* Anabaena circinalis [PhD thesis]. School of Microbiology and Immunology, The University of New South Wales, Sydney, Australia.
20. Kaebernick,M. (2001) *Regulation and Function of Microcystin Production in* Microcystis aeruginosa [PhD thesis]. School of Microbiology and Immunology, The University of New South Wales, Sydney, Australia.
21. Tillett,D. and Neilan,B.A. (2000) Xanthogenate nucleic acid isolation from cultured and environmental cyanobacteria. *J. Phycol.*, **35**, 1–8.
22. Sambrook,J., Fritsch,E.F. and Maniatis,T. (eds) (1989) *Molecular Cloning: a Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
23. Eisen,J.A. and Fraser,C.M. (2003) Phylogenomics: intersection of evolution and genomics. *Science*, **300**, 1706–1707.
24. Meeks,J.C., Elhai,J., Thiel,T., Potts,M., Larimer,F., Lamerdin,J., Predki,P. and Atlas,R. (2001) An overview of the genome of *Nostoc punctiforme*, a multicellular, symbiotic cyanobacterium. *Photosynth. Res.*, **70**, 85–106.
25. Perrin,A., Nassif,X. and Tinsley,C. (1999) Identification of regions of the chromosome of *Neisseria meningitidis* and *Neisseria gonorrhoeae* which are specific to the pathogenic *Neisseria* species. *Infect. Immun.*, **67**, 6119–6129.
26. 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.
27. Palenik,B., Brahamsha,B., Larimer,F.W., Land,M., Hauser,L., Chain,P., Lamerdin,J., Regala,W., Allen,E.E., McCarren,J. *et al.* (2003) The genome of a motile marine Synechococcus. *Nature*, **424**, 1037–1042.
28. Tillett,D., Dittmann,E., Erhard,M., von Döhren,H., Börner,T. and Neilan,B.A. (2000) Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide–polyketide synthetase system. *Chem. Biol.*, **7**, 753–764.
29. Shimizu,Y. (1993) Microalgal metabolites. *Chem. Rev.*, **93**, 1685–1698.
30. Shimizu,Y., Norte,M., Hori,A., Genenah,A. and Kobayashi,M. (1984) Biosynthesis of saxitoxin analogues: the unexpected pathway. *J. Am. Chem. Soc.*, **106**, 6433–6434.
31. Còrdova,J., Bustamante,J. and Bates,S.S. (2002) *Immunological Detection of PSP Toxin Activity from Bacteria Isolated from* Alexandrium catenella, *with Additional Investigations using Bacteria from* Prorocentrum lima *and* Pseudonitzscgia multiseries. Poster presentation, 10th International Conference on Harmful Algae, October 21–25, St Pete Beach, FL.