(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 October 2009 (29.10.2009)

PCT

(10) International Publication Number
WO 2009/129558 A1

(51) International Patent Classification:
C12Q 1/68 (2006.01) *C12N 9/02* (2006.01)
C12N 9/04 (2006.01) *C12N 9/10* (2006.01)
C12N 9/20 (2006.01) *C12N 9/78* (2006.01)

(21) International Application Number:
PCT/AU2008/001805

(22) International Filing Date:
5 December 2008 (05.12.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2008902056 24 April 2008 (24.04.2008) AU

(71) Applicant *(for all designated States except US)*: **NEWSOUTH INNOVATIONS PTY LIMITED** [AU/AU]; Rupert Myers Building, Barker Street, University of New South Wales, NSW 2052 (AU).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **NEILAN, Brett, A.** [AU/AU]; 2/15 McKeon Street, Maroubra, NSW 2035 (AU). **MIHALI, Troco, Kaan** [DE/AU]; 2/18 Carlisle Street, Tamarama, NSW 2026 (AU). **KELLMANN, Ralf** [DE/NO]; Gamle Kirkeveien 5, N-5222 Nesttun (NO). **JEON, Young, Jae** [KR/AU]; 2/11 Anzac Parade, Kensington, NSW 2033 (AU).

(74) Agent: **SPRUSON & FERGUSON**; GPO Box 3898, Syndey, NSW 2001 (AU).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2009/129558 A1

(54) **Title**: CYANOBACTERIA SAXITOXIN GENE CLUSTER AND DETECTION OF CYANOTOXIC ORGANISMS

(57) **Abstract**: The present invention relates to methods for the detection of cyanobacteria, dinoflagellates, and in particular, methods for the detection of cyanotoxic organisms. Kits for the detection of cyanobacteria, dinoflagellates, and cyanotoxic organisms are provided. The invention further relates to methods of screening for compounds that modulate the activity of polynucleotides and/or polypeptides of the saxitoxin and cylindrospermopsin biosynthetic pathways.

Cyanobacteria saxitoxin gene cluster and detection of cyanotoxic organisms

Technical Field

The present invention relates to methods for the detection of cyanobacteria, dinoflagellates, and in particular, methods for the detection of cyanotoxic organisms. Kits for the detection of cyanobacteria, dinoflagellates, and cyanotoxic organisms are provided. The invention further relates to methods of screening for compounds that modulate the activity of polynucleotides and/or polypeptides of the saxitoxin and cylindrospermopsin biosynthetic pathways.

Background

Cyanobacteria, also known as blue-green algae, are photosynthetic bacteria widespread in marine and freshwater environments. Of particular significance for water quality and human and animal health are those cyanobacteria which produce toxic compounds. Under eutrophic conditions cyanobacteria tend to form large blooms which drastically promote elevated toxin concentrations. Cyanobacterial blooms may flourish and expand in coastal waters, streams, lakes, and in drinking water and recreational reservoirs. The toxins they produce can pose a serious health risk for humans and animalsand this problem is internationally relevant since most toxic cyanobacteria have a global distribution.

A diverse range of cyanobacterial genera are well known for the formation of toxic blue-green algal blooms on water surfaces. Saxitoxin (*SXT*) and its analogues cause the paralytic shellfish poisoning (PSP) syndrome, which afflicts human health and impacts on coastal shellfish economies worldwide. PSP toxins are unique alkaloids, being produced by both prokaryotes and eukaryotes. PSP toxins are among the most potent and pervasive algal toxins and are considered a serious toxicological health-risk that may affect humans, animals and ecosystems worldwide. These toxins block voltage-gated sodium and calcium channels, and prolong the gating of potassium channels preventing the transduction of neuronal signals. It has been estimated that more than 2000 human cases of PSP occur globally every year. Moreover, coastal blooms of producing microorganisms result in millions of dollars of economic damage due to PSP toxin contamination of seafood and the continuous requirement for costly biotoxin monitoring programs. Early warning systems to anticipate paralytic shellfish toxin (PST)-producing

algal blooms, such as PCR and ELISA-based screening, are as yet unavailable due to the lack of data on the genetic basis of PST production.

SXT is a tricyclic perhydropurine alkaloid which can be substituted at various positions leading to more than 30 naturally occurring *SXT* analogues. Although *SXT* biosynthesis seems complex and unique, organisms from two kingdoms, including certain species of marine dinoflagellates and freshwater cyanobacteria, are capable of producing these toxins, apparently by the same biosynthetic route. In spite of considerable efforts none of the enzymes or genes involved in the biosynthesis and modification of *SXT* have been previously identified.

The occurrence of the cyanobacterial genus *Cylindrospermopsis* has been documented on all continents and therefore poses a significant public health threat on a global scale. The major toxin produced by *Cylindrospermopsis* is cylindrospermopsin (*CYR*). Besides posing a threat to human health, cylindrospermopsin also causes significant economic losses for farmers due to the poisoning of livestock with cylindrospermopsin-contaminated drinking water. Cylindrospermopsin has hepatotoxic, general cytotoxic and neurotoxic effects and is a potential carcinogen. Its toxicity is due to the inhibition of glutathione and protein synthesis as well as inhibiting cytochrome P450. Six cyanobacterial species have so far been identified to produce cylindrospermopsin; *Cylindrospermopsis raciborskii, Aphanizomenon ovalisporum, Aphanizomenon flos-aquae, Umezakia natans, Rhaphdiopsis curvata* and *Anabaena bergii*. Incidents of human poisoning with cylindrospermopsin have only been reported in sub-tropical Australia to date, however *C. raciborskii* and *A. flos-aquae* have recently been detected in areas with more temperate climates. The tendency of *C. raciborskii* to form dense blooms and the invasiveness of the producer organisms gives rise to global concerns for drinking water quality and necessitates the monitoring of drinking water reserves for the presence of cylindrospermopsin producers.

There is a need for rapid and accurate methods detecting cyanobacteria, and in particular those strains which are capable of producing cyanotoxins such as saxitoxin and cylindrospermopsin. Rapid and accurate methods for detecting cyanotoxic organisms are needed for assessing the potential health hazard of cyanobacterial blooms and for the implementation of effective water management strategies to minimize the effects of toxic bloom outbreaks.

Summary

In a first aspect, there is provided an isolated polynucleotide comprising a sequence according to SEQ ID NO: 1 or a variant or fragment thereof.

In one embodiment of the first aspect, the fragment comprises a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants and fragments thereof.

In a second aspect, there is provided an isolated ribonucleic acid or an isolated complementary DNA encoded by a sequence according to the first aspect.

In a third aspect, there is provided an isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 67, SEQ ID NO: 69, and variants and fragments thereof.

In one embodiment, there is provided a probe or primer that hybridises specifically with one or more of: a polynucleotide according to the first aspect, a ribonucleic acid or complementary DNA according to the second aspect, or a polypeptide according the third aspect.

In another embodiment, there is provided a vector comprising a polynucleotide according to the first aspect, or a ribonucleic acid or complementary DNA according the second aspect. The vector may be an expression vector.

In another embodiment, a host cell is provided comprising the vector.

In another embodiment, there is provided an isolated antibody capable of binding specifically to a polypeptide according to the third aspect.

In a fourth aspect, there is provided a method for the detection of cyanobacteria, the method comprising the steps of obtaining a sample for use in the method and analyzing the sample for the presence of one or more of:

(i) a polynucleotide comprising a sequence according to the first aspect

(ii) a ribonucleic acid or complementary DNA according to the second aspect

(iii) a polypeptide comprising a sequence according to third aspect

wherein said presence is indicative of cyanobacteria in the sample.

In a fifth aspect, there is provided a method for detecting a cyanotoxic organism, the method comprising the steps of obtaining a sample for use in the method and analyzing the sample for the presence of one or more of:

(i) a polynucleotide comprising a sequence selected from the group consisting of: SEQ ID NO: 14, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 36, and variants and fragments thereof

(ii) a ribonucleic acid or complementary DNA encoded by a sequence according to (i)

(iii) a polypeptide comprising a sequence selected from the group consisting of: SEQ ID NO: 15, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 37, and variants and fragments thereof,

wherein said presence is indicative of cyanotoxic organisms in the sample.

In one embodiment of the fifth aspect, the cyanotoxic organism is a cyanobacteria or a dinoflagellate.

In one embodiment of the fourth and fifth aspects, analyzing the sample comprises amplification of DNA from the sample by polymerase chain reaction and detecting the amplified sequences. The polymerase chain reaction may utilise one or more primers comprising a sequence selected from the group consisting of SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 117, SEQ ID NO: 118, SEQ ID NO: 119, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 122, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 130, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 133, SEQ ID NO: 134, and variants and fragments thereof.

In another embodiment of the fourth and fifth aspects, the method comprises further analyzing the sample for the presence of one or more of:

(i) a polynucleotide comprising a sequence selected from the group consisting of: SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, and variants and fragments thereof,

(ii) a ribonucleic acid or complementary DNA encoded by a sequence according to (i),

(iii) a polypeptide comprising a sequence selected from the group consisting of: SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92 , SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, and SEQ ID NO: 110, and variants and fragments thereof.

The further analyis of the sample may comprise amplification of DNA from the sample by polymerase chain reaction. The polymerase chain reaction may utilise one or more primers comprising a sequence selected from the group consisting of SEQ ID NO: 111, SEQ ID NO: 112, or variants or fragments thereof.

In a sixth aspect, there is provided a method for the detection of dinoflagellates, the method comprising the steps of obtaining a sample for use in the method and analyzing the sample for the presence of one or more of:

(i) a polynucleotide comprising a sequence according to the first aspect,

(ii) a ribonucleic acid or complementary DNA according to the second aspect,

(iii) a polypeptide comprising a sequence according to the third aspect,

wherein said presence is indicative of dinoflagellates in the sample.

In one embodiment of the sixth aspect, analysing the sample comprises amplification of DNA from the sample by polymerase chain reaction and detecting the amplified sequences.

In one embodiment of the fourth, fifth, and sixth aspects, the detection comprises one or both of gel electrophoresis and nucleic acid sequencing. The sample may comprise one or more isolated or cultured organisms. The sample may be an environmental sample. The environmental sample may be derived from salt water, fresh water or a blue-green algal bloom.

In a seventh aspect, there is provided a kit for the detection of cyanobacteria, the kit comprising at least one agent for detecting the presence of one or more of:

(i) a polynucleotide comprising a sequence according to the first aspect,

(ii) a ribonucleic acid or complementary DNA according to the second aspect,

(iii) a polypeptide comprising a sequence according to the third aspect,

wherein said presence is indicative of cyanobacteria in the sample.

In an eighth aspect, there is provided a kit for the detection of cyanotoxic organisms, the kit comprising at least one agent for detecting the presence of one or more of:

(i) a polynucleotide comprising a sequence selected from the group consisting of: SEQ ID NO: 14, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 36, and variants and fragments thereof,

(ii) a ribonucleic acid or complementary DNA encoded by a sequence according to (i),

(iii) a polypeptide comprising a sequence selected from the group consisting of: SEQ ID NO: 15, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 37, and variants and fragments thereof,

wherein said presence is indicative of cyanotoxic organisms in the sample.

In one embodiment of the seventh and eighth aspects, the at least one agent is a primer, antibody or probe. The primer or probe may comprise a sequence selected from the group consisting of SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 117, SEQ ID NO: 118, SEQ ID NO: 119, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 122, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 130, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 133, SEQ ID NO: 134, and variants and fragments thereof.

In another embodiment of the seventh and eighth aspects, the kit further comprises at least one additional agent for detecting the presence of one or more of:

(i) a polynucleotide comprising a sequence selected from the group consisting of: SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, and variants and fragments thereof,

(ii) a ribonucleic acid or complementary DNA encoded by a sequence according to (i),

(iii) a polypeptide comprising a sequence selected from the group consisting of: SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92 , SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, and SEQ ID NO: 110, and variants and fragments thereof.

The at least one additional agent may be a primer, antibody or probe. The primer or probe may comprise a sequence selected from the group consisting of SEQ ID NO: 109, SEQ ID NO: 110, and variants and fragments thereof.

In a ninth aspect, there is provided a kit for the detection of dinoflagellates, the kit comprising at least one agent for detecting the presence of one or more of:

(i) a polynucleotide comprising a sequence according to the first aspect,

(ii) a ribonucleic acid or complementary DNA according to the second aspect,

(iii) a polypeptide comprising a sequence according to the third aspect,

wherein said presence is indicative of dinoflagellates in the sample.

In a tenth aspect, there is provided a method of screening for a compound that modulates the expression or activity of one or more polypeptides according to the third aspect, the method comprising contacting the polypeptide with a candidate compound under conditions suitable to enable interaction of the candidate compound and the polypeptide, and assaying for activity of the polypeptide.

In one embodiment of the tenth aspect, modulating the expression or activity of one or more polypeptides comprises inhibiting the expression or activity of said polypeptide.

In another embodiment of the tenth aspect, modulating the expression or activity of one or more polypeptides comprises enhancing the expression or activity of said polypeptide.

In an eleventh aspect, there is provided an isolated polynucleotide comprising a sequence according to SEQ ID NO: 80 or a variant or fragment thereof.

In one embodiment of the eleventh aspect, the fragment comprises a sequence selected from the group consisting of SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO:107, SEQ ID NO: 109, and variants and fragments thereof.

In a twelfth aspect, there is provided a ribonucleic acid or complementary DNA encoded by a sequence according to the eleventh aspect.

In a thirteenth aspect, there is provided an isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92 , SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, SEQ ID NO: 110, and variants and fragments thereof.

In one embodiment, there is provided a probe or primer that hybridises specifically with one or more of: a polynucleotide according to the eleventh aspect, a ribonucleic acid or complementary DNA according to the twelfth aspect, or a polypeptide according to the thirteenth aspect.

In another embodiment, there is provided a vector comprising a polynucleotide according to the eleventh aspect, or a ribonucleic acid or complementary DNA according to the twelfth aspect. The vector may be an expression vector. In one embodiment, a host cell is provided comprising the vector.

In another embodiment, there is provided an isolated antibody capable of binding specifically to a polypeptide according to the thirteenth aspect.

In a fourteenth aspect, there is provided a method for the detection of cyanobacteria, the method comprising the steps of obtaining a sample for use in the method and analyzing the sample for the presence of one or more of:

(i) a polynucleotide comprising a sequence according to the eleventh aspect,

(ii) a ribonucleic acid or complementary DNA according to the twelfth aspect,

(iii) a polypeptide comprising a sequence according to thirteenth aspect,

wherein said presence is indicative of cyanobacteria in the sample.

In a fifteenth aspect, there is provided a method for detecting a cyanotoxic organism, the method comprising the steps of obtaining a sample for use in the method and analyzing the sample for the presence of one or both of:

(i) a polynucleotide comprising a sequence according to SEQ ID NO: 95 or a variant or fragment thereof,

(ii) a ribonucleic acid or complementary DNA encoded by a sequence according to (i),

(iii) a polypeptide comprising a sequence according to SEQ ID NO: 96, or a variant or fragment thereof,

wherein said presence is indicative of a cyanotoxic organism in the sample.

In one embodiment of the fifteenth aspect, the cyanotoxic organism is a cyanobacteria.

In one embodiment of the fourteenth and fifteenth aspects, analyzing the sample comprises amplification of DNA from the sample by polymerase chain reaction and detecting the amplified sequences. The polymerase chain reaction may utilise one or more primers comprising a sequence selected from the group consisting of SEQ ID NO: 111, SEQ ID NO: 112 and variants and fragments thereof.

In another embodiment of the fourteenth and fifteenth aspects, the method comprises analyzing the sample for the presence of one or more of:

(i) a polynucleotide comprising a sequence selected from the group consisting of: SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants and fragments thereof,

(ii) a ribonucleic acid or complementary DNA encoded by a sequence according to (i),

(iii) a polypeptide comprising a sequence selected from the group consisting of: SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19 SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 67, SEQ ID NO: 69, and variants and fragments thereof.

The further analysis of the sample may comprise amplification of DNA from the sample by polymerase chain reaction. The polymerase chain reaction may utilise one or more primers comprising a sequence selected from the group consisting of SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 117, SEQ ID NO:

118, SEQ ID NO: 119, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 122, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 130, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 133, SEQ ID NO: 134, and variants and fragments thereof.

In a sixteenth aspect, there is provided a method for detecting a cylindrospermopsin-producing organism, the method comprising the steps of obtaining a sample for use in the method and analyzing the sample for the presence of one or both of:

(i) a polynucleotide comprising a sequence according to SEQ ID NO: 95 or a variant or fragments thereof,

(ii) a ribonucleic acid or complementary DNA encoded by a sequence according to (i),

(iii) a polypeptide comprising a sequence according to SEQ ID NO: 96, or a variant or fragments thereof,

wherein said presence is indicative of a cylindrospermopsin-producing organism in the sample.

In one embodiment of the sixteenth aspect, the cyanotoxic organism is a cyanobacteria. In another embodiment of the sixteenth aspect, analyzing the sample comprises amplification of DNA from the sample by polymerase chain reaction and detecting the amplified sequences. The polymerase chain reaction may utilise one or more primers comprising a sequence selected from the group consisting of SEQ ID NO: 111, SEQ ID NO: 112 and variants and fragments thereof.

In one embodiment of the fourteenth, fifteenth, and sixteenth aspects, the detection comprises one or both of gel electrophoresis and nucleic acid sequencing. The sample may comprise one or more isolated or cultured organisms. The sample may be an environmental sample. The environmental sample may be derived from salt water, fresh water or a blue-green algal bloom.

In a seventeenth aspect, there is provided a kit for the detection of cyanobacteria, the kit comprising at least one agent for detecting the presence of one or more of:

(i) a polynucleotide comprising a sequence according to the eleventh aspect,

(ii) a ribonucleic acid or complementary DNA according to the twelfth aspect,

(iii) a polypeptide comprising a sequence according to the thirteenth aspect,

wherein said presence is indicative of cyanobacteria in the sample.

In an eighteenth aspect, there is provided a kit for the detection of cyanotoxic organisms, the kit comprising at least one agent for detecting the presence of one or more of:

(i) a polynucleotide comprising a sequence according to SEQ ID NO: 95 or a variant or fragment thereof,

(ii) a ribonucleic acid or complementary DNA encoded by a sequence according to (i),

(iii) a polypeptide comprising a sequence according to SEQ ID NO: 96, or a variant or fragment thereof,

wherein said presence is indicative of cyanotoxic organisms in the sample.

In one embodiment of the seventeenth and eighteenth aspects, the at least one agent is a primer, antibody or probe. The primer or probe may comprise a sequence selected from the group consisting of SEQ ID NO: 111, SEQ ID NO: 112 and variants and fragments thereof.

In another embodiment of the seventeenth and eighteenth aspects, the kit may further comprise at least one additional agent for detecting the presence of one or more nucleotide sequences selected from the group consisting of:

(i) a polynucleotide comprising a sequence selected from the group consisting of: SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants and fragments thereof,

(ii) a ribonucleic acid or complementary DNA encoded by a sequence according to (i),

(iii) a polypeptide comprising a sequence selected from the group consisting of: SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19 SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53,

SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 67, SEQ ID NO: 69, and variants and fragments thereof.

The at least one additional agent may be a primer, antibody or probe. The primer or probe may comprise a sequence selected from the group consisting of SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 117, SEQ ID NO: 118, SEQ ID NO: 119, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 122, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 130, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 133, SEQ ID NO: 134, and variants and fragments thereof.

In a nineteenth aspect, there is provided a kit for the detection of cylindrospermopsin-producing organisms, the kit comprising at least one agent for detecting the presence of one or more of:

(i) a polynucleotide comprising a sequence according to SEQ ID NO: 95 or a variant or fragment thereof,

(ii) a ribonucleic acid or complementary DNA encoded by a sequence according to (i),

(iii) a polypeptide comprising a sequence according to SEQ ID NO: 96, or a variant or fragment thereof,

wherein said presence is indicative of a cylindrospermopsin-producing organism in the sample.

In a twentieth aspect, there is provided a method of screening for a compound that modulates the expression or activity of one or more polypeptides according to the thirteenth aspect, the method comprising contacting the polypeptide with a candidate compound under conditions suitable to enable interaction of the candidate compound and the polypeptide, and assaying for activity of the polypeptide.

In one embodiment of the twentieth aspect, modulating the expression or activity of one or more polypeptides comprises inhibiting the expression or activity of said polypeptide.

In another embodiment of the twentieth aspect, modulating the expression or activity of one or more polypeptides comprises enhancing the expression or activity of said polypeptide.

Definitions

As used in this application, the singular form "a", "an" and "the" include plural references unless the context clearly dictates otherwise. For example, the term "a stem cell" also includes a plurality of stem cells.

As used herein, the term "comprising" means "including." Variations of the word "comprising", such as "comprise" and "comprises," have correspondingly varied meanings. Thus, for example, a polynucleotide "comprising" a sequence encoding a protein may consist exclusively of that sequence or may include one or more additional sequences.

As used herein, the terms "antibody" and "antibodies" include IgG (including IgG1, IgG2, IgG3, and IgG4), IgA (including IgA1 and IgA2), IgD, IgE, or IgM, and IgY, whole antibodies, including single-chain whole antibodies, and antigen-binding fragments thereof. Antigen-binding antibody fragments include, but are not limited to, Fab, Fab' and F(ab')2, Fd, single-chain Fvs (scFv), single-chain antibodies, disulfide-linked Fvs (sdFv) and fragments comprising either a VL or VH domain. The antibodies may be from any animal origin. Antigen-binding antibody fragments, including single-chain antibodies, may comprise the variable region(s) alone or in combination with the entire or partial of the following: hinge region, CH1, CH2, and CH3 domains. Also included are any combinations of variable region(s) and hinge region, CH1, CH2, and CH3 domains. Antibodies may be monoclonal, polyclonal, chimeric, multispecific, humanized, and human monoclonal and polyclonal antibodies which specifically bind the biological molecule.

As used herein, the terms "polypeptide" and "protein" are used interchangeably and are taken to have the same meaning.

As used herein, the terms "nucleotide sequence" and "polynucleotide sequence" are used interchangeably and are taken to have the same meaning.

As used herein, the term "kit" refers to any delivery system for delivering materials. In the context of the detection assays described herein, such delivery systems include systems that allow for the storage, transport, or delivery of reaction reagents (for example labels, reference samples, supporting material, etc. in the appropriate containers) and/or supporting materials (for example, buffers, written instructions for performing the assay etc.) from one location to another. For example, kits include one or more enclosures, such as boxes, containing the relevant reaction reagents and/or supporting materials.

Any discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is solely for the purpose of providing a context for the present invention. It is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the field relevant to the present invention before the priority date of this application.

For the purposes of description all documents referred to herein are incorporated by reference unless otherwise stated.

Brief Description of the Drawings

A preferred embodiment of the present invention will now be described, by way of an example only, with reference to the accompanying drawings wherein:

Figure 1A is a table showing the distribution of the *sxt* genes in toxic and non-toxic cyanobacteria. PSP, saxitoxin; CYLN, cylindrospermopsin; +, gene fragment amplified; - no gene detected.

Figure 1B is a table showing primer sequences used to amplify various *SXT* genes.

Figure 2 is a table showing *sxt* genes from the saxitoxin gene cluster of *C. raciborskii* T3, their putative length, their BLAST similarity match with similar protein sequences from other organisms, and their predicted function.

Figure 3 is a diagram showing the structural organisation of the *sxt* gene cluster from *C. raciborskii* T3. Abbreviations used are: IS4, insertion sequence 4; at, aminotransferase; dmt, drug metabolite transporter; ompR, transcriptional regulator of *ompR* family; penP, penicillin binding; smf, gene predicted to be involved in DNA uptake. The scale indicates the gene cluster length in base pairs.

Figure 4 is a flow diagram showing the pathway for *SXT* biosynthesis and the putative functions of *sxt* genes.

Figure 5 shows MS/MS spectra of selected ions from cellular extracts of *Cylindrospermopsis raciborskii* T3. The predicted fragmentation of ions and the corresponding *m/z* values are indicated. Figure 5A, arginine (*m/z* 175); Figure 5B, saxitoxin (*m/z* 300); Figure 5C, intermediate A' (*m/z* 187); Figure 5D, intermediate C' (*m/z* 211); Figure 5E, intermediate E' (*m/z* 225).

Figure 6 is a table showing the *cyr* genes from the cylindrospermopsin gene cluster of *C. raciborskii* AWT205, their putative length, their BLAST similarity match with similar protein sequences from other organisms, and their predicted function.

Figure 7 is a table showing the distribution of the sulfotransferase gene (*cyrJ*) in toxic and non-toxic cyanobacteria. 16S rRNA gene amplification is shown as a positive control. CYLN, cylindrospermopsin; SXT, saxitoxin; N.D., not detected; +, gene fragment amplified; -, no gene detected; NA, not available; AWQC, Australian Water Quality Center.

Figure 8 is a flow diagram showing the biosynthetic pathway of cylindrospermopsin biosynthesis.

Figure 9 is a diagram showing the structural organization of the cylindrospermopsin gene cluster from *C. raciborskii* AWT205. Scale indicates gene cluster length in base pairs.

Description

The inventors have identified a gene cluster responsible for saxitoxin biosynthesis (the *SXT* gene cluster) and a gene cluster responsible for cylindrospermopsin biosynthesis (the *CYR* gene cluster). The full sequence of each gene cluster has been determined and functional activities assigned to each of the genes identified therein. Based on this information, the inventors have elucidated the full saxitoxin and cylindrospermopsin biosynthetic pathways.

Accordingly, the invention provides polynucleotide and polypeptide sequences derived from each of the *SXT* and *CYR* gene clusters and in particular, sequences relating to the specific genes within each pathway. Methods and kits for the detection of cyanobacterial strains in a sample are provided based on the presence (or absence) in the sample of one or more of the sequences of the invention. The inventors have determined that certain open-reading frames present in the *SXT* gene cluster of saxitoxin-producing microorganisms are absent in the *SXT* gene cluster of microorganisms that do not produce saxitoxin. Similarly, it has been discovered that one open-reading frame present in the *CYR* gene cluster of cylindrospermopsin-producing microorganisms is absent in non-cylindrospermopsin-producing microorganisms. Accordingly, the invention provides methods and kits for the detection of toxin-producing microorganisms.

Also provided by the invention are screening methods for the identification of compounds capable of modulating the expression or activity of proteins in the saxitoxin and/or cylindrospermopsin biosynthetic pathways.

Polynucleotides and polypeptides

The inventors have determined the full polynucleotide sequence of the saxitoxin (*SXT*) gene cluster and the cylindrospermopsin (*CYR*) gene cluster.

In accordance with aspects and embodiments of the invention, the *SXT* gene cluster may have, but is not limited to, the polynucleotide sequence as set forth SEQ ID NO: 1 (GenBank accession number DQ787200), or display sufficient sequence identity thereto to hybridise to the sequence of SEQ ID NO: 1.

The *SXT* gene cluster comprises 31 genes and 30 intergenic regions.

Gene 1 of the *SXT* gene cluster is a 759 base pair (bp) nucleotide sequence set forth in SEQ ID NO: 4. The nucleotide sequence of *SXT* Gene 1 ranges from the nucleotide in position 1625 up to the nucleotide in position 2383 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 1 (*SXTD*) is set forth in SEQ ID NO: 5.

Gene 2 of the *SXT* gene cluster is a 396 bp nucleotide sequence set forth in SEQ ID NO: 6. The nucleotide sequence of *SXT* Gene 2 ranges from the nucleotide in position 2621 up to the nucleotide in position 3016 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 2 (*ORF3*) is set forth in SEQ ID NO: 7.

Gene 3 of the *SXT* gene cluster is a 360 bp nucleotide sequence set forth in SEQ ID NO: 8. The nucleotide sequence of *SXT* Gene 3 ranges from the nucleotide in position 2955 up to the nucleotide in position 3314 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 3 (*ORF4*) is set forth in SEQ ID NO: 9.

Gene 4 of the *SXT* gene cluster is a 354 bp nucleotide sequence set forth in SEQ ID NO: 10. The nucleotide sequence of *SXT* Gene 4 ranges from the nucleotide in position 3647 up to the nucleotide in position 4000 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 4 (*SXTC*) is set forth in SEQ ID NO: 11.

Gene 5 of the *SXT* gene cluster is a 957 bp nucleotide sequence set forth in SEQ ID NO: 12. The nucleotide sequence of *SXT* Gene 5 ranges from the nucleotide in position 4030 up to the nucleotide in position 4986 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 5 (*SXTB*) is set forth in SEQ ID NO: 13.

Gene 6 of the *SXT* gene cluster is a 3738 bp nucleotide sequence set forth in SEQ ID NO: 14. The nucleotide sequence of *SXT* Gene 6 ranges from the nucleotide in position 5047 up to the nucleotide in position 8784 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 6 (*SXTA*) is set forth in SEQ ID NO: 15.

Gene 7 of the *SXT* gene cluster is a 387 bp nucleotide sequence set forth in SEQ ID NO: 16. The nucleotide sequence of *SXT* Gene 7 ranges from the nucleotide in position

9140 up to the nucleotide in position 9526 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 7 (*SXTE*) is set forth in SEQ ID NO: 17.

Gene 8 of the *SXT* gene cluster is a 1416 bp nucleotide sequence set forth in SEQ ID NO: 18. The nucleotide sequence of *SXT* Gene 8 ranges from the nucleotide in position 9686 up to the nucleotide in position 11101 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 8 (*SXTF*) is set forth in SEQ ID NO: 19.

Gene 9 of the *SXT* gene cluster is an 1134 bp nucleotide sequence set forth in SEQ ID NO: 20. The nucleotide sequence of *SXT* Gene 9 ranges from the nucleotide in position 11112 up to the nucleotide in position 12245 of SEQ ID NO: 1. The polypeptide sequence encoded by *SXT* Gene 9 (*SXTG*) is set forth in SEQ ID NO: 21.

Gene 10 of the *SXT* gene cluster is a 1005 bp nucleotide sequence set forth in SEQ ID NO: 22. The nucleotide sequence of *SXT* Gene 10 ranges from the nucleotide in position 12314 up to the nucleotide in position 13318 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 10 (*SXTH*) is set forth in SEQ ID NO: 23.

Gene 11 of the *SXT* gene cluster is an 1839 bp nucleotide sequence set forth in SEQ ID NO: 24. The nucleotide sequence of *SXT* Gene 11 ranges from the nucleotide in position 13476 up to the nucleotide in position 15314 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 11 (*SXTI*) is set forth in SEQ ID NO: 25.

Gene 12 of the *SXT* gene cluster is a 444 bp nucleotide sequence set forth in SEQ ID NO: 26. The nucleotide sequence of *SXT* Gene 12 ranges from the nucleotide in position 15318 up to the nucleotide in position 15761 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 12 (*SXTJ*) is set forth in SEQ ID NO: 27.

Gene 13 of the *SXT* gene cluster is a 165 bp nucleotide sequence set forth in SEQ ID NO: 28. The nucleotide sequence of *SXT* Gene 13 ranges from the nucleotide in position 15761 up to the nucleotide in position 15925 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 13 (*SXTK*) is set forth in SEQ ID NO: 29.

Gene 14 of the *SXT* gene cluster is a 1299 bp nucleotide sequence set forth in SEQ ID NO: 30. The nucleotide sequence of *SXT* Gene 14 ranges from the nucleotide in position 15937 up to the nucleotide in position 17235 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 14 (*SXTL*) is set forth in SEQ ID NO: 31.

Gene 15 of the *SXT* gene cluster is a 1449 bp nucleotide sequence set forth in SEQ ID NO: 32. The nucleotide sequence of *SXT* Gene 15 ranges from the nucleotide in position 17323 up to the nucleotide in position 18771 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 16 (*SXTM*) is set forth in SEQ ID NO: 33.

Gene 16 of the *SXT* gene cluster is an 831 bp nucleotide sequence set forth in SEQ ID NO: 34. The nucleotide sequence of *SXT* Gene 16 ranges from the nucleotide in position 19119 up to the nucleotide in position 19949 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 16 (*SXTN*) is set forth in SEQ ID NO: 35.

Gene 17 of the *SXT* gene cluster is a 774 bp nucleotide sequence set forth in SEQ ID NO: 36. The nucleotide sequence of *SXT* Gene 17 ranges from the nucleotide in position 20238 up to the nucleotide in position 21011 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 17 (*SXTX*) is set forth in SEQ ID NO: 37.

Gene 18 of the *SXT* gene cluster is a 327 bp nucleotide sequence set forth in SEQ ID NO: 38. The nucleotide sequence of *SXT* Gene 18 ranges from the nucleotide in position 21175 up to the nucleotide in position 21501 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 18 (*SXTW*) is set forth in SEQ ID NO: 39.

Gene 19 of the *SXT* gene cluster is a 1653 bp nucleotide sequence set forth in SEQ ID NO: 40. The nucleotide sequence of *SXT* Gene 219 ranges from the nucleotide in position 21542 up to the nucleotide in position 23194 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 19 (*SXTV*) is set forth in SEQ ID NO: 41.

Gene 20 of the *SXT* gene cluster is a 750 bp nucleotide sequence set forth in SEQ ID NO: 42. The nucleotide sequence of *SXT* Gene 20 ranges from the nucleotide in position 23199 up to the nucleotide in position 23948 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 20 (*SXTU*) is set forth in SEQ ID NO: 43.

Gene 21 of the *SXT* gene cluster is a 1005 bp nucleotide sequence set forth in SEQ ID NO: 44. The nucleotide sequence of *SXT* Gene 21 ranges from the nucleotide in position 24091 up to the nucleotide in position 25095 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 21 (*SXTT*) is set forth in SEQ ID NO: 45.

Gene 22 of the *SXT* gene cluster is a 726 bp nucleotide sequence set forth in SEQ ID NO: 46. The nucleotide sequence of *SXT* Gene 22 ranges from the nucleotide in position 25173 up to the nucleotide in position 25898 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 22 (*SXTS*) is set forth in SEQ ID NO: 47.

Gene 23 of the *SXT* gene cluster is a 576 bp nucleotide sequence set forth in SEQ ID NO: 48. The nucleotide sequence of *SXT* Gene 23 ranges from the nucleotide in position 25974 up to the nucleotide in position 26549 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 23 (*ORF24*) is set forth in SEQ ID NO: 49.

Gene 24 of the *SXT* gene cluster is a 777 bp nucleotide sequence set forth in SEQ ID NO: 50. The nucleotide sequence of *SXT* Gene 24 ranges from the nucleotide in

position 26605 up to the nucleotide in position 27381 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 24 (*SXTR*) is set forth in SEQ ID NO: 51.

Gene 25 of the *SXT* gene cluster is a 777 bp nucleotide sequence set forth in SEQ ID NO: 52. The nucleotide sequence of *SXT* Gene 25 ranges from the nucleotide in position 27392 up to the nucleotide in position 28168 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 25 (*SXTQ*) is set forth in SEQ ID NO: 53.

Gene 26 of the *SXT* gene cluster is a 1227 bp nucleotide sequence set forth in SEQ ID NO: 54. The nucleotide sequence of *SXT* Gene 26 ranges from the nucleotide in position 28281 up to the nucleotide in position 29507 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 26 (*SXTP*) is set forth in SEQ ID NO: 55.

Gene 27 of the *SXT* gene cluster is a 603 bp nucleotide sequence set forth in SEQ ID NO: 56. The nucleotide sequence of *SXT* Gene 27 ranges from the nucleotide in position 29667 up to the nucleotide in position 30269 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 27 (*SXTO*) is set forth in SEQ ID NO: 57.

Gene 28 of the *SXT* gene cluster is a 1350 bp nucleotide sequence set forth in SEQ ID NO: 58. The nucleotide sequence of *SXT* Gene 28 ranges from the nucleotide in position 30612 up to the nucleotide in position 31961 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 28 (*ORF29*) is set forth in SEQ ID NO: 59.

Gene 29 of the *SXT* gene cluster is a 666 bp nucleotide sequence set forth in SEQ ID NO: 60. The nucleotide sequence of *SXT* Gene 29 ranges from the nucleotide in position 32612 up to the nucleotide in position 33277 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 29 (*SXTY*) is set forth in SEQ ID NO: 61.

Gene 30 of the *SXT* gene cluster is a 1353 bp nucleotide sequence set forth in SEQ ID NO: 62. The nucleotide sequence of *SXT* Gene 30 ranges from the nucleotide in position 33325 up to the nucleotide in position 34677 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 30 (*SXTZ*) is set forth in SEQ ID NO: 63.

Gene 31 of the *SXT* gene cluster is an 819 bp nucleotide sequence set forth in SEQ ID NO: 64. The nucleotide sequence of *SXT* Gene 31 ranges from the nucleotide in position 35029 up to the nucleotide in position 35847 of SEQ ID NO: 1. The polypeptide sequence encoded by Gene 31 (*OMPR*) is set forth in SEQ ID NO: 65.

The 5' border region of *SXT* gene cluster comprises a 1320 bp gene (*orf1*), the sequence of which is set forth in SEQ ID NO: 2. The nucleotide sequence of *orf1* ranges from the nucleotide in position 1 up to the nucleotide in position 1320 of SEQ ID NO: 1. The polypeptide sequence encoded by *orf1* is set forth in SEQ ID NO: 3.

The 3' border region of *SXT* gene cluster comprises a 774 bp gene *(hisA)*, the sequence of which is set forth in SEQ ID NO: 66. The nucleotide sequence of *hisA* ranges from the nucleotide in position 35972 up to the nucleotide in position 36745 of SEQ ID NO: 1. The polypeptide sequence encoded by *hisA* is set forth in SEQ ID NO: 67.

The 3' border region of *SXT* gene cluster also comprises a 396 bp gene *(orfA)*, the sequence of which is set forth in in SEQ ID NO: 68. The nucleotide sequence of *orfA* ranges from the nucleotide in position 37060 up to the nucleotide in position 37455 of SEQ ID NO: 1. The polypeptide sequence encoded by *orfA* is set forth in SEQ ID NO: 69.

In accordance with other aspects and embodiments of the invention, the *CYR* gene cluster may have, but is not limited to, the nucleotide sequence as set forth SEQ ID NO: 80 (GenBank accession number EU140798), or display sufficient sequence identity thereto to hybridise to the sequence of SEQ ID NO: 80.

The *CYR* gene cluster comprises 15 genes and 14 intergenic regions.

Gene 1 of the *CYR* gene cluster is a 5631 bp nucleotide sequence set forth in SEQ ID NO: 81. The nucleotide sequence of *CYR* Gene 1 ranges from the nucleotide in position 444 up to the nucleotide in position 6074 of SEQ ID NO: 80. The polypeptide sequence encoded by Gene 1 (*CYRD*) is set forth in SEQ ID NO: 82.

Gene 2 of the *CYR* gene cluster is a 4074 bp nucleotide sequence set forth in SEQ ID NO: 83. The nucleotide sequence of *CYR* Gene 2 ranges from the nucleotide in position 6130 up to the nucleotide in position 10203 of SEQ ID NO: 80. The polypeptide sequence encoded by Gene 2 (*CYRF*) is set forth in SEQ ID NO: 84.

Gene 3 of the *CYR* gene cluster is a 1437 bp nucleotide sequence set forth in SEQ ID NO: 85. The nucleotide sequence of *CYR* Gene 3 ranges from the nucleotide in position 10251 up to the nucleotide in position 11687 of SEQ ID NO: 80. The polypeptide sequence encoded by Gene 3 (*CYRG*) is set forth in SEQ ID NO: 86.

Gene 4 of the *CYR* gene cluster is an 831 bp nucleotide sequence set forth in SEQ ID NO: 87. The nucleotide sequence of *CYR* Gene 4 ranges from the nucleotide in position 11741 up to the nucleotide in position 12571 of SEQ ID NO: 80. The polypeptide sequence encoded by Gene 4 (*CYRI*) is set forth in SEQ ID NO: 88.

Gene 5 of the *CYR* gene cluster is a 1398 bp nucleotide sequence set forth in SEQ ID NO: 89. The nucleotide sequence of *CYR* Gene 5 ranges from the nucleotide in position 12568 up to the nucleotide in position 13965 of SEQ ID NO: 80. The polypeptide sequence encoded by Gene 5 (*CYRK*) is set forth in SEQ ID NO: 90.

Gene 6 of the *CYR* gene cluster is a 750 bp nucleotide sequence set forth in SEQ ID NO: 91. The nucleotide sequence of *CYR* Gene 6 ranges from the nucleotide in position 14037 up to the nucleotide in position 14786 of SEQ ID NO: 80. The polypeptide sequence encoded by Gene 6 (*CYRL*) is set forth in SEQ ID NO: 92.

Gene 7 of the *CYR* gene cluster is a 1431 bp nucleotide sequence set forth in SEQ ID NO: 93. The nucleotide sequence of *CYR* Gene 7 ranges from the nucleotide in position 14886 up to the nucleotide in position 16316 of SEQ ID NO: 80. The polypeptide sequence encoded by Gene 7 (*CYRH*) is set forth in SEQ ID NO: 94.

Gene 8 of the *CYR* gene cluster is a 780 bp nucleotide sequence set forth in SEQ ID NO: 95. The nucleotide sequence of *CYR* Gene 8 ranges from the nucleotide in position 16893 up to the nucleotide in position 17672 of SEQ ID NO: 80. The polypeptide sequence encoded by Gene 8 (*CYRJ*) is set forth in SEQ ID NO: 96.

Gene 9 of the *CYR* gene cluster is an 1176 bp nucleotide sequence set forth in SEQ ID NO: 97. The nucleotide sequence of *CYR* Gene 9 ranges from the nucleotide in position 18113 up to the nucleotide in position 19288 of SEQ ID NO: 80. The polypeptide sequence encoded by Gene 9 (*CYRA*) is set forth in SEQ ID NO: 98.

Gene 10 of the *CYR* gene cluster is an 8754 bp nucleotide sequence set forth in SEQ ID NO: 99. The nucleotide sequence of *CYR* Gene 10 ranges from the nucleotide in position 19303 up to the nucleotide in position 28056 of SEQ ID NO: 80. The polypeptide sequence encoded by Gene 10 (*CYRB*) is set forth in SEQ ID NO: 100.

Gene 11 of the *CYR* gene cluster is a 5667 bp nucleotide sequence set forth in SEQ ID NO: 101. The nucleotide sequence of *CYR* Gene 11 ranges from the nucleotide in position 28061 up to the nucleotide in position 33727 of SEQ ID NO: 80. The polypeptide sequence encoded by Gene 11 (*CYRE*) is set forth in SEQ ID NO: 102.

Gene 12 of the *CYR* gene cluster is a 5004 bp nucleotide sequence set forth in SEQ ID NO: 103. The nucleotide sequence of *CYR* Gene 12 ranges from the nucleotide in position 34299 up to the nucleotide in position 39302 of SEQ ID NO: 80. The polypeptide sequence encoded by Gene 12 (*CYRC*) is set forth in SEQ ID NO: 104.

Gene 13 of the *CYR* gene cluster is a 318 bp nucleotide sequence set forth in SEQ ID NO: 105. The nucleotide sequence of *CYR* Gene 13 ranges from the nucleotide in position 39366 up to the nucleotide in position 39683 of SEQ ID NO: 80. The polypeptide sequence encoded by Gene 13 (*CYRM*) is set forth in SEQ ID NO: 106.

Gene 14 of the *CYR* gene cluster is a 600 bp nucleotide sequence set forth in SEQ ID NO: 107. The nucleotide sequence of *CYR* Gene 14 ranges from the nucleotide in

position 39793 up to the nucleotide in position 40392 of SEQ ID NO: 80. The polypeptide sequence encoded by Gene 14 (*CYRN*) is set forth in SEQ ID NO: 108.

Gene 15 of the *CYR* gene cluster is a 1548 bp nucleotide sequence set forth in SEQ ID NO: 109. The nucleotide sequence of *CYR* Gene 15 ranges from the nucleotide in position 40501 up to the nucleotide in position 42048 of SEQ ID NO: 80. The polypeptide sequence encoded by Gene 15 (*CYRO*) is set forth in SEQ ID NO: 110.

In general, the nucleic acids and polypeptides of the invention are of an isolated or purified form.

In addition to the *SXT* and *CYR* polynucleotides and polypeptide sequences set forth herein, also included within the scope of the present invention are variants and fragments thereof.

SXT and *CYR* polynucleotides disclosed herein may be deoxyribonucleic acids (DNA), ribonucleic acids (RNA) or complementary deoxyribonucleic acids (cDNA).

RNA may be derived from RNA polymerase-catalyzed transcription of a DNA sequence. The RNA may be a primary transcript derived transcription of a corresponding DNA sequence. RNA may also undergo post-transcriptional processing. For example, a primary RNA transcript may undergo post-transcriptional processing to form a mature RNA. Messenger RNA (mRNA) refers to RNA derived from a corresponding open reading frame that may be translated into protein by the cell. cDNA refers to a double-stranded DNA that is complementary to and derived from mRNA. Sense RNA refers to RNA transcript that includes the mRNA and so can be translated into protein by the cell. Antisense RNA refers to an RNA transcript that is complementary to all or part of a target primary transcript or mRNA and may be used to block the expression of a target gene.

The skilled addresse will recognise that RNA and cDNA sequences encoded by the *SXT* and *CYR* DNA sequences disclosed herein may be derived using the genetic code. An RNA sequence may be derived from a given DNA sequence by generating a sequence that is complementary the particular DNA sequence. The complementary sequence may be generated by converting each cytosine ('C') base in the DNA sequence to a guanine ('G') base, each guanine ('G') base in the DNA sequence to a cytosine ('C') base, each thymidine ('T') base in the DNA sequence to an adenine ('A') base, and each adenine ('A') base in the DNA sequence to a uracil ('U') base.

A complementary DNA (cDNA) sequence may be derived from a DNA sequence by deriving an RNA sequence from the DNA sequence as above, then converting the RNA sequence into a cDNA sequence. An RNA sequence can be converted into a cDNA

sequence by converting each cytosine ('C') base in the RNA sequence to a guanine ('G') base, each guanine ('G') base in the RNA sequence to a cytosine ('C') base, each uracil ('U') base in the RNA sequence to an adenine ('A') base, and each adeneine ('A') base in the RNA sequence to a thymidine ('T') base.

The term "variant" as used herein refers to a substantially similar sequence. In general, two sequences are "substantially similar" if the two sequences have a specified percentage of amino acid residues or nucleotides that are the same (percentage of "sequence identity"), over a specified region, or, when not specified, over the entire sequence. Accordingly, a "variant" of a polynucleotide and polypeptide sequence disclosed herein may share at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 83% 85%, 88%, 90%, 93%, 95%, 96%, 97%, 98% or 99% sequence identity with the reference sequence.

In general, polypeptide sequence variants possess qualitative biological activity in common. Polynucleotide sequence variants generally encode polypeptides which generally possess qualitative biological activity in common. Also included within the meaning of the term "variant" are homologues of polynucleotides and polypeptides of the invention. A polynucleotide homologue is typically from a different bacterial species but sharing substantially the same biological function or activity as the corresponding polynucleotide disclosed herein. A polypeptide homologue is typically from a different bacterial species but sharing substantially the same biological function or activity as the corresponding polypeptide disclosed herein. For example, homologues of the polynucleotides and polypeptides disclosed herein include, but are not limited to those from different species of cyanobacteria.

Further, the term "variant" also includes analogues of the polypeptides of the invention. A polypeptide "analogue" is a polypeptide which is a derivative of a polypeptide of the invention, which derivative comprises addition, deletion, substitution of one or more amino acids, such that the polypeptide retains substantially the same function. The term "conservative amino acid substitution" refers to a substitution or replacement of one amino acid for another amino acid with similar properties within a polypeptide chain (primary sequence of a protein). For example, the substitution of the charged amino acid glutamic acid (Glu) for the similarly charged amino acid aspartic acid (Asp) would be a conservative amino acid substitution.

In general, the percentage of sequence identity between two sequences may be determined by comparing two optimally aligned sequences over a comparison window.

The portion of the sequence in the comparison window may, for example, comprise deletions or additions (*i.e.* gaps) in comparison to the reference sequence (for example, a polynucleotide or polypeptide sequence disclosed herein), which does not comprise deletions or additions, in order to align the two sequences optimally. A percentage of sequence identity may then be calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison, and multiplying the result by 100 to yield the percentage of sequence identity.

In the context of two or more nucleic acid or polypeptide sequences, the percentage of sequence identity refers to the specified percentage of amino acid residues or nucleotides that are the same over a specified region, (or, when not specified, over the entire sequence), when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters. Methods of alignment of sequences for comparison are well known in the art. Optimal alignment of sequences for comparison can be determined conventionally using known computer programs, including, but not limited to: CLUSTAL in the PC/Gene program (available from Intelligenetics, Mountain View, California); the ALIGN program (Version 2.0) and GAP, BESTFIT, BLAST, FASTA, and TFASTA in the GCG Wisconsin Genetics Software Package, Version 10 (available from Accelrys Inc., 9685 Scranton Road, San Diego, California, USA).

The BESTFIT program (Wisconsin Sequence Analysis Package, for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, Wis. 53711) uses the local homology algorithm of Smith and Waterman to find the best segment of homology between two sequences (Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981)). When using BESTFIT or any other sequence alignment

program to determine the degree of homology between sequences, the parameters may be set such that the percentage of identity is calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number of nucleotides in the reference sequence are allowed.

GAP uses the algorithm described in Needleman and Wunsch (1970) J. Mol. Biol. 48:443-453, to find the alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps. GAP considers all possible alignments and gap positions and creates the alignment with the largest number of matched bases and the fewest gaps. It allows for the provision of a gap creation penalty and a gap extension penalty in units of matched bases. GAP presents one member of the family of best alignments.

Another method for determining the best overall match between a query sequence and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag and colleagues (Comp. App. Biosci. 6:237-245 (1990)). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity.

The BLAST and BLAST 2.0 algorithms, may be used for determining percent sequence identity and sequence similarity. These are described in Altschul et al. (1977) Nuc. Acids Res. 25:3389-3402, and Altschul et al (1990) J. Mol. Biol. 215:403-410, respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al, supra). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the

accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) or 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) Proc. Natl, Acad. Sci USA 89:10915) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands. [0028] The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin and Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873-5787). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001.

The invention also contemplates fragments of the polypeptides disclosed herein. A polypeptide "fragment" is a polypeptide molecule that encodes a constituent or is a constituent of a polypeptide of the invention or variant thereof. Typically the fragment possesses qualitative biological activity in common with the polypeptide of which it is a constituent. The peptide fragment may be between about 5 to about 3000 amino acids in length, between about 5 to about 2750 amino acids in length, between about 5 to about 2500 amino acids in length, between about 5 to about 2250 amino acids in length, between about 5 to about 2000 amino acids in length, between about 5 to about 1750 amino acids in length, between about 5 to about 1500 amino acids in length, between about 5 to about 1250 amino acids in length, between about 5 to about 1000 amino acids in length, between about 5 to about 900 amino acids in length, between about 5 to about 800 amino acids in length, between about 5 to about 700 amino acids in length, between about 5 to about 600 amino acids in length, between about 5 to about 500 amino acids in length, between about 5 to about 450 amino acids in length, between about 5 to about 400 amino acids in length, between about 5 to about 350 amino acids in length, between about 5 to about 300 amino acids in length, between about 5 to about 250 amino acids in length, between about 5 to about 200 amino acids in length, between about 5 to about 175 amino acids in length, between about 5 to about 150 amino acids in length, between about 5 to

about 125 amino acids in length, between about 5 to about 100 amino acids in length, between about 5 to about 75 amino acids in length, between about 5 to about 50 amino acids in length, between about 5 to about 40 amino acids in length, between about 5 to about 30 amino acids in length, between about 5 to about 20 amino acids in length, and between about 5 to about 15 amino acids in length. Alternatively, the peptide fragment may be between about 5 to about 10 amino acids in length.

Also contemplated are fragments of the polynucleotides disclosed herein. A polynucleotide "fragment" is a polynucleotide molecule that encodes a constituent or is a constituent of a polynucleotide of the invention or variant thereof. Fragments of a polynucleotide do not necessarily need to encode polypeptides which retain biological activity. The fragment may, for example, be useful as a hybridization probe or PCR primer. The fragment may be derived from a polynucleotide of the invention or alternatively may be synthesized by some other means, for example by chemical synthesis.

Certain embodiments of the invention relate to fragments of SEQ ID NO: 1. A fragment of SEQ ID NO: 1 may comprise, for example, a constituent of SEQ ID NO: 1 in which the 5' gene border region gene *orf1* is absent. Alternatively, a fragment of SEQ ID NO: 1 may comprise, for example, a constituent of SEQ ID NO: 1 in which the 3' gene border region gene *hisA* is absent. Alternatively, a fragment of SEQ ID NO: 1 may comprise, for example, a constituent of SEQ ID NO: 1 in which the 3' gene border region gene *orfA* is absent. Alternatively, a fragment of SEQ ID NO: 1 may comprise, for example, a constituent of SEQ ID NO: 1 in which the 5' gene border region gene *orf1* is absent and the 3' border region gene *orfA* is absent. Alternatively, a fragment of SEQ ID NO: 1 may comprise, for example, a constituent of SEQ ID NO: 1 in which the 5' gene border region gene *orf1* is absent and the 3' border region genes *hisA* and *orfA* are absent.

In other embodiments, a fragment of SEQ ID NO: 1 may comprise one or more *SXT* open reading frames. The *SXT* open reading frame may be selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID

NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants thereof.

Additional embodiments of the invention relate to fragments of SEQ ID NO: 80. The fragment of SEQ ID NO: 80 may comprise one or more *CYR* open reading frames. The *CYR* open reading frame may be selected from the group consisting of of SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO:107, SEQ ID NO: 109, and variants thereof.

In particular embodiments, the polynucleotides of the invention may be cloned into a vector. The vector may comprise, for example, a DNA, RNA or complementary DNA (cDNA) sequence. The vector may be a plasmid vector, a viral vector, or any other suitable vehicle adapted for the insertion of foreign sequences, their introduction into cells and the expression of the introduced sequences. Typically the vector is an expression vector and may include expression control and processing sequences such as a promoter, an enhancer, ribosome binding sites, polyadenylation signals and transcription termination sequences. The invention also contemplates host cells transformed by such vectors. For example, the polynucleotides of the invention may be cloned into a vector which is transformed into a bacterial host cell, for example *E. coli*. Methods for the construction of vectors and their transformation into host cells are generally known in the art, and described in, for example, *Molecular Cloning: A Laboratory Manual* (2nd ed., Cold Spring Harbor Laboratory Press, Plainview, New York, and, Ausubel F. M. et al. (Eds) *Current Protocols in Molecular* Biology (2007), John Wiley and Sons, Inc.

Nucleotide Probes, Primers and Antibodies

The invention contemplates nucleotides and fragments based on the sequences of the polynucleotides disclosed herein for use as primers and probes for the identification of homologous sequences.

The nucleotides and fragments may be in the form of oligonucleotides. Oligonucleotides are short stretches of nucleotide residues suitable for use in nucleic acid amplification reactions such as PCR, typically being at least about 5 nucleotides to about 80 nucleotides in length, more typically about 10 nucleotides in length to about 50 nucleotides in lenght, and even more typically about 15 nucleotides in length to about 30 nucleotides in length.

Probes are nucleotide sequences of variable length, for example between about 10 nucleotides and several thousand nucleotides, for use in detection of homologous sequences, typically by hybridization. Hybridization probes may be genomic DNA fragments, cDNA fragments, RNA fragments, or other oligonucleotides.

Methods for the design and/or production of nucleotide probes and/or primers are generally known in the art, and are described in Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (2nd ed., Cold Spring Harbor Laboratory Press, Plainview, New York; Itakura K. et al. (1984) Annu. Rev. Biochem. 53:323; Innis *et al.*, (Eds) (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic Press, New York); Innis and Gelfand, (Eds) (1995) *PCR Strategies* (Academic Press, New York); and Innis and Gelfand, (Eds) (1999) *PCR Methods Manual* (Academic Press, New York). Nucleotide primers and probes may be prepared, for example, by chemical synthesis techniques for example, the phosphodiester and phosphotriester methods (see for example Narang S. A. et al. (1979) Meth. Enzymol. 68:90; Brown, E. L. (1979) et al. Meth. Enzymol. 68:109; and U.S. Patent No. 4356270), the diethylphosphoramidite method (see Beaucage S.L et al. (1981) Tetrahedron Letters, 22:1859-1862). A method for synthesizing oligonucleotides on a modified solid support is described in U.S. Patent No. 4458066.

The nucleic acids of the invention, including the above-mentioned probes and primers, may be labelled by incorporation of a marker to facilitate their detection. Techniques for labelling and detecting nucleic acids are described, for example, in Ausubel F. M. et al. (Eds) *Current Protocols in Molecular* Biology (2007), John Wiley and Sons, Inc. Examples of suitable markers include fluorescent molecules (*e.g.* acetylaminofluorene, 5-bromodeoxyuridine, digoxigenin, fluorescein) and radioactive isotopes (*e.g.* 32P, 35S, 3H, 33P). Detection of the marker may be achieved, for example, by chemical, photochemical, immunochemical, biochemical, or spectroscopic techniques.

The probes and primers of the invention may be used, for example, to detect or isolate cyanobacteria and/or dinoflagellates in a sample of interest. Additionally or alternatively, the probes and primers of the invention may be used to detect or isolate a cyanotoxic organism and/or a cylindrospermopisn-producing organism in a sample of interest. Additionally or alternatively, the probes or primers of the invention may be used to isolate corresponding sequences in other organisms including, for example, other bacterial species. Methods such as the polymerase chain reaction (PCR), hybridization, and the like can be used to identify such sequences based on their sequence homology to

the sequences set forth herein. Sequences that are selected based on their sequence identity to the entire sequences set forth herein or to fragments thereof are encompassed by the embodiments. Such sequences include sequences that are orthologs of the disclosed sequences. The term "orthologs" refers to genes derived from a common ancestral gene and which are found in different species as a result of speciation. Genes found in different species are considered orthologs when their nucleotide sequences and/or their encoded protein sequences share substantial identity as defined elsewhere herein. Functions of orthologs are often highly conserved among species.

In hybridization techniques, all or part of a known nucleotide sequence is used to generate a probe that selectively hybridizes to other corresponding nucleic acid sequences present in a given sample. The hybridization probes may be genomic DNA fragments, cDNA fragments, RNA fragments, or other oligonucleotides, and may be labelled with a detectable marker. Thus, for example, probes for hybridization can be made by labelling synthetic oligonucleotides based on the sequences of the invention.

The level of homology (sequence identity) between probe and the target sequence will largely be determined by the stringency of hybridization conditions. In particular the nucleotide sequence used as a probe may hybridize to a homologue or other variant of a polynucleotide disclosed herein under conditions of low stringency, medium stringency or high stringency. There are numerous conditions and factors, well known to those skilled in the art, which may be employed to alter the stringency of hybridization. For instance, the length and nature (DNA, RNA, base composition) of the nucleic acid to be hybridized to a specified nucleic acid; concentration of salts and other components, such as the presence or absence of formamide, dextran sulfate, polyethylene glycol etc; and altering the temperature of the hybridization and/or washing steps.

Typically, stringent hybridization conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (*e.g.*, 10 to 50 nucleotides) and at least about 60°C for long probes (*e.g.*, greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringency conditions include hybridization with a buffer solution of 30% to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulfate) at 37 °C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50°C to 55 °C. Exemplary moderate stringency conditions include hybridization in 40% to 45% formamide, 1.0 M NaCl, 1% SDS at

37 °C, and a wash in 0.5X to 1X SSC at 55°C to 60 °C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37 °C, and a final wash in 0.1X SSC at 60°C to 65 °C for at least about 20 minutes. Optionally, wash buffers may comprise about 0.1% to about 1% SDS. The duration of hybridization is generally less than about 24 hours, usually about 4 to about 12 hours.

Under a PCR approach, oligonucleotide primers can be designed for use in PCR reactions to amplify corresponding DNA sequences from cDNA or genomic DNA extracted from any organism of interest. Methods for designing PCR primers and PCR cloning are generally known in the art and are disclosed in Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2nd ed., Cold Spring Harbor Laboratory Press, Plainview, New York); Ausubel F. M. et al. (Eds) *Current Protocols in Molecular* Biology (2007), John Wiley and Sons, Inc; Maniatis et al. *Molecular Cloning* (1982), 280-281; Innis et al. (Eds) (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic Press, New York); Innis and Gelfand, (Eds) (1995) *PCR Strategies* (Academic Press, New York); and Innis and Gelfand, (Eds) (1999) *PCR Methods Manual* (Academic Press, New York). Known methods of PCR include, but are not limited to, methods using paired primers, nested primers, single specific primers, degenerate primers, gene-specific primers, vector-specific primers, partially-mismatched primers, and the like.

The skilled addressee will recognise that the primers described herein for use in PCR or RT-PCR may also be used as probes for the detection of *SXT* or *CYR* sequences.

Also contemplated by the invention are antibodies which are capable of binding specifically to the polypeptides of the invention. The antibodies may be used to qualitatively or quantitatively detect and analyse one or more *SXT* or *CYR* polypeptides in a given sample. By "binding specifically" it will be understood that the antibody is capable of binding to the target polypeptide or fragment thereof with a higher affinity than it binds to an unrelated protein. For example, the antibody may bind to the polypeptide or fragment thereof with a binding constant in the range of at least about 10^{-4}M to about 10^{-10}M. Preferably the binding constant is at least about 10^{-5}M, or at least about 10^{-6}M, more preferably the binding constant of the antibody to the *SXT* or *CYR* polypeptide or fragment thereof is at least about 10^{-7}M, at least about 10^{-8}M, or at least about 10^{-9}M or more.

Antibodies of the invention may exist in a variety of forms, including for example as a whole antibody, or as an antibody fragment, or other immunologically active fragment thereof, such as complementarity determining regions. Similarly, the antibody

may exist as an antibody fragment having functional antigen-binding domains, that is, heavy and light chain variable domains. Also, the antibody fragment may exist in a form selected from the group consisting of, but not limited to: Fv, F_{ab}, $F(ab)_2$, scFv (single chain Fv), dAb (single domain antibody), chimeric antibodies, bi-specific antibodies, diabodies and triabodies.

An antibody 'fragment' may be produced by modification of a whole antibody or by synthesis of the desired antibody fragment. Methods of generating antibodies, including antibody fragments, are known in the art and include, for example, synthesis by recombinant DNA technology. The skilled addressee will be aware of methods of synthesising antibodies, such as those described in, for example, US Patent No. 5296348 and Ausubel F. M. et al. (Eds) *Current Protocols in Molecular* Biology (2007), John Wiley and Sons, Inc.

Preferably antibodies are prepared from discrete regions or fragments of the *SXT* or *CYR* polypeptide of interest. An antigenic portion of a polypeptide of interest may be of any appropriate length, such as from about 5 to about 15 amino acids. Preferably, an antigenic portion contains at least about 5, 6, 7, 8, 9, 10, 11, 12, 13 or 14 amino acid residues.

In the context of this specification reference to an antibody specific to a *SXT* or *CYR* polypeptide of the invention includes an antibody that is specific to a fragment of the polypeptide of interest.

Antibodies that specifically bind to a polypeptide of the invention can be prepared, for example, using the purified *SXT* or *CYR* polypeptides or their nucleic acid sequences using any suitable methods known in the art. For example, a monoclonal antibody, typically containing Fab portions, may be prepared using hybridoma technology described in Harlow and Lane (Eds) *Antibodies - A Laboratory Manual*, (1988), Cold Spring Harbor Laboratory, N.Y; Coligan, *Current Protocols in Immunology* (1991); Goding, *Monoclonal Antibodies: Principles and Practice* (1986) 2nd ed; and Kohler & Milstein, (1975) Nature 256: 495-497. Such techniques include, but are not limited to, antibody preparation by selection of antibodies from libraries of recombinant antibodies in phage or similar vectors, as well as preparation of polyclonal and monoclonal antibodies by immunizing rabbits or mice (see, for example, Huse et al. (1989) Science 246: 1275-1281; Ward et al. (1989) Nature 341: 544-546).

It will also be understood that antibodies of the invention include humanised antibodies, chimeric antibodies and fully human antibodies. An antibody of the invention

may be a bi-specific antibody, having binding specificity to more than one antigen or epitope. For example, the antibody may have specificity for one or more *SXT* or *CYR* polypeptide or fragments thereof, and additionally have binding specificity for another antigen. Methods for the preparation of humanised antibodies, chimeric antibodies, fully human antibodies, and bispecific antibodies are known in the art and include, for example as described in United States Patent No. 6995243 issued February 7, 2006 to Garabedian, et al. and entitled "Antibodies that recognize and bind phosphorylated human glucocorticoid receptor and methods of using same".

Generally, a sample potentially comprising *SXT* or *CYR* polypeptides can be contacted with an antibody that specifically binds the *SXT* or *CYR* polypeptide or fragment thereof. Optionally, the antibody can be fixed to a solid support to facilitate washing and subsequent isolation of the complex, prior to contacting the antibody with a sample. Examples of solid supports include, for example, microtitre plates, beads, ticks, or microbeads. Antibodies can also be attached to a ProteinChip array or a probe substrate as described above.

Detectable labels for the identification of antibodies bound to the *SXT* or *CYR* polypeptides of the invention include, but are not limited to fluorochromes, fluorescent dyes, radiolabels, enzymes such as horse radish peroxide, alkaline phosphatase and others commonly used in the art, and colorimetric labels including colloidal gold or coloured glass or plastic beads. Alternatively, the marker in the sample can be detected using an indirect assay, wherein, for example, a second, labelled antibody is used to detect bound marker-specific antibody.

Methods for detecting the presence of or measuring the amount of, an antibody-marker complex include, for example, detection of fluorescence, chemiluminescence, luminescence, absorbance, birefringence, transmittance, reflectance, or refractive index such as surface plasmon resonance, ellipsometry, a resonant mirror method, a grating coupler wave guide method or interferometry. Radio frequency methods include multipolar resonance spectroscopy. Electrochemical methods include amperometry and voltametry methods. Optical methods include imaging methods and non-imaging methods and microscopy.

Useful assays for detecting the presence of or measuring the amount of, an antibody-marker complex include, include, for example, enzyme-linked immunosorbent assay (ELISA), a radioimmune assay (RIA), or a Western blot assay. Such methods are described in, for example, Clinical Immunology (Stites & Terr, eds., 7th ed. 1991);

Methods in Cell Biology: Antibodies in Cell Biology, volume 37 (Asai, ed. 1993); and Harlow & Lane, supra.

Methods and kits for detection

The invention provides methods and kits for the detection and/or isolation of *SXT* nucleic acids and polypeptides. Also provided are methods and kits for the detection and/or isolation *CYR* nucleic acids and polypeptides.

In one aspect, the invention provides a method for the detection of cyanobacteria. The skilled addressee will understand that the detection of "cyanobacteria" encompasses the detection of one or more cyanobacteria. The method comprises obtaining a sample for use in the method, and detecting the presence of one or more *SXT* polynucleotides or polypeptides as disclosed herein, or a variant or fragment thereof. The presence of *SXT* polynucleotides, polypeptides, or variants or fragments thereof, is indicative of cyanobacteria in the sample.

The *SXT* polynucleotide may comprise a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants and fragments thereof.

Alternatively, the *SXT* polynucleotide may be an RNA or cDNA encoded by a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66,

SEQ ID NO: 68, and variants and fragments thereof and/or polypeptides as disclosed herein, or a variant or fragment thereof.

The *SXT* polypeptide may comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 67, SEQ ID NO: 69, and variants and fragments thereof.

The inventors have determined that several genes of the *SXT* gene cluster exist in saxitoxin-producing organisms, and are absent in organisms with the *SXT* gene cluster that do not produce saxitoxin. Specifically, the inventors have identified that gene 6 (*sxtA*) (SEQ ID NO: 14), gene 9 (*sxtG*) (SEQ ID NO: 20), gene 10 (*sxtH*) (SEQ ID NO: 22), gene 11 (*sxtI*) (SEQ ID NO: 24) and gene 17 (*sxtX*) (SEQ ID NO: 36) of the *SXT* gene cluster are present only in organisms that produce saxitoxin.

Accordingly, in another aspect the invention provides a method of detecting a cyanotoxic organism. The method comprises obtaining a sample for use in the method, and detecting a cyanotoxic organism based on the detection of one or more *SXT* polynucleotides comprising a sequence set forth in SEQ ID NO: 14 (*sxtA*, gene 6), SEQ ID NO: 20 (*sxtG*, gene 9), SEQ ID NO: 22 (*sxtH,* gene 10), SEQ ID NO: 24 (*sxtI*, gene 11), SEQ ID NO: 36 (*sxtX*, gene 17), or variants or fragments thereof. Additionally or alternatively, a cyanotoxic organism may be detected based on the detection of an RNA or cDNA comprising a sequence encoded by SEQ ID NO: 14 (*sxtA*, gene 6), SEQ ID NO: 20 (*sxtG*, gene 9), SEQ ID NO: 22 (*sxtH,* gene 10), SEQ ID NO: 24 (*sxtI*, gene 11), SEQ ID NO: 36 (*sxtX*, gene 17), or variants or fragments thereof. Additionally or alternatively, a cyanotoxic organism may be detected based on the detection of one or more polypeptides comprising a sequence set forth in SEQ ID NO: 15 (*SXTA*), SEQ ID NO: 21 (*SXTG*), SEQ ID NO: 23 (*SXTH*), SEQ ID NO: 25 (*SXTI*), SEQ ID NO: 37 (*SXTX*), or variants or fragments thereof, in a sample suspected of comprising one or more cyanotoxic organisms. The cyanotoxic organism may be any organism capable of producing saxitoxin. In a preferred embodiment of the invention, the cyanotoxic organism is a cyanobacteria or a dinoflagellate.

In certain embodiments of the invention, the methods for detecting cyanobacteria or the methods for detecting cyanotoxic organisms may further comprise the detection of one or more *CYR* polynucleotides or *CYR* polypeptides as disclosed herein, or a variant or fragment thereof. The *CYR* polynucleotide may comprise a sequence selected from the group consisting of SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, and variants or fragments thereof.

Alternatively, the *CYR* polynucleotide may be an RNA or cDNA encoded by a polynucleotide sequence selected from the group consisting of SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, and variants or fragments thereof.

The *CYR* polypeptide may comprise a sequence selected from the group consisting of SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92 , SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, and SEQ ID NO: 110, and variants or fragments thereof.

The inventors have determined gene 8 (*cyrJ*) (SEQ ID NO: 95) of the *CYR* gene cluster exists in cylindrospermopsin-producing organisms, and is absent in organisms with the *CYR* gene cluster that do not produce cylindrospermopsin. Accordingly, the methods for detecting cyanobacteria or the methods for detecting cyanotoxic organisms may further comprise the detection of a cylindrospermopsin-producing organism based on the detection of a *CYR* polynucleotide comprising a sequence set forth in SEQ ID NO: 95, or a variant or fragment thereof. Additionally or alternatively, the methods for detecting cyanobacteria or the methods for detecting cyanotoxic organisms may further comprise the detection of a cylindrospermopsin-producing organism based on the detection of an RNA or cDNA comprising a sequence encoded by SEQ ID NO: 95, or a variant or fragment thereof. Additionally or alternatively, the methods for detecting cyanobacteria or the methods for detecting cyanotoxic organisms may further comprise the detection of a cylindrospermopsin-producing organism based on the detection of a *CYR* polypeptide comprising a sequence set forth in SEQ ID NO: 96, or a variant or fragment thereof.

In another aspect, the invention provides a method for the detection of cyanobacteria. The skilled addressee will understand that the detection of "cyanobacteria" encompasses the detection of one or more cyanobacteria. The method comprises obtaining a sample for use in the method, and detecting the presence of one or more *CYR* polynucleotides or polypeptides as disclosed herein, or a variant or fragment thereof. The presence of *CYR* polynucleotides, polypeptides, or variants or fragments thereof, is indicative of cyanobacteria in the sample.

The *CYR* polynucleotide may comprise a sequence selected from the group consisting of SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO:107, SEQ ID NO: 109 and variants and fragments thereof.

Alternatively, the *CYR* polynucleotide may be an RNA or cDNA encoded by a sequence selected from the group consisting of SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO:107, SEQ ID NO: 109 and variants and fragments thereof.

The *CYR* polypeptide may comprise a sequence selected from the group consisting of SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92 , SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, and SEQ ID NO: 110, and variants or fragments thereof.

In another aspect of the invention there is provided a method of detecting a cylindrospermopsin-producing organism based on the detection of *CYR* gene 8 (*cyrJ*). The method comprises obtaining a sample for use in the method, and detecting the presence of a *CYR* polynucleotide comprising a sequence set forth in SEQ ID NO: 95, or a variant or fragment thereof. Additionally or alternatively, the method for detecting a cylindrospermopsin-producing organism based on the detection of *CYR* gene 8 (*cyrJ*) may comprise the detection of an RNA or cDNA comprising a sequence encoded by SEQ ID NO: 95, or a variant or fragment thereof. Additionally or alternatively, the method for detecting a cylindrospermopsin-producing organism based on the detection of *CYR* gene 8 (*cyrJ*) may comprise the detection of a *CYR* polypeptide comprising a sequence set forth in SEQ ID NO: 96, or a variant or fragment thereof.

In certain embodiments of the invention, the methods for detecting cyanobacteria comprising the detection of *CYR* sequences or variants or fragments thereof further comprise the detection of one or more *SXT* polynucleotides or *SXT* polypeptides as disclosed herein, or a variant or fragment thereof.

The *SXT* polynucleotide may comprise a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants and fragments thereof.

Alternatively, the *SXT* polynucleotide may be an RNA or cDNA encoded by a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants and fragments thereof and/or polypeptides as disclosed herein, or a variant or fragment thereof.

The *SXT* polypeptide may comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 67, SEQ ID NO: 69, and variants and fragments thereof.

In another aspect, the invention provides a method for the detection of dinoflagellates. The skilled addressee will understand that the detection of

"dinoflagellates" encompasses the detection of one or more dinoflagellates. The method comprises obtaining a sample for use in the method, and detecting the presence of one or more *SXT* polynucleotides or polypeptides as disclosed herein, or a variant or fragment thereof. The presence of *SXT* polynucleotides, polypeptides, or variants or fragments thereof, is indicative of dinoflagellates in the sample.

The *SXT* polynucleotide may comprise a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants and fragments thereof.

Alternatively, the *SXT* polynucleotide may be an RNA or cDNA encoded by a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants and fragments thereof and/or polypeptides as disclosed herein, or a variant or fragment thereof.

The *SXT* polypeptide may comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 67, SEQ ID NO: 69, and variants and fragments thereof.

A sample for use in accordance with the methods described herein may be suspected of comprising one or more cyanotoxic organisms. The cyanotoxic organisms may be one or more cyanobacteria and/or one or more dinoflagellates. Additionally or alternatively, a sample for use in accordance with the methods described herein may be suspected of comprising one more cyanobacteria and/or one or more dinoflagellates. A sample for use in accordance with the methods described herein may be a comparative or control sample, for example, a sample comprising a known concentration or density of a cyanobacteria and/or dinoflagellates, or a sample comprising one or more known species or strains of cyanobacteria and/or dinoflagellates.

A sample for use in accordance with the methods described herein may be derived from any source. For example, a sample may be an environmental sample. The environmental sample may be derived, for example, from salt water, fresh water or a blue-green algal bloom. Alternatively, the sample may be derived from a laboratory source, such as a culture, or a commercial source.

It will be appreciated by those in the art that the methods and kits disclosed herein are generally suitable for detecting any organisms in which the *SXT* and/or *CYR* gene clusters are present. Suitable cyanobacteria to which the methods of the invention are applicable may be selected from the orders Oscillatoriales, Chroococcales, Nostocales and Stigonematales. For example, the cyanobacteria may be selected from the genera *Anabaena, Nostoc, Microcystis, Planktothrix, Oscillatoria, Phormidium*, and *Nodularia*. For example, the cyanobacteria may be selected from the species *Cylindrospermopsis raciborskii* T3, *Cylindrospermopsis raciborskii* AWT205, *Aphanizomenon ovalisporum, Aphanizomenon flos-aquae, Aphanizomenon sp., Umezakia natans, Raphidiopsis curvata, Anabaena bergii, Lyngbya wollei,* and *Anabaena* circinalis. Examples of suitable dinoflagellates to which the methods and kits of the invention are applicable may be selected from the genera *Alexandrium, Pyrodinium* and *Gymnodinium*. The methods and kits of the invention may also be employed for the discovery of novel hepatotoxic species or genera in culture collections or from environmental samples. The methods and kits of the invention may also be employed to detect cyanotoxins that accumulate in other animals, for example, fish and shellfish.

Detection of *SXT* and *CYR* polynucleotides and polypeptides disclosed herein may be performed using any suitable method. For example, methods for the detection of *SXT* and *CYR* polynucleotides and/or polypeptides disclosed herein may involve the use of a primer, probe or antibody specific for one or more *SXT* and *CYR* polynucleotides and

polypeptides. Suitable techniques and assays in which the skilled addressee may utilise a primer, probe or antibody specific for one or more *SXT* and *CYR* polynucleotides and polypeptides include, for example, the polymerase chain reaction (and related variations of this technique), antibody based assays such as ELISA and flow cytometry, and fluorescent microscopy. Methods by which the *SXT* and *CYR* polypeptides disclosed herein may be identified are generally known in the art, and are described for example in Coligan J. E. et al. (Eds) *Current Protocols in Protein Science* (2007), John Wiley and Sons, Inc; Walker, J. M., (Ed) (1988) *New Protein Techniques: Methods in Molecular Biology*, Humana Press, Clifton, N.J; and Scopes, R. K. (1987) *Protein Purification: Principles and Practice,* 3rd. Ed., Springer-Verlag, New York, N.Y. For example, *SXT* and *CYR* polypeptides disclosed herein may be detected by western blot or spectrophotometric analysis. Other examples of suitable methods for the detection of *SXT* and *CYR* polypeptides are described, for example, in US Patent No. 4683195, US Patent No. 6228578, US Patent No. 7282355, US Patent No. 7348147 and PCT publication No. W0/2007/056723.

In a preferred embodiment of the invention, the detection of *SXT* and *CYR* polynucleotides and polypeptides is achieved by amplification of DNA from the sample of interest by polymerase chain reaction, using primers that hybridise specifically to the *SXT* and/or *CYR* sequence, or a variant or fragment thereof, and detecting the amplified sequence.

Nucleic acids and polypeptides for analysis using methods and kits disclosed herein may be extracted from organisms either in mixed culture or as individual species or genus isolates. Accordingly, the organisms may be cultured prior to nucleic acid and/or polypeptide isolation or alternatively nucleic acid and/or polypeptides may be extracted directly from environmental samples, such as water samples or blue-green algal blooms.

Suitable methods for the extraction and purification of nucleic acids for analysis using the methods and kits invention are generally known in the art and are described, for example, in Ausubel F. M. et al. (Eds) *Current Protocols in Molecular* Biology (2007), John Wiley and Sons, Inc; Neilan (1995) Appl. Environ. Microbiol. 61:2286-2291; and Neilan et al. (2002) Astrobiol. 2:271-280. The skilled addressee will readily appreciate that the invention is not limited to the specific methods for nucleic acid isolation described therein and other suitable methods are encompassed by the invention. The invention may be performed without nucleic acid isolation prior to analysis of the nucleic acid.

Suitable methods for the extraction and purification of polypeptides for the purposes of the invention are generally known in the art and are described, for example, in Coligan J. E. et al. (Eds) *Current Protocols in Protein Science* (2007), John Wiley and Sons, Inc; Walker, J. M., (Ed) (1988) *New Protein Techniques: Methods in Molecular Biology*, Humana Press, Clifton, N.J; and Scopes, R. K. (1987) *Protein Purification: Principles and Practice,* 3rd. Ed., Springer-Verlag, New York, N.Y. Examples of suitable techniques for protein extraction include, but are not limited to dialysis, ultrafiltration, and precipitation. Protein purification techniques suitable for use include, but are not limited to, reverse-phase chromatography, hydrophobic interaction chromatography, centrifugation, gel filtration, ammonium sulfate precipitation, and ion exchange.

In accordance with the methods and kits of the invention, *SXT* and *CYR* polynucleotides or variants or fragments thereof may be detected by any suitable means known in the art. In a preferred embodiment of the invention, *SXT* and *CYR* polynucleotides are detected by PCR amplification. Under the PCR approach, oligonucleotide primers can be designed for use in PCR reactions to amplify *SXT* and *CYR* polynucleotides of the invention. Also encompassed by the invention is the PCR amplification of complementary DNA (cDNA) amplified from messenger RNA (mRNA) derived from reverse-transcription of *SXT* and *CYR* sequences (RT-PCR). Known methods of PCR include, but are not limited to, methods using paired primers, nested primers, single specific primers, degenerate primers, gene-specific primers, vector-specific primers, partially-mismatched primers, and the like. Methods for designing PCR and RT-PCR primers are generally known in the art and are disclosed, for example, in Ausubel F. M. et al. (Eds) *Current Protocols in Molecular* Biology (2007), John Wiley and Sons, Inc; Maniatis et al. *Molecular Cloning* (1982), 280-281; Innis et al. (Eds) (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic Press, New York); Innis and Gelfand, (Eds) (1995) *PCR Strategies* (Academic Press, New York); Innis and Gelfand, (Eds) (1999) *PCR Methods Manual* (Academic Press, New York); and Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2nd ed., Cold Spring Harbor Laboratory Press, Plainview, New York.

The skilled addressee will readily appreciate that various parameters of PCR and RT-PCR procedures may be altered without affecting the ability to achieve the desired product. For example, the salt concentration may be varied or the time and/or temperature of one or more of the denaturation, annealing and extension steps may be varied. Similarly, the amount of DNA, cDNA, or RNA template may also be varied depending on

the amount of nucleic acid available or the optimal amount of template required for efficient amplification. The primers for use in the methods and kits of the present invention are typically oligonucleotides typically being at least about 5 nucleotides to about 80 nucleotides in length, more typically about 10 nucleotides in length to about 50 nucleotides in length, and even more typically about 15 nucleotides in length to about 30 nucleotides in length. The skilled addressee will recognise that the primers described herein may be useful for a number of different applications, including but not limited to PCR, RT-PCR, and use of probes for the detection of *SXT* or *CYR* sequences.

Such primers can be prepared by any suitable method, including, for example, direct chemical synthesis or cloning and restriction of appropriate sequences. Not all bases in the primer need reflect the sequence of the template molecule to which the primer will hybridize, the primer need only contain sufficient complementary bases to enable the primer to hybridize to the template. A primer may also include mismatch bases at one or more positions, being bases that are not complementary to bases in the template, but rather are designed to incorporate changes into the DNA upon base extension or amplification. A primer may include additional bases, for example in the form of a restriction enzyme recognition sequence at the 5' end, to facilitate cloning of the amplified DNA.

The invention provides a method of detecting a cyanotoxic organism based on the detection of one or more of *SXT* gene 6 (*sxtA*), *SXT* gene 9 (*sxtG*), *SXT* gene 10 (*sxtH*), *SXT* gene 11 (*sxtI*) and *SXT* gene 17 (*sxtX*) (SEQ ID NOS: 14, 20, 22, 24, and 36 respectively), or fragments or variants thereof. Additionally or alternatively, a cyanotoxic organism may be detected based on the detection of one or more of the following *SXT* polypeptides: *SXTA* (SEQ ID NO: 15), *SXTG* (SEQ ID NO: 21), *SXTH* (SEQ ID NO: 23), *SXTI* (SEQ ID NO: 25), *SXTX* (SEQ ID NO: 37), or fragments or variants thereof.

The skilled addressee will recognise that any primers capable of the amplifying the stated *SXT* and/or *CYR* sequences, or variants or fragments thereof, are suitable for use in the methods of the invention. For example, suitable oligonucleotide primer pairs for the PCR amplification of *SXT* gene 6 (*sxtA*) may comprise a first primer comprising the sequence of SEQ ID NO: 70 and a second primer comprising the sequence of SEQ ID NO: 71, a first primer comprising the sequence of SEQ ID NO: 72 and a second primer comprising the sequence of SEQ ID NO: 73, a first primer comprising the sequence of SEQ ID NO: 74 and a second primer comprising the sequence of SEQ ID NO: 75, a first primer comprising the sequence of SEQ ID NO: 76 and a second primer comprising the

sequence of SEQ ID NO: 77, a first primer comprising the sequence of SEQ ID NO: 78 and a second primer comprising the sequence of SEQ ID NO: 79, a first primer comprising the sequence of SEQ ID NO: 113 and a second primer comprising the sequence of SEQ ID NO: 114, or a first primer comprising the sequence of SEQ ID NO: 115 or SEQ ID NO:116 and a second primer comprising the sequence of SEQ ID NO: 117.

Suitable oligonucleotide primer pairs for the amplification of *SXT* gene 9 (*sxtG*) may comprise a first primer comprising the sequence of SEQ ID NO: 118 and a second primer comprising the sequence of SEQ ID NO: 119, or a first primer comprising the sequence of SEQ ID NO: 120 and a second primer comprising the sequence of SEQ ID NO: 121.

Suitable oligonucleotide primer pairs for the amplification of *SXT* gene 10 (*sxtH*) may comprise a first primer comprising the sequence of SEQ ID NO: 122 and a second primer comprising the sequence of SEQ ID NO: 123.

Suitable oligonucleotide primer pairs for the amplification of *SXT* gene 11 (*sxtI*) may comprise a first primer comprising the sequence of SEQ ID NO: 124 or SEQ ID NO: 125 and a second primer comprising the sequence of SEQ ID NO: 126, or a first primer comprising the sequence of SEQ ID NO: 127 and a second primer comprising the sequence of SEQ ID NO: 128.

Suitable oligonucleotide primer pairs for the amplification of *SXT* gene 17 (*sxtX*) may comprise a first primer comprising the sequence of SEQ ID NO: 129 and a second primer comprising the sequence of SEQ ID NO: 130, or a first primer comprising the sequence of SEQ ID NO: 131 and a second primer comprising the sequence of SEQ ID NO: 132.

The skilled addressee will recognise that fragments and variants of the above-mentioned primer pairs may also efficiently amplify *SXT* gene 6 (*sxtA*), *SXT* gene 9 (*sxtG*), *SXT* gene 10 (*sxtH*), *SXT* gene 11 (*sxtI*) or *SXT* gene 17 (*sxtX*) sequences.

In certain embodiments of the invention, polynucleotide sequences derived from the *CYR* gene are detected based on the detection of *CYR* gene 8 (*cyrJ*) (SEQ ID NO: 95). Suitable oligonucleotide primer pairs for the PCR amplification of *CYR* gene 8 (*cyrJ*) may comprise a first primer having the sequence of SEQ ID NO: 111 or a fragment or variant thereof and a second primer having the sequence of SEQ ID NO: 112 or a fragment thereof.

Also included within the scope of the present invention are variants and fragments of the exemplified oligonucleotide primers. The skilled addressee will also recognise that the invention is not limited to the use of the specific primers exemplified, and alternative primer sequences may also be used, provided the primers are designed appropriately so as to enable the amplification of *SXT* and/or *CYR* sequences. Suitable primer sequences can be determined by those skilled in the art using routine procedures without undue experimentation. The location of suitable primers for the amplification of *SXT* and/or *CYR* sequences may be determined by such factors as G+C content and the ability for a sequence to form unwanted secondary structures.

Suitable methods of analysis of the amplified nucleic acids are well known to those skilled in the art and are described for example, in, Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2nd ed., Cold Spring Harbor Laboratory Press, Plainview, New York); Ausubel F. M. et al. (Eds) *Current Protocols in Molecular* Biology (2007), John Wiley and Sons, Inc; and Maniatis et al. *Molecular Cloning* (1982), 280-281. Suitable methods of analysis of the amplified nucleic acids include, for example, gel electrophoresis which may or may not be preceded by restriction enzyme digestion, and/or nucleic acid sequencing. Gel electrophoresis may comprise agarose gel electrophoresis or polyacrylamide gel electrophoresis, techniques commonly used by those skilled in the art for separation of DNA fragments on the basis of size. The concentration of agarose or polyacrylamide in the gel in large part determines the resolution ability of the gel and the appropriate concentration of agarose or polyacrylamide will therefore depend on the size of the DNA fragments to be distinguished.

In other embodiments of the invention, *SXT* and *CYR* polynucleotides and variants or fragments thereof may be detected by the use of suitable probes. The probes of the invention are based on the sequences of *SXT* and/or *CYR* polynucleotides disclosed herein. Probes are nucleotide sequences of variable length, for example between about 10 nucleotides and several thousand nucleotides, for use in detection of homologous sequences, typically by hybridization. Hybridization probes of the invention may be genomic DNA fragments, cDNA fragments, RNA fragments, or other oligonucleotides.

Methods for the design and/or production of nucleotide probes are generally known in the art, and are described, for example, in Robinson P. J.. et al. (Eds) *Current Protocols in Cytometry* (2007), John Wiley and Sons, Inc; Ausubel F. M. et al. (Eds) *Current Protocols in Molecular* Biology (2007), John Wiley and Sons, Inc; Sambrook et

al. (1989) *Molecular Cloning: A Laboratory Manual* (2nd ed., Cold Spring Harbor Laboratory Press, Plainview, New York; and Maniatis et al. *Molecular Cloning* (1982), 280-281. Nucleotide probes may be prepared, for example, by chemical synthesis techniques, for example, the phosphodiester and phosphotriester methods (see for example Narang S. A. et al. (1979) Meth. Enzymol. 68:90; Brown, E. L. (1979) et al. Meth. Enzymol. 68:109; and U.S. Patent No. 4356270), the diethylphosphoramidite method (see Beaucage S.L et al. (1981) Tetrahedron Letters, 22:1859-1862). A method for synthesizing oligonucleotides on a modified solid support is described in U.S. Patent No. 4458066.

The probes of the invention may be labelled by incorporation of a marker to facilitate their detection. Techniques for labelling and detecting nucleic acids are described, for example, in Ausubel F. M. et al. (Eds) *Current Protocols in Molecular* Biology (2007), John Wiley and Sons, Inc. Examples of suitable markers include fluorescent molecules (*e.g.* acetylaminofluorene, 5-bromodeoxyuridine, digoxigenin, fluorescein) and radioactive isotopes (*e.g.* 32P, 35S, 3H, 33P). Detection of the marker may be achieved, for example, by chemical, photochemical, immunochemical, biochemical, or spectroscopic techniques.

The methods and kits of the invention also encompass the use of antibodies which are capable of binding specifically to the polypeptides of the invention. The antibodies may be used to qualitatively or quantitatively detect and analyse one or more *SXT* or *CYR* polypeptides in a given sample. Methods for the generation and use of antibodies are generally known in the art and described in, for example, Harlow and Lane (Eds) *Antibodies - A Laboratory Manual*, (1988), Cold Spring Harbor Laboratory, N.Y: Coligan, *Current Protocols in Immunology* (1991); Goding, *Monoclonal Antibodies: Principles and Practice* (1986) 2nd ed; and Kohler & Milstein, (1975) Nature 256: 495-497. The antibodies may be conjugated to a fluorochrome allowing detection, for example, by flow cytometry, immunohistochemisty or other means known in the art. Alternatively, the antibody may be bound to a substrate allowing colorimetric or chemiluminescent detection. The invention also contemplates the use of secondary antibodies capable of binding to one or more antibodies capable of binding specifically to the polypeptides of the invention.

The invention also provides kits for the detection of cyanotoxic organisms and/or cyanobacteria, and/or dinoflagellates. In general, the kits of the invention comprise at least one agent for detecting the presence of one or more *SXT* and/or *CYR* polynucleotide

or polypeptides disclosed herein, or a variant or fragment thereof. Any suitable agent capable of detecting *SXT* and/or *CYR* sequences of the invention may be included in the kit. Non-limiting examples include primers, probes and antibodies.

In one aspect, the invention provides a kit for the detection of cyanobacteria, the kit comprising at least one agent for detecting the presence the presence of one or more *SXT* polynucleotides or polypeptides as disclosed herein, or a variant or fragment thereof.

The *SXT* polynucleotide may comprise a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants and fragments thereof.

Alternatively, the *SXT* polynucleotide may be an RNA or cDNA encoded by a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants and fragments thereof and/or polypeptides as disclosed herein, or a variant or fragment thereof.

The *SXT* polypeptide may comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 67, SEQ ID NO: 69, and variants and fragments thereof.

Also provided is a kit for the detection of cyanotoxic organisms. The kit comprises at least one agent for detecting the presence of one or more *SXT* polynucleotides or polypeptides as disclosed herein, or a variant or fragment thereof.

The *SXT* polynucleotide may comprise a sequence selected from the group consisting of SEQ ID NO: 14, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 36, and variants and fragments thereof.

Alternatively, the *SXT* polynucleotide may be an RNA or cDNA encoded by a sequence selected from the group consisting of SEQ ID NO: 14, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 36, and variants and fragments thereof.

The *SXT* polypeptide may comprising an amino acid sequence selected from the group consisting of consisting of SEQ ID NO: 15, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 37, and variants and fragments thereof.

The at least one agent may be any suitable reagent for the detection of *SXT* polynucleotides and/or polypeptides disclosed herein. For example, the agent may be a primer, an antibody or a probe. By way of exemplification only, the primers or probes may comprise a sequence selected from the group consisting of SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 117, SEQ ID NO: 118, SEQ ID NO: 119, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 122, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 130, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 133, SEQ ID NO: 134, and variants and fragments thereof.

In certain embodiments of the invention, the kits for the detection of cyanobacteria or cyanotoxic organisms may further comprise at least one additional agent capable of detecting one or more CYR polynucleotide and/or *CYR* polypeptide sequences as disclosed herein, or a variant or fragment thereof.

The *CYR* polynucleotide may comprise a polynucleotide comprising a sequence selected from the group consisting of: SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, and variants and fragments thereof.

Alternatively, the *CYR* polynucleotide may comprise a ribonucleic acid or complementary DNA encoded by a sequence selected from the group consisting of: SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, and variants and fragments thereof.

The *CYR* polypeptide may comprise a polypeptide comprising a sequence selected from the group consisting of: SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92 , SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, and SEQ ID NO: 110, and variants and fragments thereof.

The at least one additional agent may be selected, for example, from the group consisting of primers, antibodies and probes. A suitable primer or probe may comprise a sequence selected from the group consisting of SEQ ID NO: 111, SEQ ID NO: 112, and variants and fragments thereof.

In another aspect, the invention provides a kit for the detection of cyanobacteria, the kit comprising at least one agent for detecting the presence the presence of one or more *CYR* polynucleotides or polypeptides as disclosed herein, or a variant or fragment thereof.

The *CYR* polynucleotide may comprise a sequence selected from the group consisting of SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO:107, SEQ ID NO: 109, and variants and fragments thereof.

Alternatively, the *CYR* polynucleotide may be an RNA or cDNA encoded by a sequence selected from the group consisting of SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO:107, SEQ ID NO: 109, and variants and fragments thereof.

The *CYR* polypeptide may comprise a sequence selected from the group consisting of SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92 , SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, and SEQ ID NO: 110, and variants or fragments thereof.

In certain embodiments of the invention, the kits for detecting cyanobacteria comprising one or more agents for the detection of *CYR* sequences or variants or fragments thereof, may further comprise at least one additional agent capable of detecting one or more of the *SXT* polynucleotides and/or *SXT* polypeptides disclosed herein, or variants or fragments thereof.

The *SXT* polynucleotide may comprise a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants and fragments thereof.

Alternatively, the *SXT* polynucleotide may be an RNA or cDNA encoded by a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants and fragments thereof and/or polypeptides as disclosed herein, or a variant or fragment thereof.

The at least one agent may be any suitable reagent for the detection of *CYR* polynucleotides and/or polypeptides disclosed herein. For example, the agent may be a primer, an antibody or a probe. By way of exemplification only, the primers or probes may comprise a sequence selected from the group consisting of SEQ ID NO: 111, SEQ ID NO: 112, and variants and fragments thereof.

Also provided is a kit for the detection of dinoflagellates, the kit comprising at least one agent for detecting the presence one or more *SXT* polynucleotides or polypeptides as disclosed herein, or a variant or fragment thereof.

The *SXT* polynucleotide may comprise a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID

NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants and fragments thereof.

Alternatively, the *SXT* polynucleotide may be an RNA or cDNA encoded by a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants and fragments thereof and/or polypeptides as disclosed herein, or a variant or fragment thereof.

The *SXT* polypeptide may comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 67, SEQ ID NO: 69, and variants and fragments thereof.

In general, the kits of the invention may comprise any number of additional components. By way of non-limiting examples the additional components may include, reagents for cell culture, reference samples, buffers, labels, and written instructions for performing the detection assay.

Methods of screening

The polypeptides and polynucleotides of the present invention, and fragments and analogues thereof are useful for the screening and identification of compounds and agents that interact with these molecules. In particular, desirable compounds are those that modulate the activity of these polypeptides and polynucleotides. Such compounds may exert a modulatory effect by activating, stimulating, increasing, inhibiting or preventing expression or activity of the polypeptides and/or polynucleotides. Suitable compounds may exert their effect by virtue of either a direct (for example binding) or indirect interaction.

Compounds which bind, or otherwise interact with the polypeptides and polynucleotides of the invention, and specifically compounds which modulate their activity, may be identified by a variety of suitable methods. Non limiting methods include the two-hybrid method, co-immunoprecipitation, affinity purification, mass spectroscopy, tandem affinity purification, phage display, label transfer, DNA microarrays/gene coexpression and protein microarrays.

For example, a two-hybrid assay may be used to determine whether a candidate agent or plurality of candidate agents interacts or binds with a polypeptide of the invention or a variant or fragment thereof. The yeast two-hybrid assay system is a yeast-based genetic assay typically used for detecting protein-protein interactions (Fields and Song., *Nature* 340: 245-246 (1989)). The assay makes use of the multi-domain nature of transcriptional activators. For example, the DNA-binding domain of a known transcriptional activator may be fused to a polypeptide of the invention or a variant or fragment thereof, and the activation domain of the transcriptional activator fused to the candidate agent. Interaction between the candidate agent and the polypeptide of the invention or a variant or fragment thereof, will bring the DNA-binding and activation domains of the transcriptional activator into close proximity. Subsequent transcription of a specific reporter gene activated by the transcriptional activator allows the detection of an interaction.

In a modification of the technique above, a fusion protein may be constructed by fusing the polypeptide of the invention or a variant or fragment thereof to a detectable tag, for example alkaline phosphatase, and using a modified form of immunoprecipitation as described by Flanagan and Leder (Flanagan and Leder, *Cell* 63:185–194 (1990))

Alternatively, co-immunoprecipation may be used to to determine whether a candidate agent or plurality of candidate agents interacts or binds with polypeptide of the invention or a variant or fragment thereof. Using this technique, cyanotoxic organisms,

cyanobacteria and/or dinoflagellates may be lysed under nondenaturing conditions suitable for the preservation of protein-protein interactions. The resulting solution can then be incubated with an antibody specific for a polypeptide of the invention or a variant or fragment thereof and immunoprecipitated from the bulk solution, for example by capture with an antibody-binding protein attached to a solid support. Immunoprecipitation of the polypeptide of the invention or a variant or fragment thereof by this method facilitates the co-immunoprecipation of an agent associated with that protein. The identification an associated agent can be established using a number of methods known in the art, including but not limited to SDS-PAGE, western blotting, and mass spectrometry.

Alternatively, the phage display method may be used to to determine whether a candidate agent or plurality of candidate agents interacts or binds with a polypeptide of the invention or a variant or fragment thereof. Phage display is a test to screen for protein interactions by integrating multiple genes from a gene bank into phage. Under this method, recombinant DNA techniques are used to express numerous genes as fusions with the coat protein of a bacteriophage such the peptide or protein product of each gene is displayed on the surface of the viral particle. A whole library of phage-displayed peptides or protein products of interest can be produced in this way. The resulting libraries of phage-displayed peptides or protein products may then be screened for the ability to bind a polypeptide of the invention or a variant or fragment thereof. DNA extracted from interacting phage contains the sequences of interacting proteins.

Alternatively, affinity chromatography may be used to to determine whether a candidate agent or plurality of candidate agents interacts or binds with a polypeptide of the invention or a variant or fragment thereof. For example, a polypeptide of the invention or a variant or fragment thereof, may be immobilised on a support (such as sepharose) and cell lysates passed over the column. Proteins binding to the immobilised polypeptide of the invention or a variant or fragment thereof may then be eluted from the column and identified, for example by N-terminal amino acid sequencing.

Potential modulators of the activity of the polypeptides of the invention may be generated for screening by the above methods by a number of techniques known to those skilled in the art. For example, methods such as X-ray crystallography and nuclear magnetic resonance spectroscopy may be used to model the structure of polypeptide of the invention or a variant or fragment thereof, thus facilitating the design of potential modulating agents using computer-based modeling. Various forms of combinatorial chemistry may also be used to generate putative modulators.

Polypeptides of the invention and appropriate variants or fragments thereof can be used in high-throughput screens to assay candidate compounds for the ability to bind to, or otherwise interact therewith. These candidate compounds can be further screened against functional polypeptides to determine the effect of the compound on polypeptide activity.

The present invention also contemplates compounds which may exert their modulatory effect on polypeptides of the invention by altering expression of the polypeptide. In this case, such compounds may be identified by comparing the level of expression of the polypeptide in the presence of a candidate compound with the level of expression in the absence of the candidate compound.

It will be appreciated that the methods described above are merely examples of the types of methods that may be utilised to identify agents that are capable of interacting with, or modulating the activity of polypeptides of the invention or variants or fragments thereof. Other suitable methods will be known by persons skilled in the art and are within the scope of this invention.

Using the methods described above, an agent may be identified that is an agonist of a polypeptide of the invention or a variant or fragment thereof. Agents which are agonists enhance one or more of the biological activities of the polypeptide. Alternatively, the methods described above may identify an agent that is an antagonist of a polypeptide of the invention or a variant or fragment thereof. Agents which are antagonists retard one or more of the biological activities of the polypeptide.

Antibodies may act as agonists or antagonists of a polypeptide of the invention or a variant or fragment thereof. Preferably suitable antibodies are prepared from discrete regions or fragments of the polypeptides of the invention or variants or fragments thereof. An antigenic portion of a polynucleotide of interest may be of any appropriate length, such as from about 5 to about 15 amino acids. Preferably, an antigenic portion contains at least about 5, 6, 7, 8, 9, 10, 11, 12, 13 or 14 amino acid residues.

Methods for the generation of suitable antibodies will be readily appreciated by those skilled in the art. For example, monoclonal antibody specific for a polypeptide of the invention or a variant or fragment thereof typically containing Fab portions, may be prepared using hybridoma technology described in *Antibodies-A Laboratory Manual*, Harlow and Lane, eds., Cold Spring Harbor Laboratory, N.Y. (1988).

In essence, in the preparation of monoclonal antibodies directed toward polypeptide of the invention or a variant or fragment thereof, any technique that provides for the

production of antibody molecules by continuous cell lines in culture may be used. These include the hybridoma technique originally developed by Kohler *et al.*, *Nature,* 256:495-497 (1975), as well as the trioma technique, the human B-cell hybridoma technique (Kozbor *et al., Immunology Today*, 4:72 (1983)), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole *et al.*, in Monoclonal Antibodies and Cancer Therapy, pp. 77- 96, Alan R. Liss, Inc., (1985)). Immortal, antibody-producing cell lines can be created by techniques other than fusion, such as direct transformation of B lymphocytes with oncogenic DNA, or transfection with Epstein-Barr virus. See, for example, M. Schreier *et al., "Hybridoma Techniques"* Cold Spring Harbor Laboratory, (1980); Hammerling *et al., "Monoclonal Antibodies and T-cell Hybridomas"* Elsevier/North-Holland Biochemical Press, Amsterdam (1981); and Kennett *et al., "Monoclonal Antibodies"*, Plenum Press (1980).

In brief, a means of producing a hybridoma from which the monoclonal antibody is produced, a myeloma or other self-perpetuating cell line is fused with lymphocytes obtained from the spleen of a mammal hyperimmunised with a recognition factor-binding portion thereof, or recognition factor, or an origin-specific DNA-binding portion thereof. Hybridomas producing a monoclonal antibody useful in practicing this invention are identified by their ability to immunoreact with the present recognition factors and their ability to inhibit specified transcriptional activity in target cells.

A monoclonal antibody useful in practicing the invention can be produced by initiating a monoclonal hybridoma culture comprising a nutrient medium containing a hybridoma that secretes antibody molecules of the appropriate antigen specificity. The culture is maintained under conditions and for a time period sufficient for the hybridoma to secrete the antibody molecules into the medium. The antibody-containing medium is then collected. The antibody molecules can then be further isolated by well-known techniques.

Similarly, there are various procedures known in the art which may be used for the production of polyclonal antibodies. For the production of polyclonal antibodies against a polypeptide of the invention or a variant or fragment thereof, various host animals can be immunized by injection with a polypeptide of the invention, or a variant or fragment thereof, including but not limited to rabbits, chickens, mice, rats, sheep, goats, etc. Further, the polypeptide variant or fragment thereof can be conjugated to an immunogenic carrier (e.g., bovine serum albumin (BSA) or keyhole limpet hemocyanin (KLH)). Also, various adjuvants may be used to increase the immunological response, including but not

limited to Freund's (complete and incomplete), mineral gels such as aluminium hydroxide, surface active substances such as rysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and Corynebacterium parvum.

Screening for the desired antibody can also be accomplished by a variety of techniques known in the art. Assays for immunospecific binding of antibodies may include, but are not limited to, radioimmunoassays, ELISAs (enzyme-linked immunosorbent assay), sandwich immunoassays, immunoradiometric assays, gel diffusion precipitation reactions, immunodiffusion assays, in situ immunoassays, Western blots, precipitation reactions, agglutination assays, complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, and the like (see, for example, Ausubel *et al.*, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York (1994)). Antibody binding may be detected by virtue of a detectable label on the primary antibody. Alternatively, the antibody may be detected by virtue of its binding with a secondary antibody or reagent which is appropriately labelled. A variety of methods are known in the art for detecting binding in an immunoassay and are included in the scope of the present invention.

The antibody (or fragment thereof) raised against a polypeptide of the invention or a variant or fragment thereof, has binding affinity for that protein. Preferably, the antibody (or fragment thereof) has binding affinity or avidity greater than about $10^5 M^{-1}$, more preferably greater than about $10^6 M^{-1}$, more preferably still greater than about $10^7 M^{-1}$ and most preferably greater than about $10^8 M^{-1}$.

In terms of obtaining a suitable amount of an antibody according to the present invention, one may manufacture the antibody(s) using batch fermentation with serum free medium. After fermentation the antibody may be purified via a multistep procedure incorporating chromatography and viral inactivation/removal steps. For instance, the antibody may be first separated by Protein A affinity chromatography and then treated with solvent/detergent to inactivate any lipid enveloped viruses. Further purification, typically by anion and cation exchange chromatography may be used to remove residual proteins, solvents/detergents and nucleic acids. The purified antibody may be further purified and formulated into 0.9% saline using gel filtration columns. The formulated bulk preparation may then be sterilised and viral filtered and dispensed.

Embodiments of the invention may utilise antisense technology to inhibit the expression of a nucleic acid of the invention or a fragment or variant thereof by blocking translation of the encoded polypeptide. Antisense technology takes advantage of the fact that nucleic acids pair with complementary sequences. Suitable antisense molecules can be manufactured by chemical synthesis or, in the case of antisense RNA, by transcription *in vitro* or *in vivo* when linked to a promoter, by methods known to those skilled in the art.

For example, antisense oligonucleotides, typically of 18-30 nucleotides in length, may be generated which are at least substantially complementary across their length to a region of the nucleotide sequence of the polynucleotide of interest. Binding of the antisense oligonucleotide to their complementary cellular nucleotide sequences may interfere with transcription, RNA processing, transport, translation and/or mRNA stability. Suitable antisense oligonucleotides may be prepared by methods well known to those of skill in the art and may be designed to target and bind to regulatory regions of the nucleotide sequence or to coding (gene) or non-coding (intergenic region) sequences. Typically antisense oligonucleotides will be synthesized on automated synthesizers. Suitable antisense oligonucleotides may include modifications designed to improve their delivery into cells, their stability once inside a cell, and/or their binding to the appropriate target. For example, the antisense oligonucleotide may be modified by the addition of one or more phosphorothioate linkages, or the inclusion of one or morpholine rings into the backbone (so-called 'morpholino' oligonucleotides).

An alternative antisense technology, known as RNA interference (RNAi), may be used, according to known methods in the art (see for example WO 99/49029 and WO 01/70949), to inhibit the expression of a polynucleotide. RNAi refers to a means of selective post-transcriptional gene silencing by destruction of specific mRNA by small interfering RNA molecules (siRNA). The siRNA is generated by cleavage of double stranded RNA, where one strand is identical to the message to be inactivated. Double-stranded RNA molecules may be synthesised in which one strand is identical to a specific region of the p53 mRNA transcript and introduced directly. Alternatively corresponding dsDNA can be employed, which, once presented intracellularly is converted into dsRNA. Methods for the synthesis of suitable molecule for use in RNAi and for achieving post-transcriptional gene silencing are known to those of skill in the art.

A further means of inhibiting expression may be achieved by introducing catalytic antisense nucleic acid constructs, such as ribozymes, which are capable of cleaving

mRNA transcripts and thereby preventing the production of wild type protein. Ribozymes are targeted to and anneal with a particular sequence by virtue of two regions of sequence complementarity to the target flanking the ribozyme catalytic site. After binding the ribozyme cleaves the target in a site-specific manner. The design and testing of ribozymes which specifically recognise and cleave sequences of interest can be achieved by techniques well known to those in the art (see for example Lieber and Strauss, 1995, *Molecular and Cellular Biology*, 15:540-551.

The invention will now be described with reference to specific examples, which should not be construed as in any way limiting the scope of the invention.

Examples

The invention will now be described with reference to specific examples, which should not be construed as in any way limiting the scope of the invention.

Example 1: Cyanobacterial cultures and characterisation of the *SXT* gene cluster.

Cyanobacterial strains used in the present study (Figure 1) were grown in Jaworski medium in static batch culture at 26°C under continuous illumination (10 μmol m^{-2} s^{-1}). Total genomic DNA was extracted from cyanobacterial cells by lysozyme/SDS/proteinase K lysis following phenol-chloroform extraction as described in Neilan, B. A. 1995.. Appl Environ Microbiol 61:2286-2291. DNA in the supernatant was precipitated with 2 volumes - 20°C ethanol, washed with 70% ethanol, dissolved in TE-buffer (10:1), and stored at - 20°C. PCR primer sequences used for the amplification of *sxt* ORFs are shown in Figure 1B).

PCR amplicons were separated by agarose gel electrophoresis in TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 7.8), and visualised by UV translumination after staining in ethidium bromide (0.5 μg/ml). Sequencing of unknown regions of DNA was performed by adaptor-mediated PCR as described in Moffitt et al. (2004) Appl. Environ. Microbiol. 70:6353-6362. Automated DNA sequencing was performed using the PRISM Big Dye cycle sequencing system and a model 373 sequencer (Applied Biosystems). Sequence data were analysed using ABI Prism-Autoassembler software, and percentage similarity and identity to other translated sequences determined using BLAST in

conjunction with the National Center for Biotechnology Information (NIH), Fugue blast (http://www-cryst.bioc.cam.ac.uk/fugue/) was used to identify distant homologs via sequence-structure comparisons. The *sxt* gene clusters were assembled using the software Phred, Phrap, and Consed (http://www.phrap.org/phredphrapconsed.html), and open reading frames manually identified. GenBank accession numbers for the *sxt* gene cluster from *C. raciborskii* T3 is DQ787200.

Example 2: Mass spectrometric analysis of *SXT* intermediates.

Bacterial extracts and *SXT* standards were analysed by HPLC (Thermo Finnigan Surveyor HPLC and autosampler) coupled to an ion trap mass spectrometer (Thermo Finnigan LCQ Deca XP Plus) fitted with an electrospray source. Separation of analytes was obtained on a 2.1 mm x 150 mm Phenomenex Luna 3 micron C18 column at 100 mL/min. Analysis was performed using a gradient starting at 5% acetonitrile in10 mM heptafluorobutyric acid (HFBA) This was maintained for 10 min, then ramped to 100% acetonitrile, over 30 min. Conditions were held at 100% acetonitrile for 10 min to wash the column and then returned to 5% acetonitrile in10 mM HFBA and again held for 10 min to equilibrate the column for the next sample. This resulted in a runtime of 60 min per sample. Sample volumes of 10-100 mL were injected for each analysis. The HPLC eluate directly entered the electrospray source, which was programmed as follows: electrospray voltage 5 kV, sheath gas flow rate 30 arbitrary units, auxiliary gas flow rate 5 arbitrary units. The capillary temperature was 200°C and had a voltage of 47 V. Ion optics were optimised for maximum sensitivity before sample analysis using the instruments autotune function with a standard toxin solution. Mass spectra were acquired in the centroid mode over the m/z range 145-650. Mass range setting was 'normal', with 200 ms maximum ion injection time and automatic gain control (AGC) on. Tandem mass spectra were obtained over a m/z range relevant to the precursor ion. Collision energy was typically 20-30 ThermoFinnigan arbitrary units, and was optimised for maximal information using standards where available.

Example 3: Identification and sequencing of the *SXT* gene cluster in *Cylindrospermopsis raciborskii* T3

O-carbamoyltransferase was initially detected in *C. raciborskii* T3 via degenerate PCR, and later named *sxtI*. Further investigation showed that homologues of *sxtI* were exclusively present in *SXT* toxin-producing strains of four cyanobacterial genera (Table

1), thus representing a good candidate gene in *SXT* toxin biosynthesis. The sequence of the complete putative *SXT* biosynthetic gene cluster (*sxt*) was then obtained by genome walking up- and downstream of *sxtI* in *C. raciborskii* T3 (Figure 3). In *C. raciborskii* T3, this *sxt* gene cluster spans approximately 35000 bp, encoding 31 open reading frames (Figure 2). The cluster also included other genes encoding *SXT*-biosynthesis enzymes, including a methyltransferase (*sxtA1*), a class II aminotransferase (*sxtA4*), an amidinotransferase (*sxtG*), dioxygenases (*sxtH*), in addition to the Ocarbamoyltransferase (*sxtI*). PCR screening of selected *sxt* open reading frames in toxic and non-toxic cyanobacteria strains showed that they were exclusively present in *SXT* toxin-producing isolates (Figure 1A), indicating the association of these genes with the toxic phenotype. In the following passages we describe the open reading frames in the putative *sxt* gene cluster and their predicted functions, based on bioinformatic analysis, LCMS/ MS data on biosynthetic intermediates and *in vitro* biosynthesis, when applicable.

Example 4: Functional prediction of the parent molecule *SXT* biosynthetic genes

Bioinformatic analysis of the *sxt* gene cluster revealed that it contains a previously undescribed example of a polyketide synthase (PKS) like structure, named *sxtA*. SxtA possesses four catalytic domains, SxtA1 to SxtA4. An iterated PSI-blast search revealed low sequence homology of SxtA1 to S-adenosylmethionine (SAM)-dependent methyltransferases. Further analysis revealed the presence of three conserved sequence motifs in SxtA1 (278-ITDMGCGDG- 286, 359-DPENILHI-366, and 424-VVNKHGLMIL-433) that are specific for SAMdependent methyltransferases. SxtA2 is related to GCN5-related N-acetyl transferases (GNAT). GNAT catalyse the transfer of acetate from acetyl-CoA to various heteroatoms, and have been reported in association with other unconventional PKSs, such as PedI, where they load the acyl carrier protein (ACP) with acetate. SxtA3 is related to an ACP, and provides a phosphopantetheinyl-attachment site. SxtA4 is homologous to class II aminotransferases and was most similar to 8-amino-7-oxononanoate synthase (AONS). Class II aminotransferases are a monophyletic group of pyridoxal phosphate (PLP)-dependent enzymes, and the only enzymes that are known to perform Claisen-condensations of amino acids. We therefore reasoned that *sxtA* performs the first step in SXT biosynthesis, involving a Claisen-condensation.

The predicted reaction sequence of SxtA, based on its primary structure, is the loading of the ACP (SxtA3) with acetate from acetyl-CoA, followed by the SxtA1-

catalysed methylation of acetyl-ACP, converting it to propionyl-ACP. The class II aminotransferase domain, SxtA4, would then perform a Claisen-condensation between propionyl-ACP and arginine (Figure 4). The putative product of SxtA is thus 4-amino-3-oxoguanidinoheptane which is here designated as Compound A', (Figure 4). To verify this pathway for SXT biosynthesis based on comparative gene sequence analysis, cell extracts of *C. raciborskii* T3 were screened by LC-MS/MS for the presence of compound A' (Figure 5) as well as arginine and SXT as controls. Arginine and SXT were readily detected (Figure 5) and produced the expected fragment ions. On the other hand, LC-MS/MS data obtained from *m/z* 187 was consistent with the presence of structure A from *C. raciborskii* T3 (Figure 5). MS/MS spectra showed the expected fragment ion (m/z 170, m/z 128) after the loss of ammonia and guanidine from A'. LC-MS/MS data strongly supported the predicted function of SxtA and thus a revised initiating reaction in the SXT biosynthesis pathway.

sxtG encodes a putative amidinotransferase, which had the highest amino acid sequence similarity to L-arginine:lysine amidinotransferases. It is proposed that the product of SxtA is the substrate for the amidinotransferase SxtG, which transfers an amidino group from arginine to the a-amino group A' (Figure 4), thus producing 4,7-diguanidino-3-oxoheptane designated compound B' (Figure 3). This hypothetical sequence of reactions was also supported by the detection of C' by LC-MS/MS (Figure 4). Cell extracts from *C. raciborskii* T3, however, did not contain any measurable levels of B' (4,7-diguanidino-3- oxoheptane). A likely explanation for the failure to detect the intermediate B' is its rapid cyclisation to form C' via the action of SxtB.

The *sxt* gene cluster encodes an enzyme, sxtB, similar to the cytidine deaminase-like enzymes from g-proteobacteria. The catalytic mechanism of cytidine deaminase is a retro-aldol cleavage of ammonia from cytidine, which is the same reaction mechanism in the reverse direction as the formation of the first heterocycle in the conversion from B' to C' (Figure 4). It is therefore suggested that SxtB catalyses this retroaldol- like condensation (step 4, Figure 4).

The incorporation of methionine methyl into SXT, and its hydroxylation was studied. Only one methionine methyl-derived hydrogen is retained in SXT, and a 1,2-H shift has been observed between acetate-derived C-5 and C-6 of SXT. Hydroxylation of the methyl side-chain of the SXT precursor proceeds via epoxidation of a double-bond between the SAM-derived methyl group and the acetate derived C-6. This incorporation pattern may result from an electrophilic attack of methionine methyl on the double bond

between C-5 and C-6, which would have formed during the preceding cyclisation. Subsequently, the new methylene side-chain would be epoxidated, followed by opening to an aldehyde, and subsequent reduction to a hydroxyl. Retention of only one methionine methyl-derived hydrogen, the 1,2-H shift between C-5 and C-6, and the lacking 1,2-H shift between C-1 and C-5 is entirely consistent with the results of this study, whereby the introduction of methionine methyl precedes the formation of the three heterocycles.

sxtD encodes an enzyme with sequence similarity to sterol desaturase and is the only candidate desaturase present in the *sxt* gene cluster, SxtD is predicted to introduce a double bond between C-1 and C-5 of C', and cause a 1,2-H shift between C-5 and C-6 (compound D', Figure 3). The gene product of *sxtS* has sequence homology to non-heme iron 2-oxoglutaratedependent (2OG) dioxygenases. These are multifunctional enzymes that can perform hydroxylation, epoxidation, desaturation, cyclisation, and expansion reactions. 2OG dioxygenases have been reported to catalyse the oxidative formation of heterocycles. SxtS could therefore perform the consecutive epoxidation of the new double bond, and opening of the epoxide to an aldehyde with concomitant bicyclisation. This explains the retention of only one methionine methyl-derived hydrogen, and the lack of a 1,2-H shift between C-1 and C-5 of SXT (steps 5 to 7, Figure 4). SxtU has sequence similarity to short-chain alcohol dehydrogenases. The most similar enzyme with a known function is clavaldehyde dehydrogenase (AAF86624), which reduces the terminal aldehyde of clavulanate-9-aldehyde to an alcohol. SxtU is therefore predicted to reduce the terminal aldehyde group of the SXT precursor in step 8 (Figure 4), forming compound E'.

The concerted action of SxtD, SxtS and SxtU is therefore the hydroxylation and bicyclisation of compound C' to E' (Figure 4). In support for this proposed pathway of SXT biosynthesis, LC-MS/MS obtained from *m/z* 211 and *m/z* 225 allowed the detection of compounds C' and E' from *C. raciborskii* T3 (Figure 5). On the other hand, no evidence could be found by LC-MS/MS for intermediates B (*m/z* 216), and C (*m/z* 198). MS/MS spectra showed the expected fragment ions after the loss of ammonia and guanidine from C', as well as the loss of water in the case of E'.

The detection of E' indicated that the final reactions leading to the complete SXT molecule are the O-carbamoylation of its free hydroxyl group and a oxidation of C-12. The actual sequence of these final reactions, however, remains uncertain. The gene product of *sxtI* is most similar to a predicted Ocarbamoyltransferase from *Trichodesmium erythraeum* (accession ABG50968) and other predicted O-carbamoyltransferases from

cyanobacteria. O-carbamoyltransferases invariably transfer a carbamoyl group from carbamoylphosphate to a free hydroxyl group. Our data indicate that SxtI may catalyse the transfer of a carbamoyl group from carbamoylphosphate to the free hydroxy group of E'. Homologues of *sxtJ* and *sxtK* with a known function were not found in the databases, however it was noted that *sxtJ* and *sxtK* homologues were often encoded adjacent to O-carbamoyltransferase genes.

The *sxt* gene cluster contains two genes, *sxtH* and *sxtT*, each encoding a terminal oxygenase subunit of bacterial phenyl-propionate and related ring-hydroxylating dioxygenases. The closest homologue with a predicted function was capreomycidine hydroxylase from *Streptomyces vinaceus*, which hydroxylates a ringcarbon (C-6) of capreomycidine. SxtH and SxtT may therefore perform a similar function in SXT biosynthesis, that is, the oxidation or hydroxylation and oxidation of C-12, converting F' into *SXT*.

Members belonging to bacterial phenylpropionate and related ring-hydroxylating dioxygenases are multi-component enzymes, as they require an oxygenase reductase for their regeneration after each catalytic cycle. The *sxt* gene cluster provides a putative electron transport system, which would fulfill this function. *sxtV* encodes a 4Fe-4S ferredoxin with high sequence homology to a ferredoxin from *Nostoc punctiforme*. *sxtW* was most similar to fumarate reductase/succinate dehydrogenase-like enzymes from *A. variabilis* and *Nostoc punctiforme,* followed by AsfA from *Pseudomonas putida*. AsfA and AsfB are enzymes involved in the transport of electrons resulting from the catabolism of aryl sulfonates. SxtV could putatively extract an electron pair from succinate, converting it to fumarate, and then transfer the electrons via ferredoxin (SxtW) to SxtH and SxtT.

Example 5: Comparative sequence analysis and functional assignment of *SXT* tailoring genes

Following synthesis of the parent molecule SXT, modifying enzymes introduce various functional groups. In addition to SXT, *C. raciborskii* T3 produces N-1 hydroxylated (neoSXT), decarbamoylated (dcSXT), and N-sulfurylated (GTX-5) toxins, whereas *A. circinalis* AWQC131C produces decarbamoylated (dcSXT), O-sulfurylated (GTX-3/2, dcGTX-3/2), as well as both O-and N-sulfurylated toxins (C- 1/2), but no N-1 hydroxylated toxins.

sxtX encodes an enzyme with homology to cephalosporin hydroxylase. *sxtX* was only detected in *C. raciborskii* T3, *A. flos-aquae* NH- 5, and *Lyngbya wollei*, which produce N-1 hydroxylated analogues of *SXT*, such as neoSXT. This component of the gene cluster was not present in any strain of *A. circinalis*, and therefore probably the reason why this species does not produce N-1 hydroxylated *PSP* toxins (Figure 1A). The predicted function of SxtX is therefore the N-1 hydroxylation of *SXT*.

A. circinalis AWQC131C and *C. raciborskii* T3 also produces N- and O-sulfated analogues of SXT (GTX-5, C-2/3, (dc)GTX- 3/4). The activity of two 3'-phosphate 5'-phosphosulfate (PAPS)-dependent sulfotransferases, which were specific for the N- 21 of SXT and GTX-3/2, and O-22 of 11- hydroxy *SXT*, respectively, has been described from the *SXT* toxin-producing dinoflagellate *Gymnodinium catenatum*. The *sxt* gene cluster from *C. raciborskii* T3 encodes a putative sulfotransferase, SxtN. A PSI-BLAST search with SxtN identified only 25 hypothetical proteins of unknown function with an E value above the threshold (0.005). A profile library search, however, revealed significant structural relatedness of SxtN to estrogen sulfotransferase (1AQU) (Z-score=24.02) and other sulfotransferases. SxtN has a conserved N-terminal region, which corresponds to the adenosine 3'-phosphate 5'-phosphosulfate (PAPS) binding region in 1AQU. It is not known, however, whether SxtN transfers a sulfate group to N-21 or O-22. Interestingly, the *sxt* gene cluster encodes an adenylylsulfate kinase (APSK), SxtO, homologues of which are involved in the formation of PAPS (Figure 2). APKS phosphorylates the product of ATPsulfurylase, adenylylsulfate, converting it to PAPS. Other biosynthetic gene clusters that result in sulfated secondary metabolites also contain genes required for the production of PAPS.

Decarbamoylated analogues of *SXT* could be produced via either of two hypothetical scenarios. Enzymes that act downstream of the carbamoyltransferase, SxtI, in the biosynthesis of PSP toxins are proposed to have broad substrate specificity, processing both carbamoylated and decarbamoylated precursors of *SXT*. Alternatively, hydrolytic cleavage of the carbamoyl moiety from *SXT* or its precursors may occur. SxtL is related to GDSL-lipases, which are multifunctional enzymes with thioesterase, arylesterase, protease and lysophospholipase activities. The function of SxtL could therefore include the hydrolytic cleavage of the carbamoyl group from *SXT* analogues.

Example 6: Cluster-associated *SXT* genes involved in metabolite transport

sxtF and *sxtM* encoded two proteins with high sequence similarity to sodium-driven multidrug and toxic compound extrusion (MATE) proteins of the NorM family. Members of the NorM family of MATE proteins are bacterial sodium-driven antiporters, that export cationic substances. All of the PSP toxins are cationic substances, except for the C-toxins which are zwitterionic. It is therefore probable that SxtF and SxtM are also involved in the export of PSP toxins. A mutational study of NorM from *V. parahaematolyticus* identified three conserved negatively charged residues (D32, E251, and D367) that confer substrate specificity, however the mechanism of substrate recognition remains unknown. In SxtF, the residue corresponding to E251 of NorM is conserved, whereas those corresponding to D32 and D367 are replaced by the neutral amino acids asparagine and tyrosine, respectively. Residues corresponding to D32 and E251 are conserved in SxtM, but D367 is replaced by histidine. The changes in substrate-binding residues may reflect the differences in PSP toxin substrates transported by these proteins.

Example 7: Putative transcriptional regulators of saxitoxin synthase

Environmental factors, such as nitrogen and phosphate availability have been reported to regulate the production of PSP toxins in dinoflagellates and cyanobacteria. Two transcriptional factors, *sxtY* and *sxtZ*, related to PhoU and OmpR, respectively, as well as a two component regulator histidine kinase were identified proximal to the 3'-end of the *sxt* gene cluster in *C. raciborskii* T3. PhoU-related proteins are negative regulators of phosphate uptake whereas OmpR-like proteins are involved in the regulation of a variety of metabolisms, including nitrogen and osmotic balance. It is therefore likely that PSP toxin production in *C. raciborskii* T3 is regulated at the transcriptional level in response to the availability of phosphate, as well as, other environmental factors.

Example 8: Phylogenetic origins of the *SXT* genes

The *sxt* gene cluster from *C. raciborskii* T3 has a true mosaic structure. Approximately half of the *sxt* genes of *C. raciborskii* T3 were most similar to counterparts from other cyanobacteria, however the remaining genes had their closest matches with homologues from proteobacteria, actinomycetes, sphingobacteria, and firmicutes. There is an increasing body of evidence that horizontal gene transfer (HGT) is a major driving force behind the evolution of prokaryotic genomes, and cyanobacterial genomes are known to be greatly affected by HGT, often involving transposases and

phages. The fact that the majority of *sxt* genes are most closely related to homologues from other cyanobacteria, suggests that *SXT* biosynthesis may have evolved in an ancestral cyanobacterium that successively acquired the remaining genes from other bacteria via HGT. The structural organisation of the investigated *sxt* gene cluster, as well as the presence of several transposases related to the IS4-family, suggests that small cassettes of *sxt* genes are mobile.

Example 9: Cyanobacterial cultures and characterisation of the *CYR* gene cluster.

Cyanobacterial strains were grown in Jaworski medium as described in Example 1 above. Total genomic DNA was extracted from cyanobacterial cells by lysozyme/SDS/proteinase K lysis following phenol-chloroform extraction as described previously Neilan, B. A. 1995.. Appl Environ Microbiol 61:2286-2291. DNA in the supernatant was precipitated with 2 volumes -20°C ethanol, washed with 70% ethanol, dissolved in TE-buffer (10:1), and stored at -20°C.

Characterization of unknown regions of DNA flanking the putative cylindrospermopsin biosynthesis genes was performed using an adaptor-mediated PCR as described in Moffitt et al. (2004) Appl. Environ. Microbiol. 70:6353-6362. PCRs were performed in 20 µl reaction volumes containing 1 x *Taq* polymerase buffer 2.5 mM $MgCl_2$, 0.2 mM deoxynucleotide triphosphates, 10 pmol each of the forward and reverse primers, between 10 and 100 ng genomic DNA and 0.2 U of *Taq* polymerase (Fischer Biotech, Australia). Thermal cycling was performed in a GeneAmp PCR System 2400 Thermal cycler (Perkin Elmer Corporation, Norwalk, CT). Cycling began with a denaturing step at 94°C for 3 min followed by 30 cycles of denaturation at 94°C for 10 s, primer annealing between 55° and 65°C for 20 s and a DNA strand extension at 72°C for 1-3 min. Amplification was completed by a final extension step at 72°C for 7 min. Amplified DNA was separated by agarose gel electrophoresis in TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 7.8), and visualized by UV transillumination after staining with ethidium bromide (0.5 µg/ml).

Automated DNA sequencing was performed using the PRISM Big Dye cycle sequencing system and a model 373 sequencer (Applied Biosystems, Foster City, CA). Sequence data were analyzed using ABI Prism-Autoassembler software, while identity/similarity values to other translated sequences were determined using BLAST in conjunction with the National Center for Biotechnology Information (NIH, Bethesda,

MD). Fugue blast (http://www-cryst.bioc.cam.ac.uk/fugue/) was used to identify distant homologs via sequence-structure comparisons. The gene clusters were assembled using the software Phred, Phrap, and Consed (http://www.phrap.org/phredphrapconsed.html), open reading frames were manually identified. Polyketide synthase and non-ribosomal peptide synthetase domains were determined using the specialized databases based on crystal structures (http://www-ab.informatik.uni-tuebingen.de/software/NRPSpredictor; http://www.tigr.org/jravel/nrps/, http://www.nii.res.in/nrps-pks.html).

Example 10: Genetic screening of Cylindrospermopsin-producing and non-producing cyanobacterial strains

Cylindrospermopsin-producing and non-producing cyanobacterial strains were screened for the presence of the sulfotransferase gene cyrJ using the primer set cynsulfF (5' ACTTCTCTCCTTTCCCTATC 3') (SEQ ID NO: 111) and cylnamR (5' GAGTGAAAATGCGTAGAACTTG 3') (SEQ ID NO: 112). Genomic DNA was tested for positive amplification using the 16S rRNA gene primers 27F and 809 as described in Neilan et al. (1997) Int. J. Syst. Bacteriol. 47:693-697. Amplicons were sequenced, as described in Example 9 above, to verify the identity of the gene fragment.

The biosynthesis of cylindrospermopsin involves an amidinotransferase, a NRPS, and a PKS (AoaA, AoaB and AoaC, respectively). In order to obtain the entire sequence of the cylindrospermopsin biosynthesis gene cluster, we used adaptor-mediated 'gene-walking' technology, initiating the process from a partial sequence of the amidinotransferase gene from *C. raciborskii* AWT205. Successive outward facing primers were designed and the entire gene cluster spanning 43 kb was sequenced, together with a further 3.5 kb on either side of the toxin gene cluster.

These flanking regions encode putative accessory genes (*hyp* genes), which include molecular chaperons involved in the maturation of hydrogenases. Due to the fact that these genes are flanking the cylindrospermopsin gene cluster at both ends, we postulate that the toxin gene cluster was inserted into this area of the genome thus interrupting the HYP gene cluster. This genetic rearrangement is mechanistically supported by the presence of transposase-like sequences within the cylindrospermopsin cluster.

Bioinformatic analysis of the toxin gene cluster was performed and based on gene function inference using sequence alignments (NCBI BLAST), predicted structural homologies (Fugue Blast), and analysis of PKS and NRPS domains using specialized blast servers based on crystal structures. The cylindrospermopsin biosynthesis cluster

contains 15 ORFs, which encode all the functions required for the biosynthesis, regulation and export of the toxin cylindrospermopsin (Figure 6).

Example 11: Formation of the *CYR* carbon skeleton

The first step in formation of the carbon skeleton of cylindrospermopsin involves the synthesis of guanidinoacetate via transamidination of glycine. CyrA, the AoaA homolog, which encodes an amidinotransferase similar to the human arginine:glycine amidinotransferase GATM, transfers a guanidino group from a donor molecule, most likely arginine, onto an acceptor molecule of glycine thus forming guanidinoacetate (Figure 8, step 1).

The next step (Figure 8, step 2) in the biosynthesis is carried out by CyrB (AoaB homolog), a mixed NRPS-PKS. CyrB spans 8.7 kb and encodes the following domains; adenylation domain (A domain) and a peptidyl carrier protein (PCP) of an NRPS followed by a βketosynthase domain (KS), acyltransferase domain (AT), dehydratase domain (DH), methyltransferase domain (MT), ketoreductase domain (KR), and an acyl carrier protein (ACP) of PKS origin. CyrB therefore must catalyse the second reaction since it is the only gene containing an A domain that could recruit a starter unit for subsequent PKS extensions. The specific amino acid activated by the CyrB A domain cannot be predicted as its substrate specificity conferring residues do not match any in the available databases (http://www-ab.informatik.uni-tuebingen.de/software/NRPSpredictor; http://www.tigr.org/jravel/nrps/, http://www.nii.res.in/nrps-pks.html). So far, no other NRPS has been described that utilizes guanidinoacetate as a substrate. The A domain is thought to activate guanidinoacetate, which is then transferred via the swinging arm of the peptidyl carrier protein (PCP) to the KS domain. The AT domain activates malonyl-CoA and attaches it to the ACP. This is followed by a condensation reaction between the activated guanidinoacetate and malonyl-CoA in the KS domain. CyrB contains two reducing modules, KR and DH. Their concerted reaction reduces the keto group to a hydroxyl followed by elimination of H_2O, resulting in a double bond between C13 and C14. The methyl transferase (MT) domain identified in CyrB via the NRPS/PKS databases (Example 9 above), is homologous to S-adenosylmethionine (SAM) dependent MT. It is therefore suggested that the MT methylates C13. It is proposed that a nucleophilic attack of the amidino group at N19 onto the newly formed double bond between C13 and C14 occurs via a 'Michael addition'. The cyclization follows Baldwin's rules for ring closure (Baldwin et al. (1997) J. Org. Chem 42;3846-3852), resulting in the

formation of the first ring in cylindrospermopsin. This reaction could be spontaneous and may not require enzymatic catalysis, as it is energetically favourable. This is the first of three ring formations.

The third step (Figure 8, step 3) in the biosynthesis involves CyrC (AoaC homolog), which encodes a PKS with KS, AT, KR, and ACP domains. The action of these domains results in the elongation of the growing chain by an acetate via activation of malonyl-CoA by the AT domain, its transfer to ACP and condensation at the KS domain with the product of CyrB. The elongated chain is bound to the ACP of CyrC and the KR domain reduces the keto group to a hydroxyl group on C12. The PKS module carrying out this step contains a KR domain and does not contain a DH domain, this corresponds only to CyrC.

Following the catalysis of enzyme CyrC is CyrD (Figure 8, step 4), a PKS with five modules; KS, AT, DH, KR, and an ACP. The action of this PKS module on the product of CyrC results in the addition of one acetate and the reduction of the keto group on C10 to a hydroxyl and dehydration to a double bond between C9 and C10. This double bond is the site of a nucleophilic attack by the amidino group N19 via another Michael addition that again follows Baldwin's rules of ring closure, resulting in the formation of the second ring, the first 6-membered ring made in cylindrospermopsin.

The product of CyrD is the substrate for CyrE (step 5 in Figure 8), a PKS containing a KS, AT, DH, KR domains and an ACP. Since this sequence of domains is identical to that of CyrD, it is not possible at this stage to ascertain which PKS acts first, but as their action is proposed to be identical it is immaterial at this point. CyrE catalyzes the addition of one acetate and the formation of a double bond between C7 and C8. This double bond is attacked by N18 via a Michael addition and the third cyclisation occurs, resulting in the second 6-member ring.

CyrF is the final PKS module (step 6 of Figure 8) and is a minimal PKS containing only a KS, AT, and ACP. CyrF acts on the product of CyrE and elongates the chain by an acetate, leaving C4 and C6 unreduced.

Step 7 in the pathway (Figure 8) involves the formation of the uracil ring, a reaction that is required for the toxicity of the final cylindrospermopsin compound. The cylindrospermopsin gene cluster encodes two enzymes with high sequence similarity (87%) that have been denoted CyrG and CyrH. A Psi-blast search (NCBI) followed by a Fugue profile library search (see materials and methods) revealed that CyrG and CyrH are most similar to the enzyme family of amidohydrolases/ureases/dihydroorotases, whose

members catalyze the formation and cleavage of N-C bonds. It is proposed that these enzymes transfer a second guanidino group from a donor molecule, such as arginine or urea, onto C6 and C4 of cylindrospermopsin resulting in the formation of the uracil ring. These enzymes carry out two or three reactions depending on the guanidino donor. The first reaction consists of the formation of a covalent bond between the N of the guanidino donor and C6 of cylindrospermopsin followed by an elimination of H_2O forming a double bond between C5 and C6. The second reaction catalyses the formation of a bond between the second N on the guanidino donor and C4 of cylindrospermopsin, co-committently with the breaking of the thioester bond between the acyl carrier protein of CyrE and cylindrospermopsin, causing the release of the molecule from the enzyme complex. Feeding experiments with labeled acetate have shown that the oxygen at C4 is of acetate origin and is not lost during biosynthesis, therefore requiring the *de novo* formation of the uracil ring. The third reaction - if required - would catalyze the cleavage of the guanidino group from a donor molecule other than urea. The action of CyrG and CyrH in the formation of the uracil ring in cylindrospermopsin describes a novel biosynthesis pathway of a pyrimidine.

One theory suggest a linear polyketide which readily assumes a favorable conformation for the formation of the rings. Cyclization may thus be spontaneous and not under enzymatic control. These analyses show that this may happen step-wise, with successive ring formation of the appropriate intermediate as it is synthesized. This mechanism also explains the lack of a thioesterase or cyclization domain, which are usually associated with NRPS/PKS modules and catalyze the release and cyclization of the final product from the enzyme complex.

Example 12: *CYR* tailoring reactions

Cylindrospermopsin biosynthesis requires the action of tailoring enzymes in order to complete the biosynthesis, catalyzing the sulfation at C12 and hydroxylation at C7. Analysis of the cylindrospermopsin gene cluster revealed three candidate enzymes for the tailoring reactions involved in the biosynthesis of cylindrospermopsin, namely CyrI, CyrJ, and CyrN. The sulfation of cylindrospermopsin at C12 is likely to be carried out by the action of a sulfotransferase. CyrJ encodes a protein that is most similar to human 3'-phosphoadenylyl sulfate (PAPS) dependent sulfotransferases. The cylindrospermopsin gene cluster also encodes an adenylsulfate kinase (ASK), namely CyrN. ASKs are enzymes that catalyze the formation of PAPS, which is the sulfate donor for

sulfotransferases. It is proposed that CyrJ sulfates cylindrospermopsin at C12 while CyrN creates the pool of PAPS required for this reaction. Screening of cylindrospermopsin producing and non-producing strains revealed that the sulfotransferase genes were only present in cylindrospermopsin producing strains, further affirming the involvement of this entire cluster in the biosynthesis of cylindrospermopsin (Figure 7). The *cyrJ* gene might therefore be a good candidate for a toxin probe, as it is more unique than NRPS and PKS genes and would presumably have less cross-reactivity with other gene clusters containing these genes, which are common in cyanobacteria. The final tailoring reaction is carried out by CyrI. A Fugue search and an iterated Psi-Blast revealed that CyrI is similar to a hydroxylase belonging to the 2-oxoglutarate and Fe(II)-dependent oxygenase superfamily, which includes the mammalian Prolyl 4-hydroxylase alpha subunit that catalyze the hydroxylation of collagen. It is proposed that CyrI catalyzes the hydroxylation of C7, a residue that, along with the uracil ring, seems to confer much of the toxicity of cylindrospermopsin. The hydroxylation at C7 by CyrI is probably the final step in the biosynthesis of cylindrospermopsin.

Example 13: *CYR* toxin transport

Cylindrospermopsin and other cyanobacterial toxins appear to be exported out of the producing cells. The cylindrospermopsin gene cluster contains an ORF denoted CyrK, the product of which is most similar to sodium ion driven multi-drug and toxic compound extrusion proteins (MATE) of the NorM family. It is postulated that CyrK is a transporter for cylindrospermopsin, based on this homology and its central location in the cluster. Heterologous expression and characterization of the protein are currently being undertaken to verify its putative role in cylindrospermopsin export.

Example 14: Transcriptional regulation of the toxin gene cluster

Cylindrospermopsin production has been shown to be highest when fixed nitrogen is eliminated from the growth media (Saker et al. (1999) J. Phycol 35:599-606). Flanking the cylindrospermopsin gene cluster are "*hyp*" gene homologs involved in the maturation of hydrogenases. In the cyanobacterium *Nostoc* PCC73102 they are under the regulation of the global nitrogen regulator NtcA, that activates transcription of nitrogen assimilation genes. It is plausible that the cylindrospermopsin gene cluster is under the same regulation, as it is located wholly within the "*hyp*" gene cluster in *C. raciborskii*

AWT205, and no obvious promoter region in the cylindrospermopsin gene cluster could be identified.

Finally, the cylindrospermopsin cluster also includes an ORF at its 3' -end designated CyrO. By homology, it encodes a hypothetical protein that appears to possess an ATP binding cassette, and is similar to WD repeat proteins, which have diverse regulatory and signal transduction roles. CyrO may also have a role in transcriptional regulation and DNA binding. It also shows homology to AAA family proteins that often perform chaperone-like functions and assist in the assembly, operation, or disassembly of protein complexes. Further insights into the role of CyrO are hindered due to low sequence homology with other proteins in databases.

The foregoing describes preferred forms of the present invention. It is to be understood that the present invention should not be restricted to the particular embodiment(s) shown above. Modifications and variations, obvious to those skilled in the art can be made thereto without departing from the scope of the present invention.

Related Application

This application claims benefit from Australian Provisional Application Number 2008902056 entitled "Detection of Cyanotoxic Organisms" which was filed on 24 April 2008, the entire contents of which are incorporated herein by reference.

CLAIMS:

1. An isolated polynucleotide comprising a sequence according to SEQ ID NO: 1 or a variant or fragment thereof.

2. The polynucleotide according to claim 1, wherein said fragment comprises a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, and variants and fragments thereof.

3. An isolated ribonucleic acid or an isolated complementary DNA encoded by a sequence according to claim 1 or claim 2.

4. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 67, SEQ ID NO: 69, and variants and fragments thereof.

5. A probe or primer that hybridises specifically with one or more of:

(i) a polynucleotide according to claim 1 or 2,

(ii) a ribonucleic acid or complementary DNA according to claim 3,

(iii) a polypeptide according to claim 4.

6. A vector comprising a polynucleotide according to claim 1 or claim 2, or a ribonucleic acid or complementary DNA according to claim 3.

7. A host cell comprising the vector according to claim 6.

8. A method for the detection of cyanobacteria, the method comprising the steps of obtaining a sample for use in the method and analyzing the sample for the presence of one or more of:

(i) a polynucleotide comprising a sequence according to claim 1 or 2,

(ii) a ribonucleic acid or complementary DNA according to claim 3,

(iii) a polypeptide comprising a sequence according to claim 4,

wherein said presence is indicative of cyanobacteria in the sample.

9. A method for detecting a cyanotoxic organism, the method comprising the steps of obtaining a sample for use in the method and analyzing the sample for the presence of one or more of:

(i) a polynucleotide comprising a sequence selected from the group consisting of: SEQ ID NO: 14, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 36, and variants and fragments thereof,

(ii) a ribonucleic acid or complementary DNA encoded by a sequence according to (i),

(iii) a polypeptide comprising a sequence selected from the group consisting of: SEQ ID NO: 15, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 37, and variants and fragments thereof,

wherein said presence is indicative of cyanotoxic organisms in the sample.

10. The method according to claim 9, wherein said cyanotoxic organism is a cyanobacteria or a dinoflagellate.

11. The method according to any one of claims 8 to 10, wherein said analyzing comprises amplification of DNA from the sample by polymerase chain reaction and detecting the amplified seqeunces.

12. The method according to claim 11, wherein said polymerase chain reaction utilises one or more primers comprising a sequence selected from the group consisting of SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 117, SEQ ID NO: 118, SEQ ID NO: 119, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 122, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 130, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 133, SEQ ID NO: 134, and variants and fragments thereof.

13. The method according to any one of claims 8 to 12, further comprising analyzing the sample for the presence of one or more of:

(i) a polynucleotide comprising a sequence selected from the group consisting of: SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87,

SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, and variants and fragments thereof,

(ii) a ribonucleic acid or complementary DNA encoded by a sequence according to (i),

(iii) a polypeptide comprising a sequence selected from the group consisting of: SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92 , SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, and SEQ ID NO: 110, and variants and fragments thereof.

14. The method according to claim 13, wherein said analyzing comprises amplification of DNA from the sample by polymerase chain reaction.

15. The method according to claim 13, wherein said polymerase chain reaction utilises one or more primers comprising a sequence selected from the group consisting of SEQ ID NO: 111, SEQ ID NO: 112, and variants and fragments thereof.

16. A method for the detection of dinoflagellates, the method comprising the steps of obtaining a sample for use in the method and analyzing the sample for the presence of one or more of:

(i) a polynucleotide comprising a sequence according to claim 1 or 2,

(ii) a ribonucleic acid or complementary DNA according to claim 3,

(iii) a polypeptide comprising a sequence according to claim 4,

wherein said presence is indicative of dinoflagellates in the sample.

17. The method according to claim 16, wherein said analyzing comprises amplification of DNA from the sample by polymerase chain reaction and detecting the amplified sequences.

18. The method according to any one of claims 8 to 17, wherein said sample comprises one or more isolated or cultured organisms.

19. The method according to any one of claims 8 to 18, wherein said sample is an environmental sample.

20. The method according to claim 19, wherein said environmental sample is derived from salt water, fresh water or a blue-green algal bloom.

21. An isolated antibody capable of binding specifically to a polypeptide according to claim 4.

22. A kit for the detection of cyanobacteria, the kit comprising at least one agent for detecting the presence of one or more of:

(i) a polynucleotide comprising a sequence according to claim 1 or 2,

(ii) a ribonucleic acid or complementary DNA according to claim 3,

(iii) a polypeptide comprising a sequence according to claim 4,

wherein said presence is indicative of cyanobacteria in the sample.

23. A kit for the detection of cyanotoxic organisms, the kit comprising at least one agent for detecting the presence of one or more of:

(i) a polynucleotide comprising a sequence selected from the group consisting of: SEQ ID NO: 14, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 36, and variants and fragments thereof,

(ii) a ribonucleic acid or complementary DNA encoded by a sequence according to (i),

(iii) a polypeptide comprising a sequence selected from the group consisting of: SEQ ID NO: 15, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 37, and variants and fragments thereof,

wherein said presence is indicative of cyanotoxic organisms in the sample.

24. The kit according to claim 22 or claim 23, wherein said at least one agent is a primer, antibody or probe.

25. The kit according to claim 24, wherein said primer or probe comprises a sequence selected from the group consisting of SEQ ID NO: 70, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 78, SEQ ID NO: 79, SEQ ID NO: 113, SEQ ID NO: 114, SEQ ID NO: 115, SEQ ID NO: 116, SEQ ID NO: 117, SEQ ID NO: 118, SEQ ID NO: 119, SEQ ID NO: 120, SEQ ID NO: 121, SEQ ID NO: 122, SEQ ID NO: 123, SEQ ID NO: 124, SEQ ID NO: 125, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 128, SEQ ID NO: 129, SEQ ID NO: 130, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 133, SEQ ID NO: 134, and variants and fragments thereof.

26. The kit according to any one of claims 22 to 25, further comprising at least one additional agent for detecting the presence of one or more of:

(i) a polynucleotide comprising a sequence selected from the group consisting of: SEQ ID NO: 80, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87, SEQ ID NO: 89, SEQ ID NO: 91, SEQ ID NO: 93, SEQ ID NO: 95, SEQ ID NO: 97,

SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, and variants and fragments thereof,

(ii) a ribonucleic acid or complementary DNA encoded by a sequence according to (i),

(iii) a polypeptide comprising a sequence selected from the group consisting of: SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88, SEQ ID NO: 90, SEQ ID NO: 92 , SEQ ID NO: 94, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, and SEQ ID NO: 110, and variants and fragments thereof.

27. The kit according to claim 26, wherein said at least one additional agent is a primer, antibody or probe.

28. The kit according to claim 27, wherein said primer or probe comprises a sequence selected from the group consisting of SEQ ID NO: 109, SEQ ID NO: 110, and variants and fragments thereof.

29. A kit for the detection of dinoflagellates, the kit comprising at least one agent for detecting the presence of one or more of:

(i) a polynucleotide comprising a sequence according to claim 1 or 2,

(ii) a ribonucleic acid or complementary DNA according to claim 3,

(iii) a polypeptide comprising a sequence according to claim 4,

wherein said presence is indicative of dinoflagellates in the sample.

30. A method of screening for a compound that modulates the expression or activity of one or more polypeptides according to claim 4, the method comprising:

contacting the polypeptide with a candidate compound under conditions suitable to enable interaction of the candidate compound and the polypeptide; and

assaying for activity of the polypeptide.

31. The method according to claim 30 wherein said modulation comprises inhibiting expression or activity of said polypeptide.

32. The method according to claim 30, wherein said modulation comprises enhancing expression or activity of said polypeptide.

Cyanobacteria Strains	Toxicity (Ref)	*sxtA*	*sxtG*	*sxtH*	*sxtI*	*sxtX*
A. circinalis AWQC118C	PSP (54)	+	+	+	+	-
A. circinalis AWQC131C	PSP (25)	+	+	+	+	-
A. circinalis AWQC134C	PSP (54)	+	+	+	+	-
A. circinalis AWQC150E	PSP (54)	+	+	+	+	-
A. circinalis AWQC173A	PSP (54)	+	+	+	+	-
A. circinalis AWQC271C	- (54)	-	-	-	-	-
A. circinalis AWQC306A	- (54)	-	-	-	-	-
A. circinalis AWQC310F	- (54)	-	-	-	-	-
A. circinalis AWQC342D	- (54)	-	-	-	-	-
Aph. flos-aquaea NH-5	PSP (26)	+	+	+	+	+
Aph. ovalisporum APH028A	CYLN (46)	-	-	-	-	-
C. raciborskii T3	PSP (23)	+	+	+	+	+
C. raciborskii 23B	CYLN (58)	-	-	-	-	-
C. raciborskii GOON	CYLN (43)	-	-	-	-	-
C. raciborskii GERM1	- (30)	-	-	-	-	-
C. raciborskii MARAU1	- (30)	-	-	-	-	-
L. wollei	PSP (7)	+	+	+	+	+

FIGURE 1A

Primer	From	To	Direction	Sequence	Gene	
SEQ ID NO:133	1917	1937	-->	GCAAATTTTGCAGGAGTAATG	sterole desaturase	sxtD
SEQ ID NO:134	2744	2763	-->	AGAGATGCTATGCTTTTTCAA	InsB	orf3
SEQ ID NO:135	2889	2911	-->	TTTTGGGTAAACTTTATATGCCAT	InsB	orf3
SEQ ID NO:136	3020	3041	<--	TGGGTCTGGACAGTTGTAGATA	InsA	orf4
SEQ ID NO:137	3306	3328	<--	AAGGGGAAAACAAAAATTATCAAT	InsA	orf4
SEQ ID NO:138	3396	3415	-->	GGCGATCGCCTGCTAAAAAT	InsA	orf4
SEQ ID NO:139	3717	3739	-->	CCTCATTTTCATTTCTAGACGTT	SPUR	sxtC
SEQ ID NO:140	4201	4220	-->	CCACTTCAACTAAAACAGCA	cytidine deaminase	sxtB
SEQ ID NO:141	4362	4381	<--	AAAAATTTTGGAGGGGTAGC	cytidine deaminase	sxtB
SEQ ID NO:142	4932	4951	-->	ATCCAAGATGCGACAACACT	cytidine deaminase	sxtB
SEQ ID NO: 70	5193	5212	-->	TTAATTGCTTGGTCTAATCTC	PKS	sxtA
SEQ ID NO: 71	5206	5225	<--	CAATACCGAAGAGGAGATAG	PKS	sxtA
SEQ ID NO: 72	5345	5364	<--	TAGGCGTGTTAGTGGGAGAT	PKS	sxtA
SEQ ID NO: 73	5415	5434	-->	TGTGTAACCAATTTGTGAGT	PKS	sxtA
SEQ ID NO: 74	5478	5497	<--	TTAGCCGGATTACAGGTGAA	PKS	sxtA
SEQ ID NO: 75	6136	6155	<--	CTGGACTCGGCTTGTTGCTT	PKS	sxtA
SEQ ID NO: 76	6933	6952	-->	CAGCGAGTTACACCCACCAC	PKS	sxtA
SEQ ID NO: 76	7035	7054	<--	CTCGCACTAAATATTCTACC	PKS	sxtA
SEQ ID NO: 78	7434	7452	-->	AAAACCTCAGCTTCCACAA	PKS	sxtA
SEQ ID NO: 79	7537	7558	<--	ATGATTTTGGAGGTCCATTGTT	PKS	sxtA
SEQ ID NO: 113	7820	7841	-->	CCCAATATCTCCCTGTAAAACT	PKS	sxtA
SEQ ID NO: 114	8170	8189	<--	TGGCAATTGTCTCTCCGTAT	PKS	sxtA
SEQ ID NO: 115	8742	8761	<--	CTCGCCGATGAAAGTCCTCT	PKS	sxtA
SEQ ID NO: 116	8772	8791	<--	GCGTGTCGAGAAAAAGGTGT	PKS	sxtA
SEQ ID NO: 117	8782	8801	-->	CTCGACACGCAAGAATAACG	PKS	sxtA
SEQ ID NO: 143	9390	9410	-->	GGTCCTTGCGCAGATAGAGTG	chaperon-like	sxtE
SEQ ID NO: 144	9390	9410	<--	CACTCTATCTGCGCAAGGACC	chaperon-like	sxtE
SEQ ID NO: 145	9856	9876	<--	TGACTGCATTCGCTGTATAAA	MATE1	sxtF
SEQ ID NO: 118	10080	10100	-->	ATGCTTCTGCTTTGGCAIGGC	amidinotransferase	sxtG
SEQ ID NO: 119	11468	11488	<--	TAACTCGACGAACTTTGACCC	amidinotransferase	sxtG

FIGURE 1B

Primer	From	To	Direction	Sequence	Gene	
SEQ ID NO: 120	11551	11569	-->	GCCGCCAATCCTCGCGATG	amidinotransferase	sxtG
SEQ ID NO: 121	12256	12277	<--	GAACGTCTAATGTTGCACAGTG	amidinotransferase	sxtG
SEQ ID NO: 122	12410	12432	-->	CTGGTACGTAGTCGCAAAGGTGG	dioxygenase I	sxtH
SEQ ID NO: 123	13292	13317	-->	CTGACGGTACATGTATTTCCTGTGAC	dioxygenase I	sxtH
SEQ ID NO: 124	13540	13561	-->	cgtctcatATGCAGATCTTAGGAATTTCAG	carbamoyltransferase	sxtI
SEQ ID NO: 125	13561	13585	-->	GCTTACTACCACGATAGTGCTGCCG	carbamoyltransferase	sxtI
SEQ ID NO: 126	14451	14472	-->	TCTATGTTTAGCAGGTGGTGTC	carbamoyltransferase	sxtI
SEQ ID NO: 127	14735	14754	<--	TTCTGCAAGACGAGCCATAA	carbamoyltransferase	sxtI
SEQ ID NO: 128	15211	15230	<--	GGTTCGCCGCGGACATTAAA	carbamoyltransferase	sxtI
SEQ ID NO: 146	15709	15730	-->	TTCATAAGACGGCTGTTGAATC	hypothetical protein	sxtJ
SEQ ID NO: 147	15966	15989	<--	ctcgagTTAAAAAAGAGTGTAAATGAAAGG	hypothetical protein	sxtK
SEQ ID NO: 148	16326	16348	<--	TTCTATAACTGCTGCCAAATTTT	GDSL-lipase	sxtL
SEQ ID NO: 149	16400	16422	-->	AATTTTGGAGTGACTGGTTATGG	GDSL-lipase	sxtL
SEQ ID NO: 150	16400	16422	<--	CCATAACCAGTCACTCCAAAATT	GDSL-lipase	sxtL
SEQ ID NO: 151	16929	16949	-->	TTTTAGTTGTTACTTTTGGCG	GDSL-lipase	sxtL
SEQ ID NO: 152	17215	17234	-->	ACAGCAGATGAGAGAAAGTA	GDSL-lipase	sxtL
SEQ ID NO: 153	18054	18073	-->	GGGTTGTCTTGCTGATTTTC	MATE II	sxtM
SEQ ID NO: 154	18721	18742	<--	CATTAAAATAAGTCCGGACAGG	MATE II	sxtM
SEQ ID NO: 155	19133	19152	<--	TTAAACAGAATGAGGAGCAA	MATE II	sxtM
SEQ ID NO: 156	19260	19279	<--	AAACAACACACCCATCTAAG	sulfotransferase	sxtN
SEQ ID NO: 157	19531	19550	-->	TTAATAAGGCATCCCCAAGA	sulfotransferase	sxtN
SEQ ID NO: 158	19728	19747	<--	GAAATGGCTGTGTAAAAACT	sulfotransferase	sxtN
SEQ ID NO: 129	20584	20603	<--	ATGCTAATGCGGTGGGAGTA	cephalosporin hydroxylase	sxtX
SEQ ID NO: 130	20643	20662	-->	AAAGCAGTTCCGACGACATT	cephalosporin hydroxylase	sxtX
SEQ ID NO: 131	20831	20853	-->	CCTATTTCGATTATTGTTTTCGG	cephalosporin hydroxylase	sxtX
SEQ ID NO: 132	21252	21271	<--	GATACCGATCATAAACTACG	cephalosporin hydroxylase	sxtX
SEQ ID NO: 159	21290	21309	-->	TCTGCCATATCCCCAACCTA	ferredoxin	sxtW
SEQ ID NO: 160	21445	21464	<--	GATCGCCCGACAGGAAGACT	ferredoxin	sxtW
SEQ ID NO: 161	22020	22039	-->	TCCGGCTTGACCTGCTGGAC	succinate dehydrogenase	sxtV

FIGURE 1B (cont)

Primer	From	To	Direction	Sequence	Gene	
SEQ ID NO: 162	22715	22734	-->	TGCGATGATTTTGCCTCTGT	succinate dehydrogenase	sxtV
SEQ ID NO: 163	22801	22820	-->	AAAATTTGCACACCCACACG	succinate dehydrogenase	sxtV
SEQ ID NO: 164	22942	22968	-->	TTGGATTGAACGTGTAATTGAAAAAGC	succinate dehydrogenase	sxtV
SEQ ID NO: 186	22942	22968	<--	GCTTTTTCAATTACACGTTCAATCCAA	succinate dehydrogenase	sxtV
SEQ ID NO: 165	23434	23453	<--	GTTTAGTCGATACGCCATTT	succinate dehydrogenase	sxtV
SEQ ID NO: 166	23434	23453	-->	AAATGGCGTATCGACTAAC	succinate dehydrogenase	sxtV
SEQ ID NO: 167	24095	24115	-->	ATATAGGAGCGCATAAAGTGC	succinate dehydrogenase	sxtV
SEQ ID NO: 168	24728	24747	-->	CTTGGTATAAGTCTTGTGAT	dioxygenase II	sxtT
SEQ ID NO: 169	25426	25445	<--	AACACTCATTAGATTCATCT	phytanoyl-CoA dioxygenase	sxtS
SEQ ID NO: 170	25979	25999	<--	TCCACTAAATCCTTTGAATTG	phytanoyl-CoA dioxygenase	sxtS
SEQ ID NO: 171	26279	26299	-->	TGTTTGTCTGGATGCGATCCT	unknown protein	orf24
SEQ ID NO: 172	26451	26470	-->	GCAGTTCAGGTCCATGAAAC	unknown protein	orf25
SEQ ID NO: 173	27155	27174	-->	AGCCCAGTCACACAACCTTCGT	GNAT transferase	sxtR
SEQ ID NO: 174	27508	27528	-->	TCTGGAAGTACTTGCACTGTC	unknown protein	sxtQ
SEQ ID NO: 175	28197	28218	-->	TGTAACTCCGTCAGGACATAAA	unknown protein	sxtQ
SEQ ID NO: 176	28395	28417	<--	TGCAAATTTTAGTAGCAATAACG	RTX-toxin like	sxtP
SEQ ID NO: 177	29532	29558	<--	CTTTACTAATTATAGCGGGGGATATTAT	RTX-toxin like	sxtP
SEQ ID NO: 178	29868	29887	<--	CAGTGGGGAAATAGATGGAT	adenylylsulphate kinase	sxtO
SEQ ID NO: 179	30249	30268	<--	TGGTCATAAAAGCGGGATTC	adenylylsulphate kinase	sxtO
SEQ ID NO: 180	31745	31762	-->	GGATCTTGGCGCAATTTA	IS4	orf29
SEQ ID NO: 181	33031	33053	<--	GTTAGAGACTTGGAACGTATTGG	PhoU	sxtY
SEQ ID NO: 182	34711	34729	-->	CCAAACCCAGAAGAAATCC	histidine kinase	sxtZ
SEQ ID NO: 183	35100	35121	-->	AATCTATAGCCAAAACCCCTAA	ribotide isomerase	
SEQ ID NO: 184	36447	36465	-->	ACTGTGTGAACAATTCCCC	ribotide isomerase	
SEQ ID NO: 185	36652	36680	-->	GCAACAAGACTACATTTAGTAGATTTAGA	ribotide isomerase	

FIGURE 1B (cont)

Name	Enzyme Family	Size (bp)	Blast Similarity Match	(%)	Putative Function
orf1	unknown protein	1320	BAB76734.1 *Nostoc* PCC7120	82	unknown
sxtD	sterole desaturase-like	759	ABG52264.1 *Trichodesmium erythraeum*	63	desaturation
orf3	transposase IusB	392	CAE11915.2 *Microcystis aeruginosa*	86	transposition
orf4	transposase IusA	360	CAE11914.1 *Microcystis aeruginosa*	71	transposition
sxtC	unknown protein	354	no similarity found		regulatory
sxtB	cytidine deaminase	957	EAS64681.1 *Vibrio angustum*	62	cyclisation
sxtA	methyltransferase	1506	ABF89568.1 *Idiococcus xanthus*	64	methylation
	GNAT	633	AAT70096.1 CurA *Lyngbya majuscula*	64	loading of ACP
	acyl carrier protein	324	AAV97870 OnuB *Theonella swinhoei*	59	ACP
	AONS	1275	ABD13093.1 *Frankia* sp. Cel3	61	Claisen condensation
sxtE	unknown protein	387	ABE53436.1 *Shewanella denitrificans*	52	unknown
sxtF	MATE	1416	NorM ABC44739.1 *Salinibacter ruber*	52	export of PSTs
sxtG	amidinotransferase	1134	ABA05575.1 *Nitrobacter winogradskyi*	71	amidinotransfer
sxtH	phenylpropionate dioxygenase	1005	ZP_00243439.1 *Rubrivivax gelatinosus*	50	C-12 hydroxylation
sxtI	carbamoyltransferase	1839	ABG50968.1 *Trichodesmium erythraeum*	82	carbamoylation

FIGURE 2

sxtJ	unknown protein	444	EAM51043.1 *Cyanothece watsonii*	72	regulatory
sxtK	unknown protein	165	ABG50954.1 *Trichodesmium erythraeum*	81	regulatory
sxtL	GDSL-lipase	1299	ABG50952.1 *Trichodesmium erythraeum*	60	cyclisation
sxtM	MATE	1449	NorM ABC44739.1 *Salinibacter ruber*	53	export of PSTs
sxtN	sulfotransferase	831	ABG53102.1 *Trichodesmium erythraeum*	57	sulfotransfer
sxtX	cephalosporin hydroxylase	774	ABG50679.1 *Trichodesmium erythraeum*	77	N-1 hydroxylation
sxtW	ferredoxin	327	ZP_00106179.2 *Nostoc punctiforme*	99	electron carrier
sxtV	succinate dehydrogenase	1653	ABA24604.1 *Anabaena variabilis*	92	dioxygenase reductase
sxtU	alcohol dehydrogenase	750	ZP_00111652.1 *Nostoc punctiforme*	83	reduction of C-1
sxtT	phenylpropionate dioxygenase	1005	ZP_00243439.1 *Rubrivivax gelatinosus*	48	C-12 hydroxylation
sxtS	phytanoyl-CoA dioxygenase	726	ABG30370.1 *Roseobacter denitrificans*	41	ring formation
orf24	unknown protein	576	no similarity found		unknown
sxtR	acyl transferase	777	AAU26161.1 *Legionella pneumophila*	54	unknown
sxtQ	unknown protein	777	EAR64935.1 *Bacillus* sp. NRRL B-14911	46	unknown
sxtP	RTX-toxin	1227	ABA20206.1 *Anabaena variabilis*	68	binding of PSTs
sxtO	adenylylsulfate kinase	603	ZP_00053494.2 *Magnetospirillum magnetotacticum*	76	PAPS biosynthesis
orf29	transposase, IS4	1350	EAO22567.1 *Symtrophobacter fumaroxidans*	61	transposition
sxtY	PhoU	666	BAB76200.1 *Nostoc* PCC7120	87	signal transduction
sxtZ	histidine kinase	1353	ABA22975.1 *Anabaena variabilis*	78	signal transduction
ompR	OmpR	819	ZP_00108178.2 *Nostoc punctiforme*	91	signal transduction
hisA	PROFAR isomerase	774	ABA22979.1 *Anabaena variabilis*	90	histidine biosynthesis
orf34	unknown protein	396	ZP_00345366.1 *Nostoc punctiforme*	84	unknown

FIGURE 2 (cont)

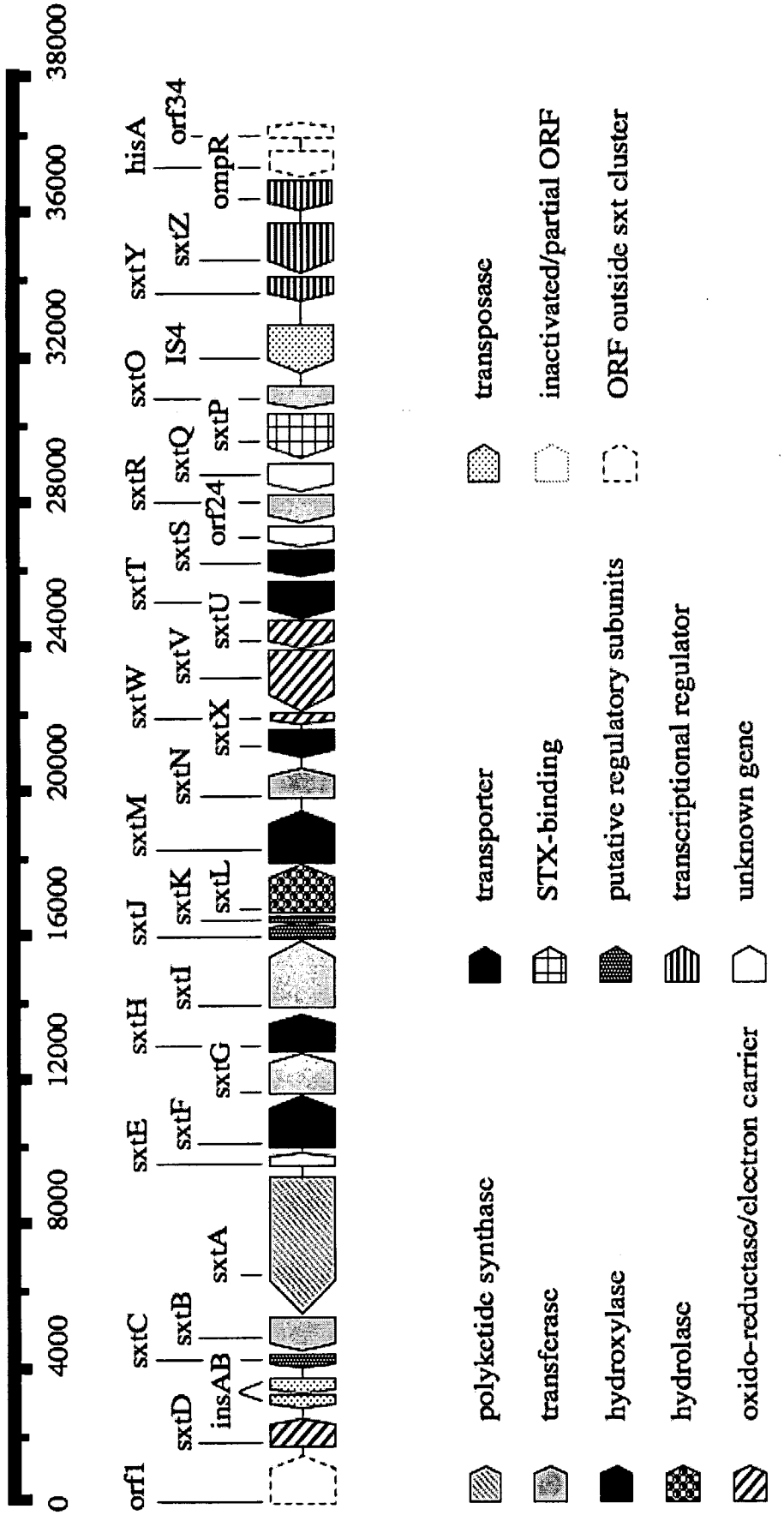
0
4000
8000
12000
16000
20000
24000
28000
32000
36000
38000
orf1
sxtD
insAB
sxtC
sxtB
sxtA
sxtE
sxtF
sxtG
sxtH
sxtI
sxtJ
sxtK
sxtL
sxtM
sxtN
sxtX
sxtW
sxtV
sxtU
sxtT
sxtS
orf24
sxtR
sxtQ
sxtP
sxtO
IS4
sxtY
sxtZ
ompR
hisA
orf34
polyketide synthase
transporter
transposase
transferase
STX-binding
inactivated/partial ORF
hydroxylase
putative regulatory subunits
ORF outside sxt cluster
hydrolase
transcriptional regulator
oxido-reductase/electron carrier
unknown gene

FIGURE 3

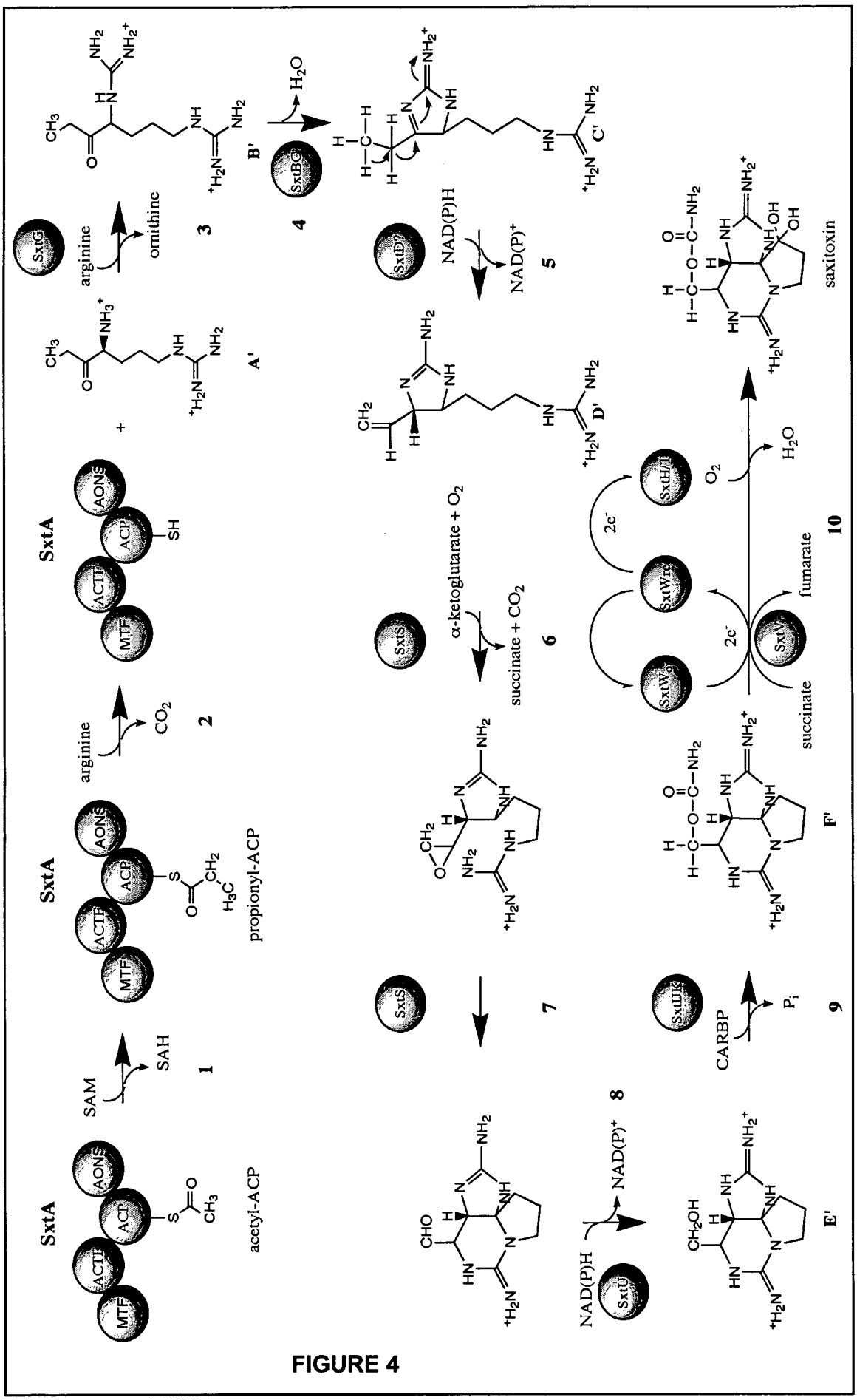
SxtA
MTF
ACTF
ACP
AONS
SAM
SAH
1
acetyl-ACP
propionyl-ACP
arginine
CO2
2
A'
SxtG
ornithine
3
B'
SxtBG
H2O
4
C'
SxtD?
NAD(P)H
NAD(P)+
5
D'
SxtS
α-ketoglutarate + O2
succinate + CO2
6
7
SxtU
NAD(P)H
NAD(P)+
8
E'
SxtH/I
SxtW
SxtV
2e-
O2
H2O
fumarate
succinate
F'
10
SxtH/I/K
CARBP
Pi
9
saxitoxin

FIGURE 4

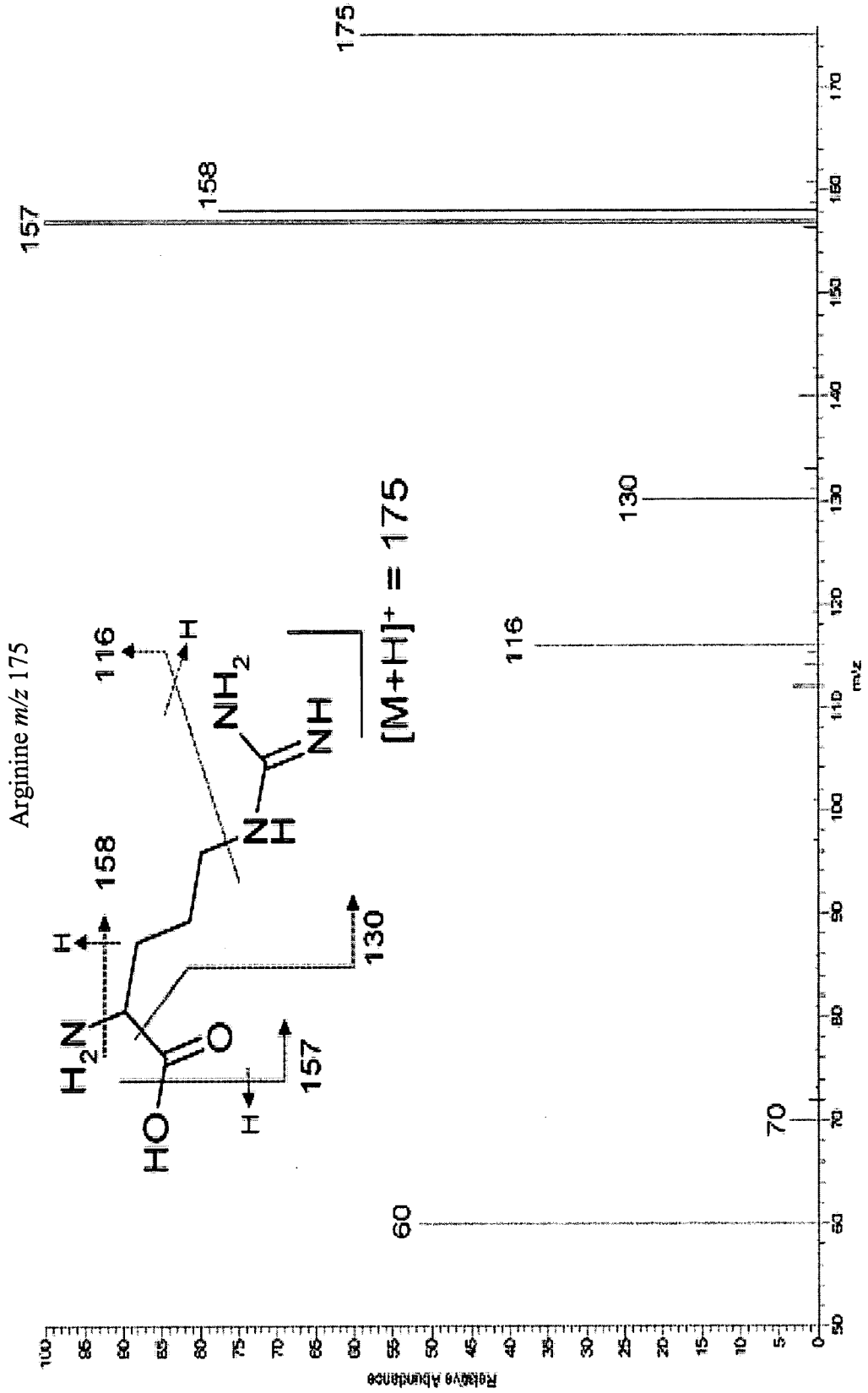
Arginine m/z 175
[M+H]+ = 175
175
158
157
130
116
70
60
m/z
Relative Abundance

FIGURE 5A

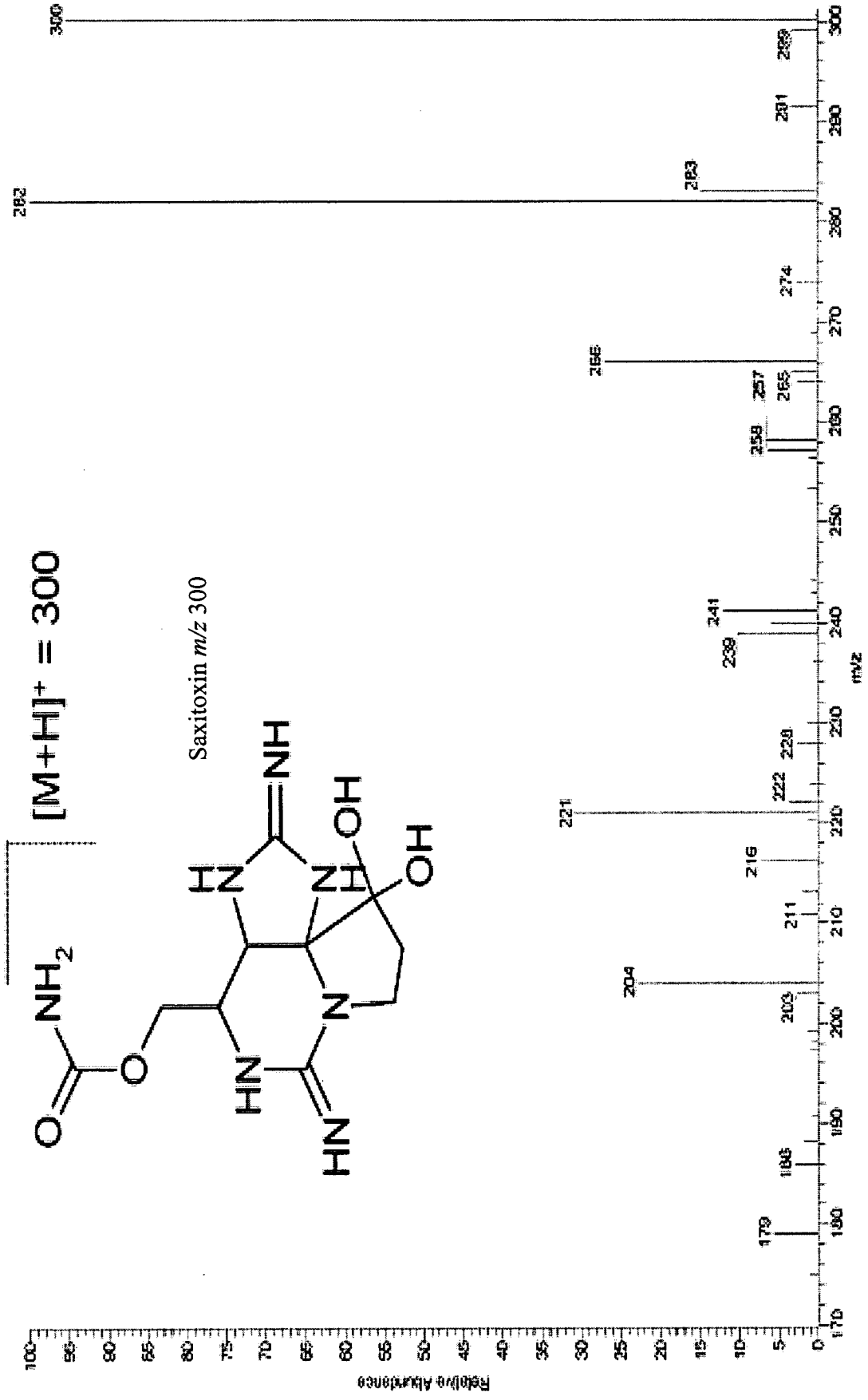
[M+H]+ = 300
Saxitoxin m/z 300
NH2
O
O
HN
HN
N
NH
OH
OH
Relative Abundance
m/z

FIGURE 5B

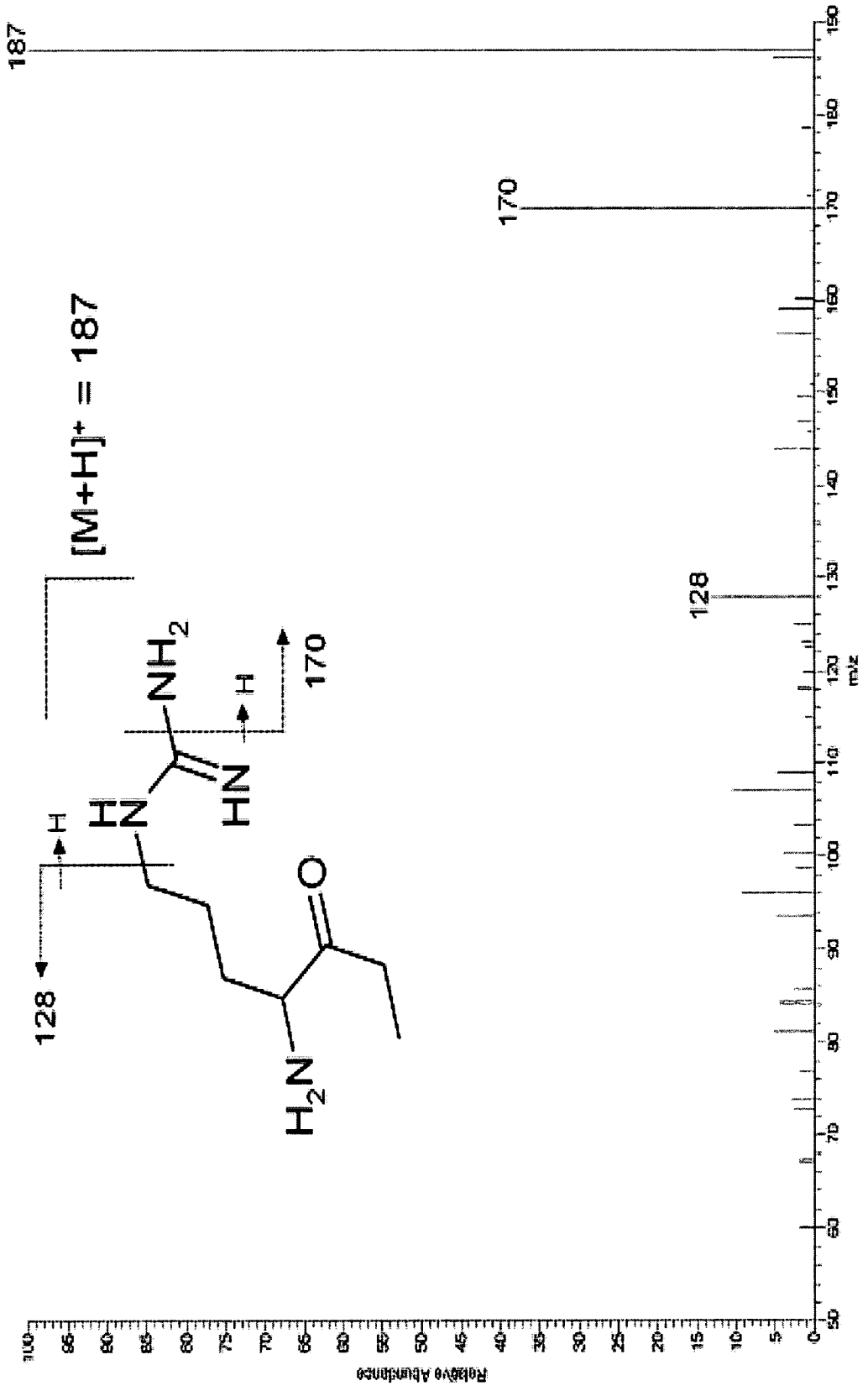
187
170
128
[M+H]+ = 187
NH2
HN
HN
H
170
H
128
O
H2N
m/z
Relative Abundance

FIGURE 5C

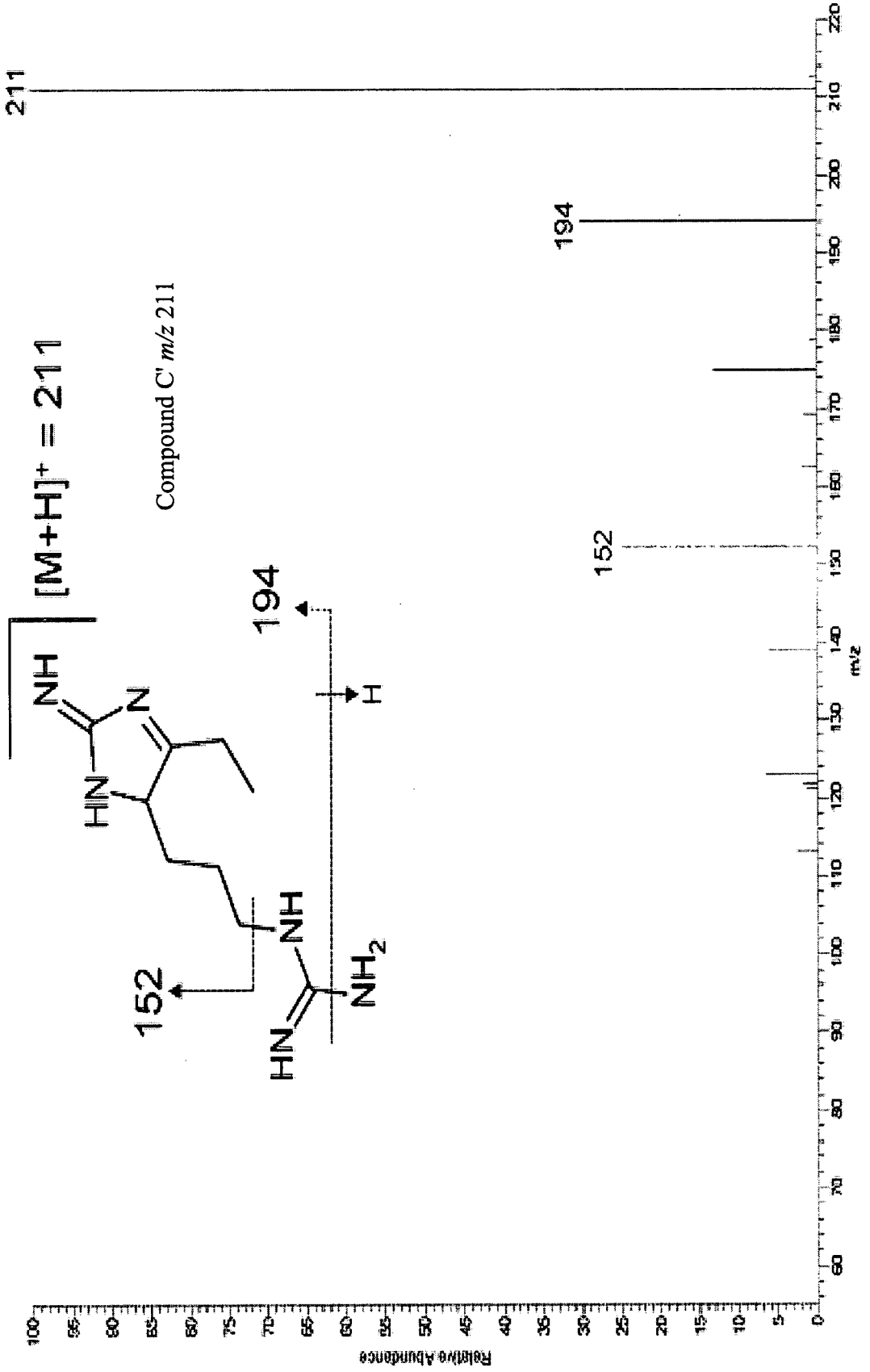
211
194
152
[M+H]+ = 211
Compound C' m/z 211
194
152
H
NH
N
HN
NH
NH2
HN
Relative Abundance
m/z

FIGURE 5D

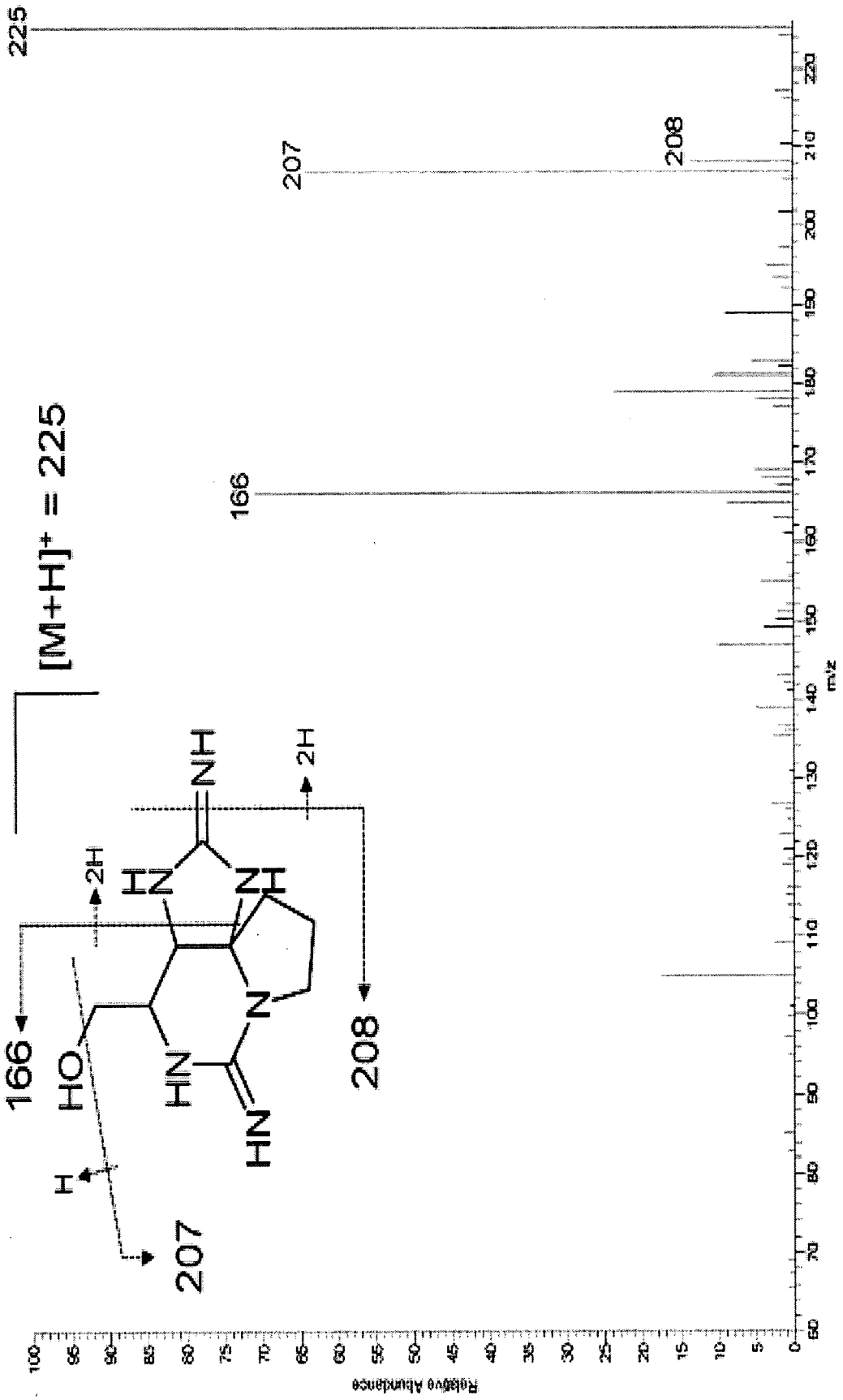
[M+H]+ = 225
225
207
208
166
166
HO
HN
HN
NH
N
2H
2H
H
208
207
m/z
Relative Abundance

FIGURE 5E

Name	Enzyme Family	Size (bp)	Psi-Blast similarity match	% ID	Putative function
cyrD	PKS CrpB	5631	ABM21570.1 *Nostoc sp.* ATCC 53789	58	PKS KS-AT-DH-KR-ACP
cyrF	PKS CrpB	4074	ABM21570.1 *Nostoc sp.* ATCC 53789	68	PKS KS-AT-ACP
cyrG	cytosine deaminase /Aminohydrolase/ Dihydroorotase	1437	BAF59909.1 *Pelotomaculum thermopropionicum* SI	50	Uracil ring formation
cyrI	Prolyl 4-Hydroxylase	831	ABB06365.1 *Burkholderia* sp. 383	43	Hydroxylation of C7
cyrK	MatE Na+-driven multidrug efflux pump	1398	EAW39051.1 *Lyngbya* sp. PCC 8106	65	Exporter
cyrL	Transposase	750	ABG50981.1 *Trichodesmium erythraeum* IMS101	70	Transposase
cyrH	cytosine deaminase /Aminohydrolase/ Dihydroorotase	1431	BAF59909.1 *Pelotomaculum thermopropionicum* SI	50	Uracil ring formation
cyrJ	branched-chain amino acid aminotransferase	780	*Trichodesmium erythraeum* IMS101	53	sulfotransferase
cyrA	Amidinotransferase AoaA	1176	AAX81898.1 *Cylindrospermopsis raciborskii*	100	amidinotransferase
cyrB	NRPS/PKS AoaB	8754	AAM33468.1 *Aphanizomenon ovalisporum*	97	NRPS/PKS A-domain, pp, KS, AT, DH, Met, KR, ACP
cyrE	PKS	5667	ABA23591.1 *Anabaena variabilis* ATCC 29413	62	PKS KS-AT-DH-KR-ACP
cyrC	PKS AoaC	5005	AAM33470.1 *Aphanizomenon ovalisporum*	97	PKS KS-AT-KR-ACP
cyrM	Partial Transposase	318	ABG50981.1 *Trichodesmium erythraeum* IMS101	70	Transposase
cyrN	Adenylylsulfate kinase (PAPS)	600	CAM76460.1 *Magnetospirillum gryphiswaldense* MSR-1	75	Adenylylsulfate kinase (PAPS)
cyrO	hypothetical protein	1548	EAW46978.1 *Nodularia spumigena* CCY9414	74	Regulator

FIGURE 6

Cyanobacterial Strain	**16s rRNA**	***cyrJ***	**Toxicity**	**Reference**
Cylindrospermopsis raciborskii T3	+	-	SXT	Lagos et al. (1999)
Anabaena circinalis 344B	+	-	N.D.	AWQC
Cylindrospermopsis raciborskii Germ1	+	-	N.D.	Neilan et al. (2003)
Anabaena circinalis 310F	+	-	N.D.	AWQC
Cylindrospermopsis raciborskii 44D	+	-	N.D.	NA
Anabaena circinalis 118C	+	-	SXT	Fergusson et al. (2000)
Anabaena circinalis 323A	+	-	N.D.	AWQC
Anabaena circinalis 323H	+	-	N.D.	AWQC
Cylindrospermopsis raciborskii VOLL2	+	-	N.D.	Neilan et al. (2003)
Cylindrospermopsis raciborskii VOLL1	+	-	N.D.	Neilan et al. (2003)
Cylindrospermopsis raciborskii HUNG1	+	-	N.D.	NA
Cylindrospermopsis raciborskii 023B	+	+	CYLN	Wilson et al. (2000)
Cylindrospermopsis raciborskii 05E	+	+	CYLN	Schembri et al. (2001)
Cylindrospermopsis raciborskii 4799	+	+	CYLN	Neilan et al. (2003)
Cylindrospermopsis raciborskii 24C	+	+	CYLN	Schembri et al. (2001)
Cylindrospermopsis raciborskii AWT 205	+	+	CYLN	Hawkins et al. (1997)
Aphanizomenon ovalisporum AO/QH	+	+	CYLN	NA

FIGURE 7

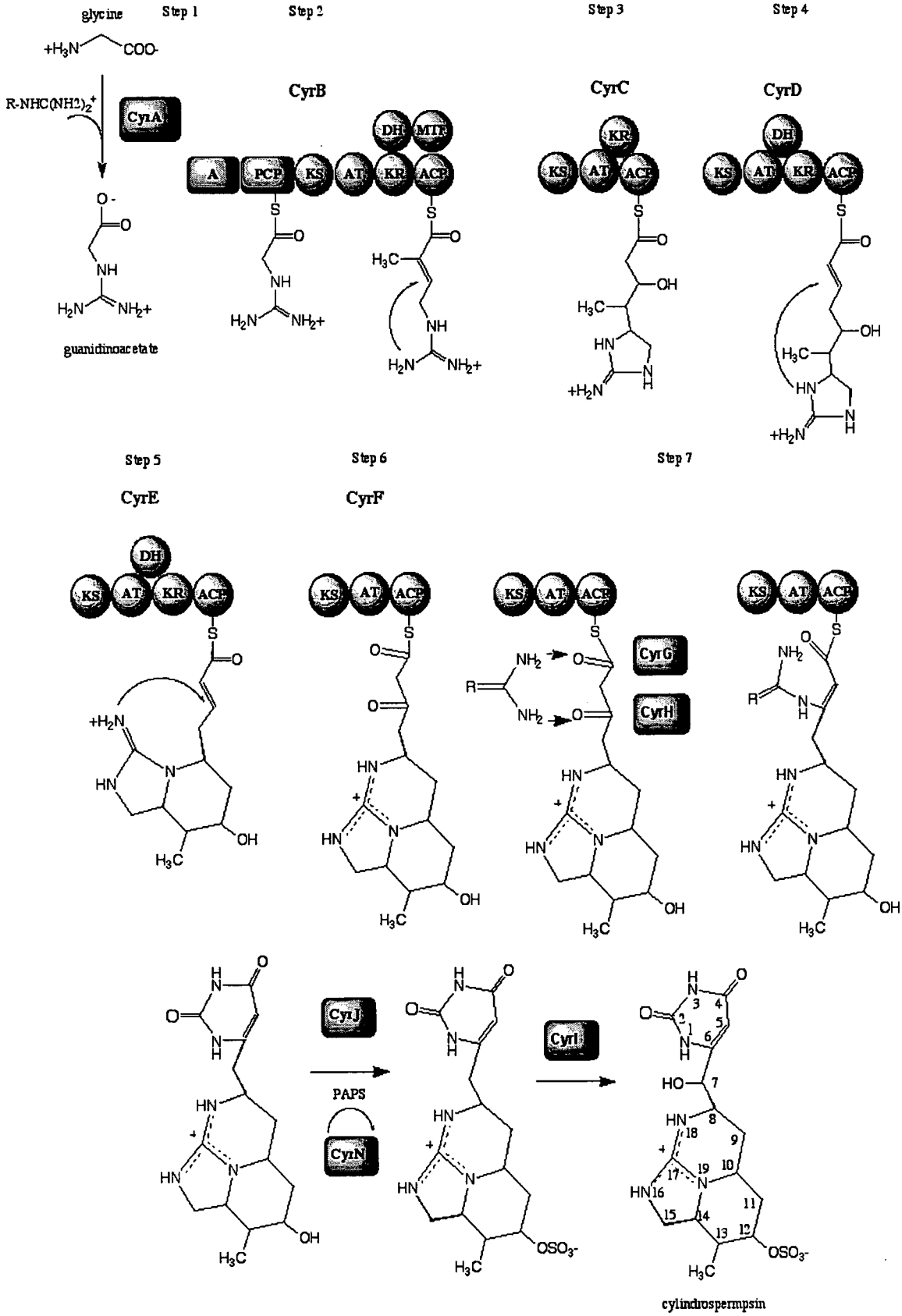
glycine
Step 1
Step 2
Step 3
Step 4
CyrA
CyrB
CyrC
CyrD
guanidinoacetate
Step 5
Step 6
Step 7
CyrE
CyrF
CyrG
CyrH
CyrJ
PAPS
CyrN
CyrI
cylindrospermpsin

FIGURE 8

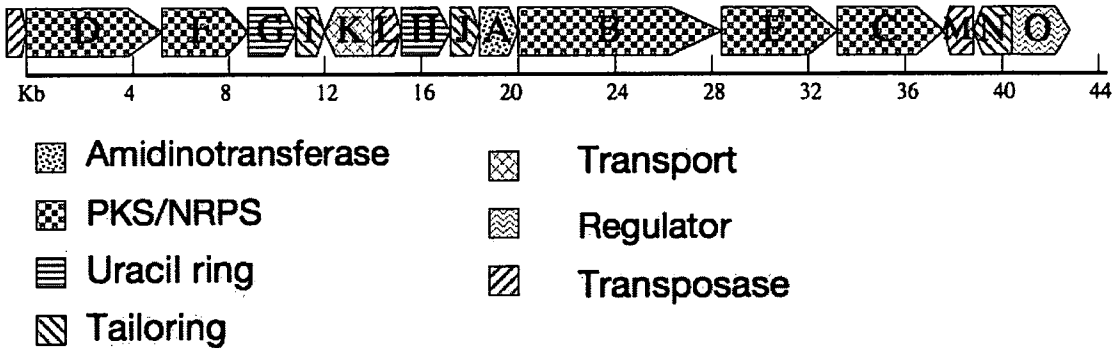
D
F
G
I
K
L
H
J
A
B
E
C
M
N
O
Kb
4
8
12
16
20
24
28
32
36
40
44
Amidinotransferase
PKS/NRPS
Uracil ring
Tailoring
Transport
Regulator
Transposase

FIGURE 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2008/001805

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

C12Q 1/68 (2006.01) *C12N 9/04* (2006.01) *C12N 9/20* (2006.01)
C12N 9/02 (2006.01) *C12N 9/10* (2006.01) *C12N 9/78* (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Genomequest : SEQ ID NO: 2-69,

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X	KELLMANN, R. et al. Biosynthetic intermediate analysis and functional homology reveal a saxitoxin gene cluster in cyanobacteria. Appl Environ Microbiol. 2008 Jul;74(13):4044-53. Epub May 16 2008. Whole document.	1-33

[X] Further documents are listed in the continuation of Box C [X] See patent family annex

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search
4 March 2009

Date of mailing of the international search report
20 MAR 2009

Name and mailing address of the ISA/AU
AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
E-mail address: pct@ipaustralia.gov.au
Facsimile No. +61 2 6283 7999

Authorized officer
SOPHINA CALANNI
AUSTRALIAN PATENT OFFICE
(ISO 9001 Quality Certified Service)
Telephone No : +61 2 6283 2038

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	WO 2007/087815 A2 (METANOMICS GMBH) 9 August 2007 SEQ ID Nos: 8331, 10767 and 11456 display 75% identity with SEQ ID NO 64 of the present application. SEQ ID NOs: 8332, 10768 and 11457 display 73% identity with SEQ ID NO: 65 of the present application.	
X	Abstract; Claims 6-12, 15, 17. WO 2006/092449 A2 (METANOMICS GMBH) 8 September 2006 SEQ ID NO: 10, 028 displays 74% identity with SEQ ID NO: 66 of the present application.	1-7, 21, 30-32
X	Abstract; Claims 11, 12, 16-20, 23, 25. US 20070020625 A1 (DUSCHAND, E. et al.) 25 January 2007 SEQ ID NO: 8723 displays 67% identity with SEQ ID NO: 20 of the present application.	1-7, 21, 30-32
X	Claims 58-67, 78, 86-88, 93 and 94. US 7314974 B2 (CAO, Y. et al.) 1 January 2008 SEQ ID NO: 43612 displays 71% identity with SEQ ID NO: 42 of the present application.	1-7, 21
X	Table 13A; Column 6 lines 5-11; Column 11 lines 11-22. Genbank Accession # BAB76734 (alr5035) 21 December 2007	1-7
X	Sequence displays 69% identity with SEQ ID NO: 3 of the present application. Genbank Accession # BAB76200 (phosphate regulon transcriptional regulator) 21 December 2007.	4, 21, 30-32
X	Sequence displays 78% identity with SEQ ID NO: 61 of the present application. Genbank Accession # BAB76202 (two-component response regulator) 21 December 2007	4, 21, 30-32
X	Sequence displays 73% identity with SEQ ID NO: 65 of the present application.	4, 21, 30-32

C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	Genbank Accession # BAB76205 (phosphorybosilformimino-5-amino-phosphorybosil-4-imidazolecarboxamideisomerase)	
X	Sequence displays 81% identity with SEQ ID NO: 67 of the present application.	4, 21, 30-32
	Genbank Accession # ABX60164 (adenylylsulfate kinase) 19 February 2008	
X	Sequence displays 78.5% identity with SEQ ID NO: 57 of the present application.	4, 21, 30-32
	Genbank Accession # ZP_01619109 (nodulation protein) 15 December 2006	
X	Sequence displays 78% identity with SEQ ID NO: 25 of the present application.	4, 21, 30-32
	Genbank Accession # ZP_01727402 (transposase, IS4) 7 March 2007	
X	Sequence displays68% identity with SEQ ID NO: 59 of the present application.	4, 21, 30-32
	Genbank Accession # ED287359 (Ascaris suum whole genome shotgun library (PMAJ_4 GSS)) 23 August 2006	
X	Sequence displays 65% identity with SEQ ID NO: 26 of the present application.	1-3, 5-7

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/AU2008/001805

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member					
WO	2007087815	AR	053108	AR	058785	AU	2005321630
		AU	2005339243	BR	PI0512818	BR	PI0519994
		CA	2559760	CA	2585798	EP	1765857
		EP	1974049	WO	2006069610		
WO	2006092449	AU	2006219823	CA	2598792	EP	1777296
		EP	1871883	US	2007118916		
US	20070020625	AU	2002339119	CA	2434323	EP	1379549
		WO	02094867				
US	7314974	US	2003233675	US	2008229451		

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

END OF ANNEX