

## 念珠藻 *Nostoc punctiforme* 胞外多醣生合成調控機制之預測

曾志彧<sup>1</sup> 楊萌皓<sup>1</sup> 李冠群<sup>1\*</sup>

<sup>1</sup> 國立台灣師範大學生命科學系

(收稿日期: 2017.03.20, 接受日期: 2017.05.20)

### 摘 要

藍菌念珠藻屬 (*Nostoc*) 的胞外多醣 (exopolysaccharides, EPS) 生合成相關之遺傳與生化方面的瞭解, 目前研究並不多, 為探討念珠藻屬 EPS 生合成基因可能受到那些轉錄因子調控, 進而了解 EPS 生合成與環境因子的關係, 本研究以念珠藻屬中基因體序列已被完整解碼的 *Nostoc punctiforme* ATCC 29133 為研究對象, 針對其基因體中 10 個可能參與 EPS 生合成的基因啟動子區域做序列分析, 共預測出 15 種轉錄因子結合位, 藉由與大腸桿菌 (*Escherichia coli*) 轉錄因子結合位序列的相似性比對, 預測最有可能會調控 EPS 生合成基因的轉錄因子為 ArgR、Fnr 與 LexA。最後, 利用 *E. coli* K12 之 ArgR、Fnr 與 LexA 胺基酸序列於 *N. punctiforme* 基因體中進行 BLAST 比對搜尋, 結果顯示 *N. punctiforme* 具有與 *E. coli* K12 相似度極高的轉錄因子 LexA, 但不具有相似的 ArgR 與 Fnr。因此推測念珠藻 *N. punctiforme* 的 EPS 生合成基因可能會受轉錄因子 LexA 與其相關環境因子的調控, 例如紫外線。

**關鍵詞:** 藍菌、念珠藻 *Nostoc punctiforme*、胞外多醣、轉錄因子、啟動子

### 緒 論

藍菌 (Cyanobacteria) 是一種分佈極廣的光合自營性微生物, 具有典型的原核生物特徵, 可行光合作用產生氧氣 (Whitton and Potts 2000), 有些藍菌種類會固定大氣中的氮氣以供自身利用 (Fay 1992, Bergman *et al.*, 1997)。大多數的藍菌可合成並分泌胞外聚合物, 主要以多醣成份為主, 藍菌的胞外多醣 (exopolysaccharides, EPS) 會以共價連接或鬆散附著在細胞表面, 有些 EPS 會被釋放到週遭環境中 (De Philippis and Vincenzini 2003)。藍菌的 EPS 中含量最多的單糖單元通常是葡萄糖, 不過也存在其它型式的單糖, 例如木糖、阿拉伯糖、半乳糖或岩藻糖等 (Tease *et al.*, 1991, Bender *et al.*, 1994, Gloaguen *et al.*, 1995, Fischer *et al.*, 1997, De Philippis and Vincenzini 1998, Parikh and Madamwar 2006)。EPS 的組成會隨著不同的微生物種類、營養狀況、生長階段、與環境狀況而改變 (De Vuyst and Degeest 1999, Ricciardi *et al.*, 2002, Fischer *et al.*, 2003, Bahat-Samet *et al.*, 2004, Vaningelgem *et al.*, 2004, Panhota *et al.*, 2007)。

EPS 的功能包括能幫助藍菌抵抗乾旱、紫外線、螯合金屬離子、以及保護細胞避免發生生物性礦化 (Pereira *et al.*, 2009)。藍菌 EPS 現在也已

經被廣泛應用於紡織、清潔劑、膠粘劑、採油、廢水處理、藥品以及食品添加等用途上。目前對於藍菌 EPS 生合成的相關研究, 主要針對細胞能量的可用性及養分碳氮比 (C:N ratio) (De Philippis and Vincenzini 1998, Li *et al.*, 2001) 對 EPS 生合成的影響, 而對於影響 EPS 生合成的其他環境因子以及 EPS 生合成基因可能受到那些轉錄因子調控, 則相對較少被探討。近年來已有一些論文報告開始研究其他營養源含量(例如無機鹽類)、不同光照強度、溫度、重金屬或是氧化壓力對 EPS 生合成的影響 (BI Yong-hong and Zheng-yu 2004, Monshupanee and Incharoensakdi 2014, Cassier-Chauvat and Chauvat 2015, Mota *et al.*, 2015), 此外, 對於不同光質的研究, 例如紫外光已經被研究證實會增加藍菌念珠藻屬 (*Nostoc*) 的地木耳 (*N. commune*) 及髮菜 (*N. flagelliforme*) 的 EPS 生合成量達三倍之多 (Ehling-Schulz *et al.*, 1997, Feng *et al.*, 2012)。

本實驗的目的在於利用生物資訊技術分析念珠藻 *N. punctiforme* ATCC 29133 的基因體, 找出可能會調控 EPS 生合成基因之轉錄因子, 進而推測可能影響 EPS 合成機制的環境因子。

\* 通信作者: 李冠群(Guan-Chiun Lee); FAX: 886-2-29312904; E-mail: gclee@ntnu.edu.tw

## 材料與方法

### *N. punctiforme* 的 EPS 生合成基因序列分析

Pereira 等人以生物資訊學分析 *N. punctiforme* ATCC 29133 基因體 (GenBank: CP001037.1) 內與 EPS 生合成相關的基因，預測出五種 *wz\_* 基因 (包括 *Npun\_F0449*, *wzb1*; *Npun\_R5507*, *wza1*; *Npun\_F5505*, *wzc1*; *Npun\_F5499*, *wzx1*; and *Npun\_F5501*, *wzy1*) 與三種 glycosyl transferases 基因 (*Npun\_F5500*, *Npun\_F5502*, and *Npun\_F5504*) 是 EPS 合成的重要基因 (Pereira *et al.*, 2009)。另外，我們也使用 NCBI Nucleotide BLAST 比對搜尋 *N. punctiforme* ATCC 29133 基因體中與上述 8 種基因相似的其他基因，找出可能與 EPS 生合成相關的基因，以便進行啟動子區域中轉錄因子結合位置的分析預測。

### EPS 生合成基因啟動子內轉錄因子結合位預測

使用原核生物啟動子預測程式 Bacterial promoter prediction program, BPROM (<http://linux1.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb>) 來預測啟動子區域內可能存在的轉錄因子結合位。我們將上述 EPS 相關基因的啟動子區域之序列，輸入至 BPROM 的網站中分析，並使用網站預設參數 (預設的資料庫為 DP Interact database，其中涵蓋 55 種已知 *E. coli* K12 之轉錄因子資訊)，該網站分析的結果會包含預測的轉錄因子結合位序列 (以 8 個鹼基序列長度呈現)，以及轉錄因子名稱。

### 調控 EPS 生合成的轉錄因子之預測

利用蛋白質資料庫網站 Uniprot ([www.uniprot.org/](http://www.uniprot.org/))，對預測到的轉錄因子進行身份與功能的確認，輸入轉錄因子的名稱，並選定搜尋菌種為 *E. coli* K12，獲得轉錄因子的詳細資訊。再使用原核生物轉錄因子結合位序列資料庫 PRODORIC (<http://www.prodoric.de/>)，輸入轉錄因子的名稱，利用網站預設的搜尋參數，並選定菌種為 *E. coli* K12，獲得完整的結合位序列，以及轉錄因子會調控哪些下游基因資訊。另外，將上述找出的 *E. coli* K12 轉錄因子結合位序列，與 BPROM 預測的 *N. punctiforme* 轉錄因子結合位序列，利用 ClustalW Multiple alignment 程式進行序列相似性比對。另外，統計所預測到的轉錄因

子在 10 個 EPS 生合成基因啟動子上出現的次數，找出在基因的啟動子中出現次數較多者，並配合序列相似性比對之結果，以篩選出最有可能會調控 *N. punctiforme* EPS 生合成基因的轉錄因子。

為了進一步確認預測的轉錄因子是否存在 *N. punctiforme*，在 Uniprot 資料庫中，輸入上述較可能會調控 EPS 生合成的轉錄因子名稱，找出 *E. coli* K12 的轉錄因子蛋白質序列，將該蛋白質序列輸入到 NCBI Protein BLAST 作序列搜尋比對，並將序列比對限定於 *N. punctiforme* 基因體，以確定 *N. punctiforme* 是否有該轉錄因子。

## 結果

### *N. punctiforme* 的 EPS 生合成基因

除了 Pereira 等人預測的 8 個 EPS 生合成相關基因外，我們使用 NCBI Nucleotide BLAST 在 *N. punctiforme* ATCC 29133 基因體上搜尋與此 8 個基因序列相似的其他基因，結果額外發現兩個基因 (*Npun\_F1381*, *wzc2* 與 *Npun\_R1496*, *wzc3*) 與 *Npun\_F5505* 有 66% 的相似程度，推測可能也與 EPS 合成相關。自 *N. punctiforme* 基因體選取這 10 個基因的啟動子區域序列 (Table 1)，以便進行轉錄因子結合位置的分析。然而，其中 9 個基因的啟動子區域有預測到 Pribnow box 序列，可以視為具有功能的啟動子，但 *Npun\_F0449* (*wzb1*) 的啟動子區域分析結果並無發現到 Pribnow box，然而在其上游的 *Npun\_F0448* 基因啟動子區域確實發現到 Pribnow box 序列，所以我們認為其可能與前一個基因形成一操縱子 (operon)，因此，取 *Npun\_F0448* 的啟動子來代表 *Npun\_F0449* (*wzb1*) 的啟動子作分析。

這些基因的功能說明如下：*Wza*，將藉由 *Wzc* 運送的多醣片段，從細胞週質空間 (periplasmic space) 轉移到細胞外膜的表面上；*Wzb*，主要對 *Wzc* 進行磷酸轉移作用與去磷酸化作用，這兩者是對於調控多醣合成與輸出到細胞外是必需的作用；*Wzc*，幫助多醣片段從細胞膜轉到細胞週質空間；*Wzx*，把合成在脂質載體上的多醣片段，從細胞膜內側轉向到細胞膜外側；*Wzy*，將細胞膜外側的多醣片段轉移到 *Wzc* 上；glycosyl transferase, group 1，主要是把細胞質中合成的多醣片段，轉移到細胞膜上的脂質載體。

表一、*N. punctiforme* ATCC29133 胞外多醣生合成相關基因。

**Table 1.** EPS biosynthesis-related genes of *N. punctiforme* ATCC29133

Gene locus name <sup>1</sup>	Protein	Location in chromosome	Gene length(bp)	Length of putative promoter (bp) <sup>2</sup>
Npun_F0449	Wzb1	549061-549570	510	198
Npun_F1381	Wzc2	1683881-1685983	2103	213
Npun_R1496	Wzc3	1836227-1838431	2205	1262
Npun_F5499	Wzx1	6795780-6797207	1428	395
Npun_F5500	glycosyl transferase	6797288-6798388	1101	80
Npun_F5501	Wzy1	6798542-6799762	1221	153
Npun_F5502	glycosyl transferase	6800105-6801205	1101	342
Npun_F5504	glycosyl transferase	6802969-6804156	1188	422
Npun_F5505	Wzc1	6804967-6807213	2247	810
Npun_R5507	Wza1	6809850-6811319	1470	536

**1:** Npun is the abbreviation of *N. punctiforme*; F means that the gene is on the forward DNA strand and R is on the reverse strand.

**2:** The putative promoter length is obtained by measuring the length between the ORF start of the later gene and the ORF end of the former gene.

### EPS 生合成基因啟動子中轉錄因子結合位預測

BPROM 分析 10 個 EPS 生合成基因的啟動子區域，共預測到 15 種轉錄因子的結合位 (Table 2)，這些轉錄因子結合位出現在各基因啟動子中的頻率如 Table 3 所示，發現有些轉錄因子只出現於一種啟動子中，例如 CynR、Fis、HipB、NagC、PhoB3、PhoB 與 PurR；有些轉錄因子則同時出現於二至四種啟動子中，例如 Fnr 出現於四種啟動子內，而 ArgR 與 LexA 皆出現於三種啟動子中。另外，在同一基因的啟動子中可能同時存在多種轉錄因子，例如 Npun\_R5507 (Wza<sub>1</sub>) 具有七種轉錄因子結合位 (Fig. 1)，而 Npun\_F1381 (Wzc<sub>2</sub>) 與 Npun\_F5502 (glycosyl transferase) 則具有四種轉錄因子結合位，顯示 *N. punctiforme* EPS 的合成可能受到多種環境因子的調控，而且不同的環境因子可以藉由特殊轉錄因子調控單一或多種 EPS 生合成基因。

### 調控 EPS 生合成的轉錄因子之預測

由於 BPROM 分析啟動子的結果僅能獲得轉錄因子簡寫名稱，以及轉錄因子可能的 8 個鹼基序列結合位，然而 BPROM 的分析依據是以已知的 *E. coli* K12 轉錄因子結合位序列，到查詢啟動子區域中比對搜尋，為了要篩選出較有可能會調控 EPS 生合成的轉錄因子，必須比對 *N. punctiforme* 與 *E. coli* K12 的啟動子結合位的序列，亦即與 *E. coli* K12 的序列相似性愈高，代表該轉錄因子愈可能會結合在 *N. punctiforme* EPS

生合成基因的啟動子上。我們利用 PRODORIC 分析軟體獲得這些預測的轉錄因子在 *E. coli* K12 所控制的下游基因，並且獲得下游基因啟動子區域上轉錄因子的完整結合位序列，再利用 ClustalW Multiple Alignment 進行與 *N. punctiforme* EPS 基因啟動子之結合位序列的相似度比對，結果顯示相似度達 60% 以上的轉錄因子結合位有 ArgR、Fnr、CynR、PurR、Crp 與 LexA，且序列相似度較高的前三種轉錄因子分別為 ArgR (72% identity)、LexA (70% identity) 與 Fnr (64% identity) (Table 4)。另外，由 Table 3 的結果亦可以看出 ArgR、Fnr、LexA 與 Lrp 在 10 個 EPS 基因的啟動子序列上出現的次數 (含重覆出現) 皆超過 3 次以上。所以，我們認為 ArgR、Fnr 與 LexA 三種轉錄因子是較有可能參與 *N. punctiforme* EPS 生合成的調控。

為了進一步確認預測的轉錄因子是否存在 *N. punctiforme*，我們分別以 *E. coli* K12 的 ArgR (UniParc : P0A6D0-1)、Fnr (UniParc : P0A9E5-1) 以及 LexA (UniParc : P0A7C2-1) 之蛋白質序列作 Protein BLAST (NCBI) 比對搜尋，並選定菌種 *N. punctiforme* 為搜尋對象。結果顯示，ArgR 相似蛋白質的比對分數都是在 30 以下，而 Fnr 相似蛋白質的最高分數是 58.9，然而 LexA 相似蛋白質的最高分數則為 102，而且其序列覆蓋率高達 97%，序列一致性有 34%，並且該蛋白質資料中指出該蛋白質為 SOS-response transcriptional regulator (NCBI Reference Sequence: WP\_012410809.1)，因此我們認為 *N. punctiforme*

可能不具有與 ArgR 及 Fnr 相似的轉錄因子，但具有轉錄因子 LexA，而且根據 Table 2，LexA 可能會藉由調控 Npun\_F5502 (glycosyl transferase)。

Npun\_F5504 (glycosyl transferase)與 Npun\_R5507 (Wza<sub>1</sub>)，而達到調控 *N. punctiforme* EPS 的生成。

表二、在 10 個 *N. punctiforme* ATCC29133 胞外多醣生成基因的啟動子區域中預測到的轉錄因子結合位。

**Table 2.** The predicted transcription factor binding sites in 10 promoter regions of the EPS biosynthesis-related genes of *N. punctiforme* ATCC 29133

Transcription factor	Sequence of transcription factor binding site predicted by BPROM1	EPS-related gene whose promoter contains the binding site	Location in promoter
ArgR	AATAATTA	Npun_F5501	-119 to -112
	CATAAAAA		-89 to -82
	ATAAAAAT	Npun_F5504	-88 to -81
	AATTAATA		-72 to -65
	TTTTTTAT	Npun_R5507	-243 to -236
ArgR2	ATATAAAT	Npun_F5499	-191 to -184
	CATATTTT	Npun_R5507	-208 to -201
Crp	AGATCACA		-65 to -58
	ATCACAAA	Npun_F1381	-63 to -56
Fis	AAAAATAA	Npun_F1381	-11 to -4
Fnr	ATAAATAT	Npun_F5499	-189 to -182
	TCAAGAGT	Npun_F5504	-132 to -125
	CCTACCTC	Npun_R1496	-575 to -568
	TTTTTTGA	Npun_R5507	-204 to -197
HipB	CCTTTTTA	Npun_R5507	-263 to -256
Ihf	AGATATAT	Npun_F5505	-290 to -283
	ACAGACAA	Npun_R5507	-251 to -244
LexA	ATAGATAA	Npun_F5502	-306 to -299
	TATATATA	Npun_F5504	-145 to -138
	TATATAAA	Npun_F5504	-143 to -136
		Npun_R5507	-238 to -231
		Npun_F5504	-142 to -135
Lrp	TATTCTTA	Npun_F1381	-83 to -76
	ATTTATTA	Npun_F5502	-271 to -264
	TATTTTTT	Npun_R5507	-226 to -219
NagC	TTTAATTT	Npun_F1381	-55 to -48
PhoB3	TCCTTACA	Npun_R1496	-562 to -555
PhoB	TAATATAT	Npun_F5502	-301 to -294
	TTTATTTAA		-270 to -263
PurR	TTTTTCGTT	Npun_F5505	-351 to -344
	TTTCGTTT		-350 to -343
TyrR		Npun_F5502,	-300 to -293
	AATATATA	Npun_F5504	-99 to -92
CynR	ATAAGTAA	Npun_F5500	-21 - -14

**1:** BPROM is a tool for predicting of bacterial promoters and the average length of predicted TF binding site is eight nucleotides.

表三、BPRM 所預測的轉錄因子結合位在 *N. punctiforme* ATCC29133 胞外多醣生合成基因啟動子區域中出現的頻率。

**Table 3.** Appearance frequency of the BPRM-predicted transcription factor binding sites in the promoters of the *N. punctiforme* EPS biosynthesis-related genes.

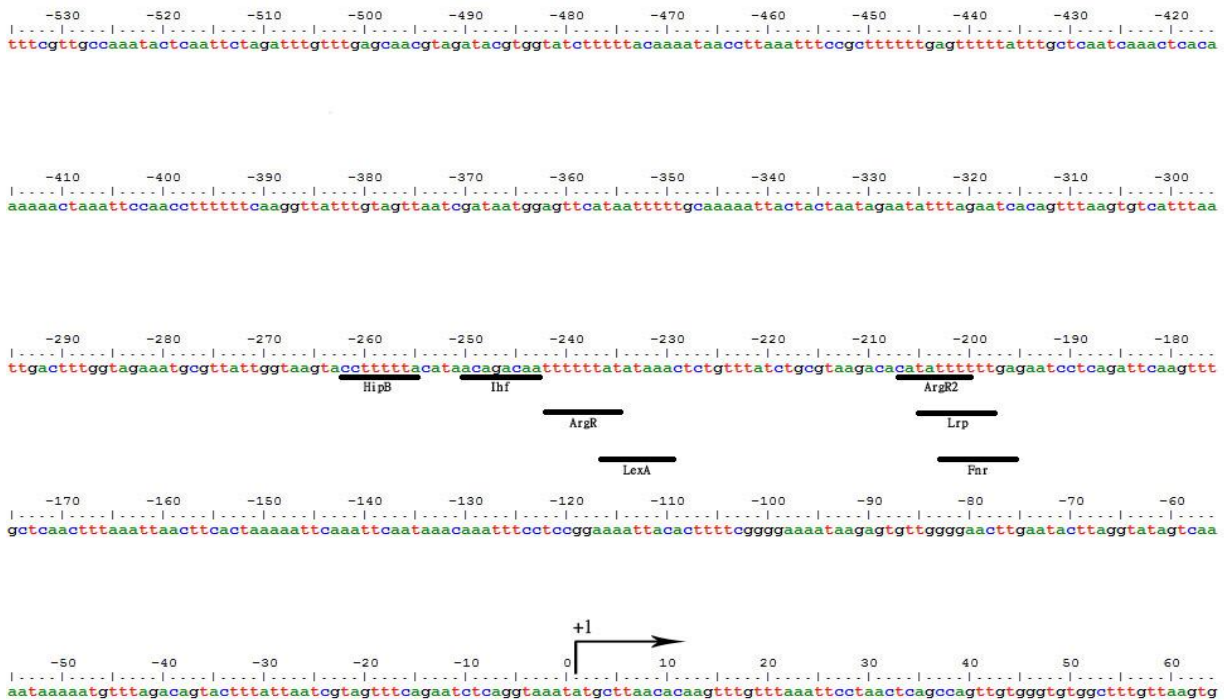
	F0449 (F0448)1	F1381	R1496	F5499	F5500	F5501	F5502	F5504	F5505	R5507	Total
CynR					1						1
ArgR						1		2		1	4
ArgR2				1						1	2
Crp	1	1									2
Fis		1									1
Fnr			1	1				1		1	4
HipB										1	1
Ihf									1	1	2
LexA							1	1		1	3
Lrp		1					1			1	3
NagC		1									1
PhoB3			1								1
PhoB							2				2
PurR									1		1
TyrR							1	1			2
Total	1	4	2	2	1	1	5(4) 2	5(4)	2	7	

1 : Because Npun\_F0448 and Npun\_F0449 are within the same operon and under the control of Npun\_F0448 promoter, we analyze the promoter of Npun\_F0448 instead of Npun\_F0449 promoter.

2 : The parenthesis indicates the numbers of various TF factors.

圖一、Npun\_R5507 基因啟動子內轉錄因子結合位分佈圖。+1 標示 Npun\_R5507 基因編碼起始位置。

**Fig. 1:** Locations of the predicted transcription factor binding sites in the promoter region of Npun\_R5507 gene. +1 indicates the translation start site of Npun\_R5507 gene.



表四、*E. coli* K12 與 *N. punctiforme* 的轉錄因子結合位序列相似度比對。

Table 4. Comparison of transcription factor binding site between *E. coli* K-12 and *N. punctiforme*

Transcription factor	Function of transcription factor(Uniprot) <sup>1</sup>	Sequence of transcription factor binding site predicted by BPROM in the promoter of <i>N. punctiforme</i> EPS-related genes	Sequence of transcription factor binding site in the promoter of the downstream genes (downstream genes in <i>E. coli</i> K-12 from PRODORIC) <sup>2</sup>	Identity (bold letters/total seq. letters)	Reference
ArgR	Arginine repressor	AATAATTA	AATGAATAAATACACATA <sup>3</sup> (argF operon)	72.22%	(Charlier <i>et al.</i> , 1992) (Tian <i>et al.</i> , 1992) (Maas 1994)
		CATAAAAA	TTTGCATAAAAAATTCATC (argR operon)	61.11%	(Lim <i>et al.</i> , 1987) (Charlier <i>et al.</i> , 1992). (Maas 1994)
			TTTGCATAAAAAATTCAGT (argG operon)	61.11%	
		ATAAAAAT	TTTGCATAAAAAATTCATC (argR operon)	55.55%	(Lim <i>et al.</i> , 1987) (Charlier <i>et al.</i> , 1992) (Maas 1994)
			TATGAATAAAAAATACACT (argECBH operon)	66.66%	
			ACAGAATAAAAAATACACT (argA operon)	55.55%	
			TTTGCATAAAAAATTCAGT (argG operon)	61.11%	
		AATTAATA	TGTGAATAAATATGCAAA (carAB operon)	55.55%	(Bouvier <i>et al.</i> , 1984) (Charlier <i>et al.</i> , 1992) (Charlier <i>et al.</i> , 1995).
		TTTTTAT	AGTGATTTTTATGCATA (argD operon)	61.11%	(Charlier <i>et al.</i> , 1992) (Maas 1994)
Fnr	Fumarate and nitrate reduction regulatory protein	ATAAATAT	ATGATAAATATCAA (narK operon)	64.28%	(Li <i>et al.</i> , 1994)
		TCAAGAGT	GTTTGATATTATCATATT (hlyE operon)	31.57%	(Westenmark <i>et al.</i> , 2000)
			TTGAGATACATCAA (hmp operon)	42.85%	(Membrillo-Hernandez <i>et al.</i> , 1997)
		CCTACCTC	TCGATCTCGTCAA (fndA operon)	46.15%	(Li <i>et al.</i> , 1994)
		TTTTTTGA	TTGATTTTTATGTC (viaKLMNOlyxKsgbHUEoperon)	42.85%	(Ibanez <i>et al.</i> , 2000)
LexA	LexA repressor	ATAGATAA	CACTGGATAGATAACCGCA (yjdQ operon)	60.00%	(Fernandez De Henestrosa <i>et al.</i> , 2000)
		TATATATA	TACTGTCTGTATATATAAGT (yjnN operon)	55.00%	(Fernandez De Henestrosa <i>et al.</i> , 2000) (Easton and Kushner 1983)
			CTGTATATATACCCAG (uvrD operon)	50.00%	
		TATATAAA (F5504)	ACAGACTACTGTATATAAAACAGT ATAACTTCA (umuDC operon)	47.05%	(Kitagawa <i>et al.</i> , 1985) (Rostas <i>et al.</i> , 1987) (Ohmori <i>et al.</i> , 1995) (Fernandez De Henestrosa <i>et al.</i> , 2000)
			TACTGTATATAAAACAGTT (recN operon)	60.00%	
			AACTGTATATAATACAGTT (dinD operon)	55.00%	
			TACTGTATATAAAACAGTA (sbmC operon)	60.00%	
		TATATAAA (R5507)	ACAGACTACTGTATATAAAACAGT ATAACTTCA (umuDC operon)	61.76%	
			TACTGTATATAAAACAGTT (recN operon)	70.00%	(Kitagawa <i>et al.</i> , 1985, Rostas <i>et al.</i> , 1987, Ohmori <i>et al.</i> , 1995, Fernandez De Henestrosa <i>et al.</i> , 2000)
			AACTGTATATAATACAGTT (dinD operon)	65.00%	
			TACTGTATATAAAACAGTA (sbmC operon)	65.00%	
		ATATAAAA	ACAGACTACTGTATATAAAACAGT ATAACTTCA (umuDC operon)	47.05%	(Kitagawa <i>et al.</i> , 1985) (Rostas <i>et al.</i> , 1987) (Fernandez De Henestrosa <i>et al.</i> , 2000)
			TACTGTATATAAAACAGTT (recN operon)	60.00%	
			TACTGTATATAAAACAGTA (sbmC operon)	60.00%	

- 1: These protein names are obtained from Uniprot (<http://www.uniprot.org/>). It is the most comprehensive catalog of information on proteins and the central repository of protein sequence and function.
- 2: The information of *E. coli* K-12 operons is obtained from PRODORIC (<http://www.prodoric.de/index.php>), which is a comprehensive database about gene regulation and gene expression in prokaryotes.
- 3: The bold letters indicates the same nucleotides which appear in both the binding sites of the *E. coli* K-12 downstream genes and *N. punctiforme* EPS-related genes. The *Italics* letters indicates the binding site sequences which are conserved with the predicted sequences from BPROM as shown in Table 2.

## 討論

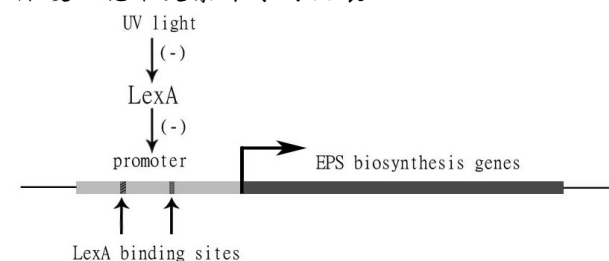
根據本研究的生物資訊分析結果，ArgR、Fnr、LexA 這三種轉錄因子有可能對 EPS 生合成基因進行調控。目前關於這三種轉錄因子之基因所受到的調控方式，以及它們對下游基因的調控方式，大致有以下的研究，ArgR 為一種 autorepressor，由 ArgR gene 表達產生，而該基因的啟動子區域具有 upstream ArgR binding site (ARG BOX)，因此該基因會受到自身基因的產物調控，並也會受到 L-arginine 的負調控，而 ArgR 對其下游基因的調控方式，必需先與 L-arginine 進行結合，形成有活性的 repressor，才能抑制與 arginine 生合成相關的基因 (Maas 1994, Van Duyne *et al.*, 1996)。

Fnr 會受氧氣的濃度調控，其活性與蛋白質構型有關，蛋白質構型分為兩種，即具有活性的還原態 4Fe-FNR 與不具活性的氧化態 2Fe-FNR，Fnr gene 會在無氧的環境下被 4Fe-FNR 進行自我負調控，並且會在缺氧的情況下，正調控與無氧呼吸相關的基因 (例如 aspartase, formate hydrogenase, fumarate reductase, and pyruvate formatylase genes) 或是負調控與有氧呼吸相關的下流基因 (Unden and Trageser 1991, Jervis and Green 2007, Mettert and Kiley 2007)。

已知 LexA 為細菌啟動 SOS 機制修補 DNA 的因子，LexA 為一種 repressor，平時會結合在 SOS genes 的操作子區域上，但在細胞的 DNA 受到輻射線破壞時，例如紫外線 (UV light) 照射等，造成雙股 DNA 斷裂，或是形成單股的 DNA，此 DNA 受損的訊息會先活化 RecA，RecA 對 LexA 進行水解切割而使得 LexA 失去活性，藉此啟動 SOS 基因，進行一連串 DNA 修復過程 (Walker 1984, Belov *et al.*, 2009)。

根據本研究結果推測，最有可能會調控 *N. punctiforme* EPS 生合成的轉錄因子為 LexA，其它兩種轉錄因子 ArgR 與 Fnr 則不確定。因此推測，當 *N. punctiforme* 受輻射與 UV 的破壞時，其 LexA 不但會參與 DNA 修補機制，也可能會藉由調控 EPS 的生合成來保護細胞。目前已經發現，

在藍菌 EPS 構成的鞘 (sheath) 中含有能夠吸收 UV 的物質，例如 scytonemin 或是 mycosporine-like amino acid compounds (Scherer *et al.*, 1988, Ehling-Schulz *et al.*, 1997)，可以形成避免有害輻射穿透進入細胞的保護屏障。過去有些論文證明紫外線 UV-B 會促進同屬的念珠藻，如地木耳 (*N. commune*) 及髮菜 (*N. flagelliforme*) 的 EPS 生合成達三倍之多，同時，也促進吸收 UV 物質的合成 (Ehling-Schulz *et al.*, 1997, Feng *et al.*, 2012)。我們根據 *N. punctiforme* ATCC 29133 (GenBank: CP001037.1)、*N. commune* AHNG0605 (GenBank: EU178142.1) 與 *N. flagelliforme* IMGA0408 (GenBank: EU178143.1) 的 16S ribosomal RNA 基因序列，進行親緣關係分析，結果顯示地木耳、髮菜與 *N. punctiforme* 的親緣關係極為相近。因此推測 UV-B 可能會藉由轉錄因子 LexA 來調控念珠藻 *N. punctiforme* 的 EPS 生合成基因 (Fig. 2)，促進 EPS 的生合成，以保護細胞不受紫外線的破壞。



圖二、念珠藻 EPS 生合成基因的可能調控方式。當念珠藻受 UV light 傷害時，EPS 生合成相關基因會被活化而大量表現。

Fig. 2: The proposed regulation mechanism of *Nostoc* EPS biosynthesis-related genes.

## 參考文獻

- Bahat-Samet E, Castro-Sowinski S, Okon Y. 2004. Arabinose content of extracellular polysaccharide plays a role in cell aggregation of *Azospirillum brasilense*. *FEMS Microbiol Lett.* **237**(2): 195-203.
- Belov OV, Krasavin EA, Parkhomenko AY. 2009. Model of SOS-induced mutagenesis in bacteria

- Escherichia coli under ultraviolet irradiation. *J Theor Biol.* **261**(3): 388-395.
- Bender J, Rodriguez-Eaton S, Ekanemesang UM, Phillips P. 1994. Characterization of metal-binding bioflocculants produced by the cyanobacterial component of mixed microbial mats. *Appl Environ Microbiol.* **60**(7): 2311-2315.
- Bergman B, Gallon JR, Rai AN, Stal LJ. 1997. N<sub>2</sub>-fixation by non-heterocystous cyanobacteria. *FEMS Microbiol Rev.* **19**: 139-185.
- BI Yong-hong, Zheng-yu H. 2004. Influence of Temperature, Nutrients and Light Intensity on the Growth of *Nostoc flagelliforme*. *The Chinese Journal of Process Engineering.* **4**(3): 245-249.
- Bouvier J, Patte JC, Stragier P. 1984. Multiple regulatory signals in the control region of the *Escherichia coli* carAB operon. *Proc Natl Acad Sci U S A.* **81**(13): 4139-4143.
- Cassier-Chauvat C, Chauvat F. 2015. Responses to oxidative and heavy metal stresses in cyanobacteria: recent advances. *Int J Mol Sci.* **16**(1): 871-886.
- Charlier D, Gigot D, Huysveld N, Roovers M, Pierard A, Glansdorff N. 1995. Pyrimidine regulation of the *Escherichia coli* and *Salmonella typhimurium* carAB operons: CarP and integration host factor (IHF) modulate the methylation status of a GATC site present in the control region. *J Mol Biol.* **250**(4): 383-391.
- Charlier D, Roovers M, Van Vliet F, Boyen A, Cunin R, Nakamura Y, Glansdorff N, Pierard A. 1992. Arginine regulon of *Escherichia coli* K-12. A study of repressor-operator interactions and of in vitro binding affinities versus in vivo repression. *J Mol Biol.* **226**(2): 367-386.
- De Philippis R, Vincenzini M. 1998. Exocellular polysaccharides from cyanobacteria and their possible applications. *FEMS Microbiol Rev.* **22**: 151-175.
- De Philippis R, Vincenzini M. 2003. Outermost polysaccharidic investments of cyanobacteria: nature, significance and possible applications. *Recent Res Dev Microbiol.* **7**: 13-22.
- De Vuyst L, Degeest B. 1999. Heteropolysaccharides from lactic acid bacteria. *FEMS Microbiol Rev.* **23**(2): 153-177.
- Easton AM, Kushner SR. 1983. Transcription of the uvrD gene of *Escherichia coli* is controlled by the lexA repressor and by attenuation. *Nucleic Acids Res.* **11**(24): 8625-8640.
- Ehling-Schulz M, Bilger W, Scherer S. 1997. UV-B-induced synthesis of photoprotective pigments and extracellular polysaccharides in the terrestrial cyanobacterium *Nostoc commune*. *J Bacteriol.* **179**(6): 1940-1945.
- Fay P. 1992. Oxygen relations of nitrogen fixation in cyanobacteria. *Microbiol Rev.* **56**(2): 340-373.
- Feng YN, Zhang ZC, Feng JL, Qiu BS. 2012. Effects of UV-B radiation and periodic desiccation on the morphogenesis of the edible terrestrial cyanobacterium *Nostoc flagelliforme*. *Appl Environ Microbiol.* **78**(19): 7075-7081.
- Fernandez De Henestrosa AR, Ogi T, Aoyagi S, Chafin D, Hayes JJ, Ohmori H, Woodgate R. 2000. Identification of additional genes belonging to the LexA regulon in *Escherichia coli*. *Mol Microbiol.* **35**(6): 1560-1572.
- Fischer D, Schlösser UG, Pohl P. 1997. Exopolysaccharide production by cyanobacteria grown in closed photobioreactors and immobilized using white cotton towelling. *J Appl Phycol.* **9**: 205-213.
- Fischer SE, Miguel MJ, Mori GB. 2003. Effect of root exudates on the exopolysaccharide composition and the lipopolysaccharide profile of *Azospirillum brasilense* Cd under saline stress. *FEMS Microbiol Lett.* **219**(1): 53-62.
- Gloaguen V, Morvan H, Hoffmann L. 1995. Released and capsular polysaccharides of Oscillatoriaceae (Cyanophyceae, Cyanobacteria). *Alg Studies.* **78**: 53-69.
- Ibanez E, Campos E, Baldoma L, Aguilar J, Badia J. 2000. Regulation of expression of the yiaKLMNOPQRS operon for carbohydrate utilization in *Escherichia coli*: involvement of the main transcriptional factors. *J Bacteriol.* **182**(16): 4617-4624.
- Jervis AJ, Green J. 2007. In vivo demonstration of FNR dimers in response to lower O<sub>2</sub> availability. *J Bacteriol.* **189**(7): 2930-2932.
- Kitagawa Y, Akaboshi E, Shinagawa H, Horii T, Ogawa H, Kato T. 1985. Structural analysis of the umu operon required for inducible mutagenesis in *Escherichia coli*. *Proc Natl Acad Sci U S A.* **82**(13): 4336-4340.
- Li J, Kustu S, Stewart V. 1994. In vitro interaction of nitrate-responsive regulatory protein NarL with DNA target sequences in the fdnG, narG, narK and frdA operon control regions of *Escherichia coli* K-12. *J Mol Biol.* **241**(2): 150-165.



- Li P, Harding SE, Liu Z. 2001. Cyanobacterial exopolysaccharides: their nature and potential biotechnological applications. *Biotechnol Genet Eng Rev.* **18**: 375-404.
- Lim DB, Oppenheim JD, Eckhardt T, Maas WK. 1987. Nucleotide sequence of the *argR* gene of *Escherichia coli* K-12 and isolation of its product, the arginine repressor. *Proc Natl Acad Sci U S A.* **84**(19): 6697-6701.
- Maas WK. 1994. The arginine repressor of *Escherichia coli*. *Microbiol Rev.* **58**(4): 631-640.
- Membrillo-Hernandez J, Cook GM, Poole RK. 1997. Roles of RpoS ( $\sigma$ S), IHF and ppGpp in the expression of the *hmp* gene encoding the flavohemoglobin (Hmp) of *Escherichia coli* K-12. *Mol Gen Genet.* **254**(5): 599-603.
- Mettert EL, Kiley PJ. 2007. Contributions of [4Fe-4S]-FNR and integration host factor to *fnr* transcriptional regulation. *J Bacteriol.* **189**(8): 3036-3043.
- Monshupanee T, Incharoensakdi A. 2014. Enhanced accumulation of glycogen, lipids and polyhydroxybutyrate under optimal nutrients and light intensities in the cyanobacterium *Synechocystis* sp. PCC 6803. *J Appl Microbiol.* **116**(4): 830-838.
- Mota R, Pereira SB, Meazzini M, Fernandes R, Santos A, Evans CA, De Philippis R, Wright PC, Tamagnini P. 2015. Effects of heavy metals on *Cyanothece* sp. CCY 0110 growth, extracellular polymeric substances (EPS) production, ultrastructure and protein profiles. *J Proteomics.* **120**: 75-94.
- Ohmori H, Saito M, Yasuda T, Nagata T, Fujii T, Wachi M, Nagai K. 1995. The *pcsA* gene is identical to *dinD* in *Escherichia coli*. *J Bacteriol.* **177**(1): 156-165.
- Panhota RS, Bianchini I, Vieira AAH. 2007. Glucose uptake and extracellular polysaccharides (EPS) produced by bacterioplankton from an eutrophic tropical reservoir (Barra Bonita, SP-Brazil). *Hydrobiologia.* **583**: 223-230.
- Parikh A, Madamwar D. 2006. Partial characterization of extracellular polysaccharides from cyanobacteria. *Bioresour Technol.* **97**(15): 1822-1827.
- Pereira S, Zille A, Micheletti E, Moradas-Ferreira P, De Philippis R, Tamagnini P. 2009. Complexity of cyanobacterial exopolysaccharides: composition, structures, inducing factors and putative genes involved in their biosynthesis and assembly. *FEMS Microbiol Rev.* **33**(5): 917-941.
- Ricciardi A, Parente E, Crudele MA, Zanetti F, Scolari G, Mannazzu I. 2002. Exopolysaccharide production by *Streptococcus thermophilus* SY: production and preliminary characterization of the polymer. *J Appl Microbiol.* **92**(2): 297-306.
- Rostas K, Morton SJ, Picksley SM, Lloyd RG. 1987. Nucleotide sequence and LexA regulation of the *Escherichia coli* *recN* gene. *Nucleic Acids Res.* **15**(13): 5041-5049.
- Scherer S, Chen TW, Boger P. 1988. A New UV-A/B Protecting Pigment in the Terrestrial Cyanobacterium *Nostoc commune*. *Plant Physiol.* **88**(4): 1055-1057.
- Tease B, Jurgens UJ, Golecki JR, Heinrich UR, Rippka R, Weckesser J. 1991. Fine-structural and chemical analyses on inner and outer sheath of the cyanobacterium *Gloeotheca* sp. PCC 6909. *Antonie Van Leeuwenhoek.* **59**(1): 27-34.
- Tian G, Lim D, Carey J, Maas WK. 1992. Binding of the arginine repressor of *Escherichia coli* K12 to its operator sites. *J Mol Biol.* **226**(2): 387-397.
- Uden G, Trageser M. 1991. Oxygen regulated gene expression in *Escherichia coli*: control of anaerobic respiration by the FNR protein. *Antonie Van Leeuwenhoek.* **59**(2): 65-76.
- Van Duyne GD, Ghosh G, Maas WK, Sigler PB. 1996. Structure of the oligomerization and L-arginine binding domain of the arginine repressor of *Escherichia coli*. *J Mol Biol.* **256**(2): 377-391.
- Vaningelgem F, Zamfir M, Mozzi F, Adriany T, Vancanneyt M, Swings J, De Vuyst L. 2004. Biodiversity of exopolysaccharides produced by *Streptococcus thermophilus* strains is reflected in their production and their molecular and functional characteristics. *Appl Environ Microbiol.* **70**(2): 900-912.
- Walker GC. 1984. Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. *Microbiol Rev.* **48**(1): 60-93.
- Westermarck M, Oscarsson J, Mizunoe Y, Urbonaviciene J, Uhlin BE. 2000. Silencing and activation of *ClyA* cytotoxin expression in *Escherichia coli*. *J Bacteriol.* **182**(22): 6347-6357.
- Whitton BA, Potts M. 2000. Introduction to the cyanobacteria. *The Ecology of Cyanobacteria*: 1-11.

# Putative gene regulation involved in the biosynthesis of *Nostoc punctiforme* exopolysaccharides

Chih-Yu Tseng<sup>1</sup>, Ming-Ho Iunn<sup>1</sup>, Guan-Chiun Lee<sup>1\*</sup>

<sup>1</sup>Department of Life Science, National Taiwan Normal University,  
Taipei, Taiwan

(Received: 20 March 2017, accepted: 20 May 2017)

## ABSTRACT

In cyanobacteria *Nostoc*, the genetics and biochemistry of exopolysaccharide (EPS) biosynthesis have not been thoroughly investigated. To identify the transcription factors that may regulate EPS biosynthesis genes and to understand the relationship between the EPS biosynthesis and the environment factors, in this study, we carried out the in silico promoter analysis of ten EPS biosynthesis-related genes in *Nostoc punctiforme* ATCC 29133, whose genomic sequence has been determined. Total 15 putative transcription factor binding sites were found in the ten promoter regions. Three transcription factors, ArgR, Fnr, and LexA, were predicted to be the most possible ones by comparing binding site sequence similarity between *E. coli* K12 and *N. punctiforme*. Because no ArgR and Fnr homologs were found in *N. punctiforme* genome by the Basic Local Alignment Search Tool (BLAST) using the amino acid sequences of *E. coli* K12 ArgR, Fnr and LexA as query sequences, we predicted that *N. Punctiforme* EPS biosynthetic genes may be regulated by the transcription factor LexA and its associated environmental factors such as UV light.

**Key words:** Cyanobacteria, *Nostoc punctiforme*, exopolysaccharide, transcription factor, promoter

---

\*Corresponding author: Guan-Chiun Lee ; FAX: 886-2-29312904 ; E-mail: gcleee@ntnu.edu.tw