



มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
รายงานวิจัยฉบับสมบูรณ์
ทุนวิจัยหมวดเงินอุดหนุน (ว.1)
ประจำปีงบประมาณ 2558

เรื่อง การอนุมานความสัมพันธ์เชิงเครือข่ายการควบคุมการแสดงออกของยีนใน
Arthrospira platensis C1 ภายใต้สภาวะอุณหภูมิสูง

Inferring transcriptional regulatory network of *Arthrospira*
platensis C1 in response to temperature stress

คณะผู้วิจัย

หัวหน้าโครงการ	นางสาวศวรรณี สุธีร์วรพงศ์
สังกัด	สถาบันพัฒนาและฝึกอบรมโรงงานต้นแบบ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี (บางขุนเทียน)
ผู้ร่วมโครงการ	นางสาวกัลยาณี ไพฑูรย์รังสฤษฎ์
สังกัด	หน่วยปฏิบัติการวิจัยและพัฒนาวิศวกรรมชีวเคมีและ โรงงานต้นแบบ สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ
ที่ปรึกษาโครงการ	นางสุภาภรณ์ ชีวะธนรักษ์
สังกัด	คณะทรัพยากรชีวภาพและเทคโนโลยี มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี (บางขุนเทียน)

เดือน กุมภาพันธ์ พ.ศ. 2559

บทคัดย่อ

Arthrospira platensis (*A. platensis*) C1 คือไซยาโนแบคทีเรียชนิดหนึ่งซึ่งมีปริมาณโปรตีนมากและสารมูลค่าสูงหลายชนิด อุตสาหกรรมจำนวนมากให้ความสนใจในการใช้ *A. platensis* C1 เป็นแหล่งผลิตสารมูลค่าสูงที่ต้องการ เป้าหมายของอุตสาหกรรมส่วนมาก ต้องการเพิ่มปริมาณน้ำหนักรวมในการผลิตซึ่งจำเป็นต้องเลี้ยงในกลางแจ้งซึ่ง *A. platensis* C1 ต้องเผชิญกับความกดดันต่างๆ เช่น อุณหภูมิสูง การขาดสารอาหาร และอื่นๆ ในการศึกษาครั้งนี้ เราได้มุ่งเน้นในการศึกษาการตอบสนองของ *A. platensis* C1 ต่อสภาวะอุณหภูมิสูง ในขั้นแรกเราได้วิเคราะห์ข้อมูลไมโครอาร์เรย์ของ *A. platensis* C1 ภายใต้สภาวะอุณหภูมิสูง และได้ยีนจำนวน 250 ยีนที่ตอบสนองต่อสภาวะดังกล่าว ยีนเหล่านี้ถูกแบ่งออกเป็น 5 กลุ่มตามรูปแบบการแสดงออกของยีน จากนั้น เราค้นหา Motif ที่มีร่วมกันของยีนที่แสดงออกร่วมกัน และเปรียบเทียบ Motif ที่ทำนายได้ด้วย Motif ของสิ่งมีชีวิตชนิดโปรคาริโอตที่มีการศึกษาไว้แล้ว เราจึงทำนายความสัมพันธ์เชิงควบคุมของยีน โดยยีนที่มีการแสดงออกร่วมกันและมี Motif ร่วมกันจะถือว่ามีความสัมพันธ์กัน และมีการทำนาย Regulator ที่มีแนวโน้มที่จะจับกับ Motif ที่ทำนายได้สุดท้าย เราได้สร้างเครือข่ายควบคุมการแสดงออกของยีนที่เป็นกลไกการตอบสนองต่อสภาวะอุณหภูมิสูงของ *A. platensis* C1 ภายใต้สภาวะอุณหภูมิสูง

Abstract

Arthrospira platensis (*A. platensis*) C1 are kind of cyanobacteria that contain high protein content and high valuable substances. Many industries are interested in using *A. platensis* C1 as a cell factory to produce desired substances. The aim is to increase their biomass that is usually grown in outdoor cultivation. *A. platensis* C1 have to face several stresses including high temperature, nutrient sufficiency, and etc. In this study, we focused on *A. platensis* C1 in response to high temperature stress. We first analyzed microarray data of *A. platensis* C1 under high temperature and 250 responsive genes were obtained. Genes were clustered into 5 groups on the basis of similar gene expression profiles. We searched for common motifs of genes expressing similarly and then compared predicted motifs with known motifs of prokaryotic organisms. We have inferred the regulatory relationship from the information whether genes that express similarly and share a common motif would be connected with a predicted regulator. Finally, we constructed a regulatory network of responsive mechanism under high temperature stress of *A. platensis* C1.

กิตติกรรมประกาศ

คณะผู้วิจัยขอขอบคุณมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรีที่ได้สนับสนุนงบประมาณสำหรับการดำเนินงานวิจัย ซึ่งเป็นทุนวิจัยหมวดเงินอุดหนุนที่ได้รับการจัดสรรจากรัฐ และขอขอบคุณสถาบันพัฒนาและฝึกอบรมโรงงานต้นแบบ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ที่ได้อำนวยความสะดวกในการใช้พื้นที่และเครื่องมือต่างๆ สำหรับการศึกษาวิจัย

คณะผู้วิจัย

กุมภาพันธ์ 2559

สารบัญ

บทคัดย่อ.....	2
Abstract.....	3
สารบัญ.....	5
สารบัญรูปภาพ	7
สารบัญตาราง.....	8
คำย่อ (Abbreviation).....	9
บทที่ 1 บทนำ.....	10
1.1 ความสำคัญและที่มาของปัญหาที่ทำการวิจัย.....	10
1.2 วัตถุประสงค์ของโครงการวิจัย.....	11
1.3 ขอบเขตของโครงการวิจัย	11
1.4 ทฤษฎี สมมติฐาน และกรอบแนวความคิดของโครงการวิจัย	11
บทที่ 2 การทบทวนวรรณกรรม/สารสนเทศที่เกี่ยวข้อง	12
2.1 การตอบสนองต่อสภาวะอุณหภูมิสูงของสาหร่ายสไปรูลีนา (<i>A. platensis</i> C1)	12
2.1.1 สาหร่ายสไปรูลีนา (<i>A. platensis</i> C1) และประโยชน์	12
2.1.2 การตอบสนองต่อสภาวะกดดันภายนอกของไซยาโนแบคทีเรีย.....	12
2.1.3 การตอบสนองต่อสภาวะอุณหภูมิสูงของสาหร่ายสไปรูลีนา (<i>A. platensis</i> C1)	13
2.2 การสร้างเครือข่ายการควบคุมยีน (Regulatory Network Reconstruction).....	13
2.2.1 นิยามและความหมายของเครือข่ายการควบคุมยีน.....	13
2.2.2 การอนุมานความสัมพันธ์เชิงควบคุมของยีน (Inference of Regulatory Relationship).....	14
2.3 เครือข่ายการควบคุมยีนในไซยาโนแบคทีเรีย	16
บทที่ 3 วิธีการดำเนินงานวิจัย	17
3.1 ข้อมูลที่ใช้ในงานวิจัย	17
3.2 การวิเคราะห์ข้อมูลไมโครอาร์เรย์ของ <i>A. platensis</i> C1 ภายใต้สภาวะอุณหภูมิสูง	17
3.3 การจัดกลุ่มยีนที่มีการแสดงออกร่วมกัน	17
3.4 การทำนาย Operon ของ <i>A. platensis</i> C1.....	17
3.5 การวิเคราะห์ความสัมพันธ์ของยีนจากการมี <i>cis</i> -regulatory element ร่วมกัน	18
3.6 การทำนาย Regulator ที่สามารถจับกับ <i>cis</i> -regulatory element	18
3.7 การวิเคราะห์ความสัมพันธ์ของยีนที่แสดงออกร่วมกัน.....	18
3.8 การอนุมานความสัมพันธ์เชิงเครือข่ายควบคุมการทำงานของยีน	18
บทที่ 4 ผลและการอภิปรายผล	20
4.1 ยีนที่แสดงออกแตกต่างกันอย่างมีนัยสำคัญภายใต้สภาวะอุณหภูมิสูง	20
4.2 กลุ่มยีนที่มีการแสดงออกร่วมกันภายใต้สภาวะอุณหภูมิสูง	20
4.3 Operon ของ <i>A. platensis</i> C1 สำหรับยีนที่ตอบสนองต่อสภาวะอุณหภูมิสูง	21

4.4 การวิเคราะห์การแสดงออกร่วมกัน	23
4.5 การค้นหา <i>cis-regulatory element</i> บริเวณหน้ายีนที่แสดงออกอย่างมีนัยสำคัญภายใต้สภาวะอุณหภูมิสูง	24
4.5.1 การค้นหา Known Motif บริเวณหน้ายีนที่แสดงออกอย่างมีนัยสำคัญภายใต้สภาวะอุณหภูมิสูง	24
4.5.2 Novel <i>cis-regulatory element</i> ที่ทำนายได้ซึ่งเกี่ยวข้องกับการควบคุมยีนในสภาวะอุณหภูมิสูง	25
4.6 เครือข่ายควบคุมยีนของ <i>A. platensis</i> C1 ภายใต้สภาวะอุณหภูมิสูง	33
บทที่ 5 บทสรุป	35
เอกสารอ้างอิง	36

สารบัญรูปภาพ

รูปที่ 2.1 แผนภาพการรับรู้สิ่งเร้า (Perception) การควบคุมยีน (Transduction) และการปรับตัวของเซลล์ (Acclimation) ต่อสิ่งเร้า.....	13
รูปที่ 3.1 แผนภาพแสดงภาพรวมการทำงานวิจัย.....	19
รูปที่ 4.1 รูปแบบการแสดงออกของยีนในแต่ละกลุ่ม.....	21
รูปที่ 4.2 ผลการทำนาย CIRCE element บริเวณหน้ายีนที่แสดงออกอย่างมีนัยสำคัญภายใต้สภาวะอุณหภูมิสูง.....	24
รูปที่ 4.3 ผลการทำนาย K-box element บริเวณหน้ายีนที่แสดงออกอย่างมีนัยสำคัญภายใต้สภาวะอุณหภูมิสูง.....	25
รูปที่ 4.4 เครือข่ายควบคุมยีนของ <i>A. platensis</i> C1 ภายใต้สภาวะอุณหภูมิสูง.....	33
รูปที่ 4.5 กลไกการทำงานของ Regulator สองชนิดได้แก่ SPLC1_S100480 (XRE Transcriptional Regulator) และ SPLC1_S500640 (HrcA Transcriptional Repressor).....	34

สารบัญตาราง

ตารางที่ 2.1 ฐานข้อมูลการแสดงออกร่วมกัน (Co-expression) ที่เกี่ยวข้องกับพืช.....	15
ตารางที่ 2.2 เครื่องมือในการทำนายโปรโมเตอร์และ <i>cis</i> -regulatory Element สำหรับสิ่งมีชีวิตชนิดโปรคาริโอต.....	16
ตารางที่ 4.1 ผลการจัดกลุ่มยีนด้วยอัลกอริทึม K-means Clustering	21
ตารางที่ 4.2 ผลการทำนาย Operon ของ <i>A. platensis</i> C1 สำหรับยีนที่ตอบสนองต่อสภาวะอุณหภูมิสูง.....	22
ตารางที่ 4.3 ผลการทำนาย Motif ของกลุ่มยีนและรูปแบบของ Motif.....	25
ตารางที่ 4.4 ผลการทำนาย Regulator ที่อาจจะจับกับ Motif	31

คำย่อ (Abbreviation)

Abbreviation	
TF	Transcription Factor
TFBS	Transcription Factor Binding Site
<i>A. platensis</i> C1	<i>Arthrospira platensis</i> C1
HSP	Heat Shock Protein
GLA	Gamma-Linolenic Acid
COG	Cluster of Orthologous Group

บทที่ 1 บทนำ

1.1 ความสำคัญและที่มาของปัญหาที่ทำการวิจัย

Arthrospira (Spirulina) เป็นสาหร่ายสีเขียวแกมน้ำเงินหรือไซยาโนแบคทีเรียที่มีการนำมาใช้ประโยชน์อย่างกว้างขวางทั้งในด้านอาหารเสริมสุขภาพของมนุษย์ อาหารสัตว์ และการวิจัยทางการแพทย์ เนื่องจาก *Spirulina* เป็นไซยาโนแบคทีเรียที่มีคุณค่าทางโภชนาการสูง โดยมีโปรตีนสูงถึง 60-70% ของน้ำหนักเซลล์ มีวิตามินต่าง ๆ โดยเฉพาะวิตามินบี 12 และโปรวิตามินเอ (เบต้าแคโรทีน) มีกรดไขมันจำเป็น Gamma-Linolenic Acid (GLA) มีแร่ธาตุต่างๆ โดยเฉพาะธาตุเหล็กและ phytonutrient จากการวิจัยทางการแพทย์พบว่า สารสกัดที่ได้จาก *Spirulina* สามารถช่วยเสริมสร้างภูมิคุ้มกันให้กับร่างกาย ลดความดันโลหิตสูง ลดโคเลสเตอรอลในกระแสเลือด ป้องกันโรคหัวใจ มะเร็ง และต้านทานต่อเชื้อแบคทีเรีย และไวรัสต่าง ๆ เป็นต้น (Ciferri et al., 1983; Belay et al., 2002; Ciferri et al., 1985) นอกจากนี้ เนื่องจาก *Spirulina* เป็นไซยาโนแบคทีเรียที่สามารถสังเคราะห์แสงและมีลักษณะคล้ายกับคลอโรพลาสต์ในเซลล์ของพืช เจริญเติบโตเร็ว ไม่ก่อให้เกิดโรค และไม่ผลิตสารพิษ สามารถเพาะเลี้ยงได้ตลอดทั้งปี โดยการเพาะเลี้ยงมีไปจนถึงระดับอุตสาหกรรม (Ciferri et al., 1985) จึงทำให้ *Spirulina* ได้รับความสนใจที่จะนำมาใช้เป็นแบบจำลองของพืช และไซยาโนแบคทีเรียสายพันธุ์ต่างๆ รวมถึงเป็นแหล่งผลิตสารชีวเคมีมูลค่าสูงให้กับอุตสาหกรรม

อย่างไรก็ตาม การเพาะเลี้ยง *Spirulina* ในสภาวะกลางแจ้งซึ่งความเข้มแสงและอุณหภูมิมีการเปลี่ยนแปลงตลอดเวลา โดยจะเพิ่มขึ้นจากตอนเช้าและจะสูงสุดในตอนเที่ยง และจะลดลงจากบ่ายจนถึงตอนเย็นที่หมดแสง การเปลี่ยนแปลงของสภาวะแวดล้อมนี้ ส่งผลให้การเจริญเติบโต การสร้างสารชีวเคมีต่างๆ ภายในเซลล์ ตลอดจนผลผลิตเซลล์มีความแตกต่างกันออกไปตามลักษณะภูมิอากาศในแต่ละวัน นอกจากนี้ กลไกทางชีววิทยาในการตอบสนองต่ออุณหภูมิสูงของ *Spirulina* มีความซับซ้อนสูงและยังไม่เป็นที่เข้าใจเพียงพอ ดังนั้น การควบคุมการเพิ่มผลผลิตการเพาะเลี้ยง *Spirulina* จึงทำได้ยาก

เพื่อที่จะควบคุมหรือเพิ่มปริมาณผลผลิตชีวมวลของ *Spirulina* จึงจำเป็นต้องเข้าใจการทำงานภายในเซลล์ของ *Spirulina* ในการตอบสนองต่อสภาวะแวดล้อมที่เปลี่ยนแปลงไป โดยการศึกษาการควบคุมการทำงานของ *Spirulina* ในระดับยีน (gene regulation) จากการศึกษาพบว่าสิ่งมีชีวิตปรับตัวอย่างต่อเนื่องเพื่อตอบสนองต่อสภาวะที่เปลี่ยนแปลงไป โดยการเปลี่ยนรูปแบบการแสดงออกของยีนซึ่งถูกควบคุมได้โดย Transcription Factor (TF) และรูปแบบการแสดงออกของยีนที่น่าสนใจคือ การแสดงออกร่วมกัน (Co-expression) ซึ่งเป็นรูปแบบที่มักจะเกิดขึ้นเมื่อยีนมีความเกี่ยวข้องกันทางด้านหน้าที่ หรือเชิงวิวัฒนาการ (Heyer et al., 1999; Spellman et al., 1998; Eisen et al., 1998) โดยยีนจะถูกควบคุมร่วมกัน (Co-regulation) เมื่อยีนคู่หนึ่งมีการแสดงออกร่วมกัน (Co-expression) และส่วนใหญ่เป็นยีนที่มี cis-regulatory element ที่เหมือนกันบนบริเวณของโปรโมเตอร์ ซึ่งจะช่วยให้ถูกควบคุมด้วย TF ชนิดเดียวกัน นอกจากนี้ จากการสังเกตยังพบอีกว่า ยีนที่เรียงตัวใกล้กันบนโครโมโซม (Co-localization) มีแนวโน้มที่จะมีการแสดงออกร่วมกันอีกด้วย (Dottorini et al., 2013; Hershberg et al., 2005; Boutanaev et al., 2002)

ดังนั้น งานวิจัยนี้จะสร้างความสัมพันธ์เชิงเครือข่ายการควบคุมการทำงานของยีน (Gene Regulatory Network) โดยอาศัยข้อมูลลำดับเบสจีโนมของ *A. platensis* C1 ข้อมูลการแสดงออกของยีนภายใต้สภาวะกดดันอุณหภูมิสูงแบบที่ศึกษาแบบ High-throughput ในระดับจีโนม และข้อมูลการทำนาย cis-regulatory element มาสร้างความสัมพันธ์ โดยผลที่ได้จะนำไปสู่ความเข้าใจกลไกการทำงานของ Signaling Pathway ที่ตอบสนองต่อสภาวะกดดันอุณหภูมิสูง ชี้ให้เห็นถึง TF ที่ถูกกระตุ้นให้ทำงาน ยีนใดบ้างที่เป็นเป้าหมายของ TF และการควบคุมยีนให้แสดงออกตามเวลาที่กำหนดไว้ได้อย่างไร

1.2 วัตถุประสงค์ของโครงการวิจัย

เพื่อศึกษากลไกการควบคุมการทำงานของยีนใน *Arthrospira* (*Spirulina*) *platensis* C1 ภายใต้สภาวะอุณหภูมิสูงโดยอาศัยการอนุมานความสัมพันธ์เชิงเครือข่ายควบคุมการแสดงออกของยีน

1.3 ขอบเขตของโครงการวิจัย

งานวิจัยนี้จะค้นหาความสัมพันธ์เชิงเครือข่ายควบคุมการทำงานของยีนของ *A. platensis* C1 โดยอาศัยข้อมูลสามอย่าง ได้แก่ ข้อมูลจีโนมและข้อมูลการแสดงออกของยีนภายใต้สภาวะอุณหภูมิสูง โดยทำการจัดกลุ่มด้วยอัลกอริทึม K-means Clustering จากนั้นกลุ่มยีนจะถูกนำไปทำนาย *cis*-regulatory element เพื่อหาความสัมพันธ์ของยีนและนำไปใช้ในการสร้างเครือข่ายควบคุมการทำงานของยีน และยืนยันความสัมพันธ์ของยีนในเครือข่ายโดยการทบทวนวรรณกรรม

1.4 ทฤษฎี สมมติฐาน และกรอบแนวความคิดของโครงการวิจัย

การเปลี่ยนแปลงระดับการแสดงออกของยีนเป็นกลไกที่เซลล์ *Spirulina* ใช้ในการตอบสนองต่อสิ่งแวดล้อม ซึ่งในท้ายที่สุดการเปลี่ยนแปลงนี้ส่งผลต่อปริมาณชีวมวลของสาหร่าย ดังนั้นเพื่อที่จะควบคุมปริมาณผลผลิตของการเพาะเลี้ยง *Spirulina* จึงจำเป็นต้องมีการศึกษากลไกการควบคุมการทำงานของยีนในระดับยีน โดยทั่วไปยีนมักจะถูกควบคุมพร้อมกันเมื่อมี *cis*-regulatory element ที่เหมือนกันบนบริเวณ promoter ซึ่งจะทำให้ TF ตัวเดียวกันสามารถมาจับได้ แต่ไม่จำเป็นต้องจับในเวลาเดียวกัน ดังนั้น *cis*-regulatory element จึงบ่งบอกถึงการปฏิสัมพันธ์กันทางกายภาพเท่านั้น ที่จริงแล้วการแสดงออกของยีนจะเกิดขึ้นภายใต้สภาวะหนึ่งๆ เท่านั้น ด้วยเหตุนี้การศึกษาการแสดงออกของยีนภายใต้สภาวะต่างๆ จึงมีบทบาทสำคัญเพื่อนำไปสู่ความเข้าใจกลไกการควบคุมการทำงานของยีนมากขึ้น กล่าวคือเมื่อยีนถูกควบคุมพร้อมกันภายใต้สภาวะหนึ่ง ยีนจะมีการแสดงออกร่วมกัน นอกจากนี้ยังมีรายงานอีกว่ายีนที่แสดงออกพร้อมกัน มักจะเรียงใกล้กันบนโครโมโซม หรือเรียงตัวใกล้กันบนคนละข้างของสาย DNA โดยอาศัย bidirectional promoter ดังนั้นงานวิจัยนี้จึงต้องการค้นหาความสัมพันธ์เชิงเครือข่ายการควบคุมการทำงานของยีน (transcriptional regulatory network) ด้วยข้อมูลลำดับเบสจีโนมของ *A. platensis* C1 (Supapon et al., 2012) และข้อมูลการแสดงออกของยีนภายใต้สภาวะอุณหภูมิสูงเพื่อศึกษากลไกการควบคุมการแสดงออกของยีนภายใต้สภาวะอุณหภูมิสูง โดยอนุมานความสัมพันธ์ของยีนจากการที่ยีนแสดงออกร่วมกันหรือถูกควบคุมด้วย TF ชนิดเดียวกัน ความสัมพันธ์เชิงเครือข่ายการควบคุมการทำงานของยีนนี้จะแสดงให้เห็นถึงกลุ่มของยีนที่ถูกควบคุมพร้อมกันโดย TF ตัวหนึ่งๆ และบ่งบอกถึงลำดับการควบคุมของ TF อีกด้วย

บทที่ 2 การทบทวนวรรณกรรม/สารสนเทศที่เกี่ยวข้อง

2.1 การตอบสนองต่อสภาวะอุณหภูมิสูงของสาหร่ายสไปรูลิน่า (*A. platensis* C1)

2.1.1 สาหร่ายสไปรูลิน่า (*A. platensis* C1) และประโยชน์

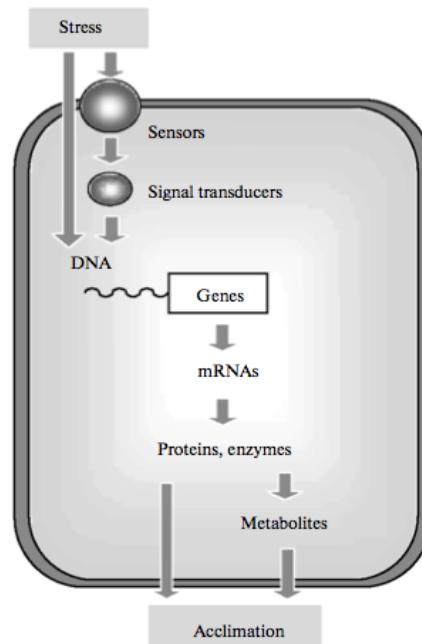
สาหร่ายสไปรูลิน่า เป็นสาหร่ายที่มีความสำคัญทางการค้าอย่างมาก เพราะเป็นสาหร่ายที่มีประโยชน์หลากหลาย เช่น การผลิตอาหารเสริม อาหารสัตว์ เครื่องสำอาง และผลิตภัณฑ์เกี่ยวกับยา ทั้งนี้เนื่องจากมีปริมาณโปรตีนและคุณค่าทางอาหารสูง ซึ่งเกี่ยวข้องกับ กรดไขมันจำเป็น ไพโคไซยานิน และเบต้าแคโรทีน (Spolaore et al., 2006)

สาหร่ายสไปรูลิน่าเป็นสาหร่ายสีเขียวแกมน้ำเงินหรือไซยาโนแบคทีเรียชนิดหนึ่ง ซึ่งสามารถกินได้ ไซยาโนแบคทีเรียมีลักษณะและโครงสร้างคล้ายคลึงกับคลอโรพลาสต์ในพืช โดยมีโครงสร้างของเมมเบรน (Membrane) และกลไกการใช้ออกซิเจนที่คล้ายกับพืช และยังสามารถสังเคราะห์แสงได้คล้ายพืชอีกด้วย ดังนั้น ไซยาโนแบคทีเรียมักจะถูกใช้เป็นสิ่งมีชีวิตต้นแบบในการศึกษากระบวนการสังเคราะห์แสง แทนการศึกษาจากพืชโดยตรง ทั้งนี้เนื่องจากไซยาโนแบคทีเรียมีการเจริญเติบโตที่รวดเร็วกว่าเมื่อเทียบกับพืช และมีความซับซ้อนน้อยกว่าพืชมาก และเนื่องจากพืชเป็นสิ่งมีชีวิตที่อยู่กับที่ จึงทำให้มันไม่สามารถเลือกสภาพแวดล้อมที่เหมาะสมสำหรับการเจริญเติบโต และการแพร่พันธุ์ของมันได้ วงจรชีวิตของมันจึงค่อนข้างเป็นไปตามสภาพแวดล้อม ทำให้มันไวต่อปัจจัยกดดันจากภายนอก เช่น ความร้อน ความเย็น แสง และการขาดสารอาหารอีกด้วย แต่พืชเป็นสิ่งมีชีวิตที่ซับซ้อนสูง เมื่อเทียบกับไซยาโนแบคทีเรีย ดังนั้น การศึกษากลไกในระดับโมเลกุลที่ตอบสนองต่อปัจจัยกดดันจากภายนอก จึงมักมีการใช้ไซยาโนแบคทีเรียมาศึกษาการตอบสนองต่อสิ่งแวดล้อม

นอกจากนี้ ปัจจุบันได้มีการหาลำดับเบสของจีโนมของสาหร่ายสไปรูลิน่าทั้ง 5 สายพันธุ์แล้ว ได้แก่ *Arthrospira platensis* NIES-39 (Fujisawa et al., 2010), *Arthrospira* PCC-8005 (Janssen et al., 2010), *Arthrospira maxima* CS-328 (ref), *Arthrospira platensis* Paraca (Lefort et al., 2014) และ *Arthrospira platensis* C1 (Cheevadhanarak et al., 2012) ซึ่งถูกเก็บไว้ใน FTP ของฐานข้อมูล NCBI (<ftp://ftp.ncbi.nlm.nih.gov/genomes>) *Arthrospira platensis* NIES39 เป็นเพียงสายพันธุ์เดียวเท่านั้น ที่มีลำดับเบสจีโนมที่สมบูรณ์แล้ว

2.1.2 การตอบสนองต่อสภาวะกดดันภายนอกของไซยาโนแบคทีเรีย

การตอบสนองของต่อสภาวะกดดันภายนอกของไซยาโนแบคทีเรีย มีกลไกพื้นฐานดังรูปที่ 2.1 โดยที่โปรตีนเซนเซอร์ (Sensory Protein or Sensor) ได้รับสัมผัสจากสิ่งแวดล้อมภายนอก ทำให้มีการเปลี่ยนแปลงภายในโปรตีนเซนเซอร์ โปรตีนเซนเซอร์จะส่งสัญญาณเกี่ยวกับการเปลี่ยนแปลงไปยัง โพลีเปปไทด์ตัวอื่นๆ หรือโปรตีนทรานสดิวเซอร์ (Transducer) ซึ่งจะควบคุมการแสดงออกของยีนที่ตอบสนองต่อสภาวะกดดันภายนอก โปรตีนทรานสดิวเซอร์สามารถจดจำส่วนของ DNA โดยตรง และจับกับส่วนของ DNA และควบคุมการแสดงออกของยีน ในที่สุด โปรตีนหรือเมตาบอไลต์ปกป้อง (Protection Protein or Metabolite) ถูกสร้างขึ้นเพื่อช่วยสิ่งมีชีวิตให้ปรับตัวต่อสิ่งแวดล้อมใหม่ อย่างไรก็ตาม อุณหภูมิสูงอาจจะมีผลกระทบโดยตรงต่อโครงสร้างของ DNA และเป็นสาเหตุให้เกิดการเปลี่ยนแปลงของการแสดงออกของยีนจำนวนมาก



รูปที่ 2.1 แผนภาพการรับรู้สิ่งเร้า (Perception) การควบคุมยีน (Transduction) และการปรับตัวของเซลล์ (Acclimation) ต่อสิ่งเร้า (รูปจาก Zorina et al., 2011)

2.1.3 การตอบสนองต่อสภาวะอุณหภูมิสูงของสาหร่ายสไปรูลีนา (*A. platensis* C1)

ปัจจุบันมีการศึกษาการแสดงออกของยีนในสาหร่ายสไปรูลีนาภายใต้สภาวะอุณหภูมิสูง (Panyakampol et al., 2014) จากการศึกษาพบว่า ความสามารถในการสังเคราะห์แสงลดลงทันที เมื่อสาหร่ายสไปรูลีนาถูกเปลี่ยนจากอุณหภูมิ 35 °C ไปยัง 42 °C เป็นเวลา 2 ชั่วโมง นอกจากนี้ หลังจากให้ความร้อนเป็นเวลา 24 ชั่วโมง มีการลดลงของน้ำหนักแห้ง (Biomass) และปริมาณโปรตีน และยังเกิดการสะสมคาร์โบไฮเดรตเพิ่มมากขึ้นด้วย การตอบสนองของสไปรูลีนาเกี่ยวข้องกับยีนในกลุ่ม Heat Shock Protein (HSP) ยีนในกลุ่ม Signal Transduction และยีนในกลุ่มคาร์บอนและไนโตรเจน และยีนที่เกี่ยวข้องกับ Inorganic Ion Transporter สำหรับแมกนีเซียม ไนเตรตและไนไตรต์

2.2 การสร้างเครือข่ายการควบคุมยีน (Regulatory Network Reconstruction)

2.2.1 นิยามและความหมายของเครือข่ายการควบคุมยีน

เครือข่ายการควบคุมยีน (Regulatory Network) เป็นเครือข่ายทางชีววิทยาที่ได้รับความสนใจมาก ซึ่งเป็นการระบุความสัมพันธ์เชิงควบคุมระหว่างยีนที่ทำหน้าที่ควบคุม (Regulatory Gene) กับยีนเป้าหมาย (Target Gene) การค้นหาเครือข่ายการควบคุมยีนเป็นงานที่มีบทบาทสูงและมีความสำคัญมาก ในด้านชีวโมเลกุล เพราะช่วยให้เข้าใจกลไกที่ซ่อนอยู่ภายในเซลล์ของสิ่งมีชีวิตที่มีความซับซ้อนสูง และยังทำให้เข้าใจกลไกการตอบสนองต่อสิ่งเร้าทั้งภายในและภายนอกของสิ่งมีชีวิต เครือข่ายการควบคุมยีนมีลักษณะเหมือนเครือข่ายทั่วไป คือประกอบด้วย Node และ Edge ซึ่งหมายถึงยีนหรือโปรตีนควบคุม และความสัมพันธ์เชิงควบคุมระหว่างยีนกับโปรตีนควบคุมตามลำดับ

2.2.2 การอนุมานความสัมพันธ์เชิงควบคุมของยีน (Inference of Regulatory Relationship)

ปัจจุบันมีวิธีการมากมายที่ถูกสร้างขึ้นมาเพื่อใช้ในการอนุมานเครือข่ายการควบคุมการทำงานของยีน เช่น Lee และคณะ (2002) และ Bar-Joseph (2003) ทำการวิเคราะห์หา *cis*-regulatory element บริเวณโปรโมเตอร์ที่ทำหน้าที่จับกับ TF ในยีสต์ (*Saccharomyces cerevisiae*) และใช้ข้อมูลดังกล่าว ในการอนุมานเครือข่ายการควบคุมยีน อย่างไรก็ตาม TF จะทำงานเฉพาะเวลาหนึ่งๆ เท่านั้นซึ่งขึ้นอยู่กับสภาวะแวดล้อม การค้นหา *cis*-regulatory element ในแบบดังกล่าว จึงเป็นเพียงการเปิดเผยการจับกันเชิงกายภาพ (Physical Interaction) ไม่ได้บ่งชี้ถึงการจับกันตามสภาวะที่แท้จริง ดังนั้น การใช้ข้อมูลการแสดงออกของยีนในการอนุมานเครือข่ายการควบคุมการแสดงออกของยีนจึงเข้ามามีบทบาทมากขึ้น มีการศึกษาอย่างกว้างขวางในการใช้ข้อมูลการแสดงออกของยีน เช่น ข้อมูล DNA ไมโครอาร์เรย์ ข้อมูลลำดับเบส DNA และข้อมูลเกี่ยวกับหน้าที่ของยีน มาเพื่ออนุมานเครือข่ายการควบคุมการทำงานของยีน (Pilpel et al., 2001; Wang et al., 2002; Segal et al., 2003; Beer and Tavazoie, 2004) โดย Xing และ van der Laan (2005) เป็นกลุ่มแรกที่ใช้วิธีทางสถิติในการสร้างเครือข่ายการควบคุมการทำงานของยีนที่มีความสอดคล้องกับการเปลี่ยนแปลงการแสดงออกของยีนภายใต้เงื่อนไขต่างๆ ทางทางทดลองโดยอาศัยข้อมูลการแสดงออกของยีน ข้อมูลลำดับเบสบนโปรโมเตอร์ และข้อมูล Transcription Factor Binding Sites (TFBS)

2.2.2.1 การใช้ข้อมูลการแสดงออกของยีนเพื่อสร้างเครือข่ายการควบคุมยีน

เครือข่ายควบคุมยีนยังนิยามสร้างด้วยความสัมพันธ์แบบ Co-expression (Co-expression Network) เนื่องจากการแสดงออกร่วมกันของยีนมักจะแสดงถึงหน้าที่ที่เกี่ยวข้องกัน อาจจะสามารถให้เอนไซม์ที่ทำงานบนวิถีเมตาบอลิซึม (Pathway) เดียวกัน หรือยีนทั้งสองอาจจะเกี่ยวข้องกันเชิงวิวัฒนาการ และหนึ่งในสาเหตุที่สำคัญที่ทำให้กลุ่มยีนมีการแสดงออกร่วมกันเพราะยีนเหล่านี้มี Transcription Factor Binding Site (TFBS) ที่เหมือนกัน จึงถูกควบคุมด้วย TF ชนิดเดียวกัน แต่ยีนมี *cis*-regulatory element ที่เหมือนกันไม่จำเป็นต้องถูกควบคุมพร้อมกันเสมอไป (Heyer et al., 1999; Spellman et al., 1998; Eisen et al., 1998) นอกจากนี้ยังมีรายงานที่แสดงให้เห็นว่าการแสดงออกร่วมกันของยีน อาจจะเป็นเพราะลักษณะการเรียงตัวของยีนเหล่านั้นที่อยู่ใกล้กันบนจีโนม อาจเรียกว่ายีนกลุ่มนี้มีการ Co-localized กัน จากการศึกษาของ Dai และคณะ (2011) พบว่าการจัดเรียงตัวของยีนบนจีโนมนั้นไม่ได้เป็นไปอย่างสุ่ม และมีงานวิจัยอื่นๆ อีกมากที่ยืนยันความสัมพันธ์ระหว่างการแสดงออกร่วมกันของยีนกับตำแหน่งของยีนบนโครโมโซม (Dottorini et al., 2013; Hershberg et al., 2005; Boutanaev et al., 2002) ในสิ่งมีชีวิตแบบยูคาริโอตหลายชนิด ยีนที่แสดงออกร่วมกันมักจะถูกจัดเรียงตัวในลักษณะคล้าย Operon ซึ่งแสดงให้เห็นว่ายีนบนจีโนมไม่ได้เรียงตัวอย่างสุ่ม แต่ขึ้นอยู่กับกระบวนการทำงานของยีนใน Signaling Pathway และ Metabolic Pathway (Hurst et al., 2004) ด้วยความสำคัญของการแสดงออกร่วมกันของยีน จึงมีนักชีววิทยาให้ความสนใจเป็นอย่างมากและพัฒนาเว็บแอปพลิเคชันที่วิเคราะห์และเก็บรวบรวมยีนที่มีการแสดงออกร่วมกันไว้เป็นจำนวนมาก ดังตารางที่ 2.1

ตารางที่ 2.1 ฐานข้อมูลการแสดงออกร่วมกัน (Co-expression) ที่เกี่ยวข้องกับพืช

Species	Database	URL
Cross species		
	PlaNet	http://aranet.mpimp-golm.mpg.de/
Arabidopsis		
	ATTED-II	http://atted.jp/
	AraNet	http://www.functionalnet.org/aranet/
	BAR	http://bar.utoronto.ca/welcome.htm
	COP	http://webs2.kazusa.or.jp/kagiana/cop
	CORNet	http://bioinformatics.psb.ugent.be/cornet
	CSB.DB	http://csbdb.mpimp-golm.mpg.de/csbdb/dbcor/ath.html
	CressExpress	http://www.cressexpress.org/index
	GeneCAT	http://genecat.mpg.de/cgi-bin/Ainitiator.py
	GeneVestigator	https://www.genevestigator.com/gv/plant.jsp
	KappaViewer 4	http://kpv.kazusa.or.jp/en/
	SeedNet	http://bree.cs.nott.ac.uk/arabidopsis/
<i>Oryza sativa</i>		
	OryzaExpress	http://bioinf.mind.meiji.ac.jp/Rice_network_public/script/index.html
	ATTED-II	http://atted.jp/
<i>Populus trichocarpa</i>		
	COP	http://webs2.kazusa.or.jp/kagiana/cop/

2.2.2.2 การทำนายโปรโมเตอร์และ cis-regulatory Element (Prediction of Promoter and cis-regulatory Element)

ปัจจุบันมีเครื่องมือที่ใช้ในการทำนายโปรโมเตอร์และ cis-regulatory Element ในสิ่งมีชีวิตชนิดโปรคาริโอตจำนวนมาก ดังแสดงในตารางที่ 2.2 ซึ่งประกอบด้วยเครื่องมือในการทำนายและฐานข้อมูลเกี่ยวกับ Regulatory Sites และ TF ของโปรคาริโอต

ตารางที่ 2.2 เครื่องมือในการทำนายโปรโมเตอร์และ *cis*-regulatory Element สำหรับสิ่งมีชีวิตชนิดโปรคาริโอต

Name	Type	Description	URL
Softberry	Tool	- Predict prokaryote promoter - Find DNA motif	http://www.softberry.com
PePPER	Tool	- Predict prokaryote promoter - Find DNA motif - TFBS search	http://pepper.molgenrug.nl/
MEME	Tool	- Find DNA motif	http://meme-suite.org/
PromPredict	Tool	- Predict promoter	
RegTransBase	Database	- a database of regulatory interactions based on literature and an expertly curated database of TFBS	http://regtransbase.lbl.gov
PRODORIC	Database	- a database about gene regulation and gene expression in prokaryotes	http://www.prodoric.de/
SwissRegulon	Database	- a database of genome-wide annotations of regulatory sites	http://swissregulon.unibas.ch/cgi/sr

2.3 เครือข่ายการควบคุมยีนในไซยาโนแบคทีเรีย

ในโปรคาริโอต (Prokaryote) ยีนที่อยู่ติดกันหลายๆ ยีนบน DNA สายเดียวกัน มักจะอยู่ในรูปแบบของ Operon และมีการแสดงออกร่วมกันเป็นแบบ Polycistronic mRNA ซึ่งยีนเหล่านี้จะมีจุดเริ่มต้นของการแสดงออก (Transcription Initiation) ที่ตำแหน่งเดียวกัน กลไกและสัญญาณการควบคุมการสิ้นสุดการแสดงออก (Transcription Termination) เป็นแบบเดียวกันด้วย โดยทั่วไปถ้าจีโนมไหนมีจำนวน TF น้อยกว่าจำนวน Operon จะทำให้โปรตีน TF หนึ่งตัวต้องทำหน้าที่ควบคุมหลาย Operon หรือยีนที่ถูกควบคุมด้วย TF เดียวกันนี้จึงถูกเรียกว่า Regulon ของ TF นั้น และ Operon หนึ่งๆ สามารถอยู่ในหลาย Regulon กลุ่มของ *cis*-regulatory element ที่ถูกจดจำด้วย TF เดียวกัน เรียกว่า *cis*-regulatory motif หรือ binding site motif การค้นหา TFBS ในไซยาโนแบคทีเรียได้ถูกทำอย่างแพร่หลาย เช่น Arun และคณะ (2013) ทำนาย *cis*-regulatory element ใน *Synechococcus elongatus* sp. PCC6301 โดยใช้วิธี phylogenetic footprinting ร่วมกับ MEME suite (Bailey et al., 2009) ซึ่งเป็นเครื่องมือที่เปิดให้ใช้ฟรีบนอินเทอร์เน็ต (Arun et al., 2013) ปัจจุบันมีฐานข้อมูลที่เก็บรวบรวมข้อมูล *cis*-regulatory element ของไซยาโนแบคทีเรียไว้ซึ่งชื่อว่า CyanoCis (<http://cyanocis.uohbif.in/>) ซึ่งได้ทำนาย *cis*-regulatory element โดยใช้วิธี Comparative Phylogenetic Footprinting จากข้อมูลลำดับเบสของไซยาโนแบคทีเรีย 14 ชนิด การค้นหา *cis*-regulatory element มีหลากหลายวิธี แต่วิธีที่มีความนิยมมากที่สุดในปัจจุบันคือ Comparative Phylogenetic Footprinting

บทที่ 3 วิธีการดำเนินงานวิจัย

การเปลี่ยนแปลงตามเวลาหรือสภาพแวดล้อม ทำให้สิ่งมีชีวิตปรับตัวเพื่ออยู่รอดในสภาวะที่เปลี่ยนแปลงไป เพื่อที่จะเข้าใจกลไกการทำงานของกระบวนการทางชีววิทยาภายในเซลล์ของ *A. platensis* C1 ในการตอบสนองต่อสิ่งเร้าจากภายนอกและสภาพแวดล้อมที่เปลี่ยนแปลงไป เราจึงทำการศึกษาความสัมพันธ์เชิงเครือข่ายควบคุมยีน ในงานวิจัยนี้ได้ทำการสร้างเครือข่ายควบคุมยีน (Gene Regulatory Network) ของ *A. platensis* C1 ที่ตอบสนองต่อสภาวะอุณหภูมิสูง โดยสร้างความสัมพันธ์จากยีนที่มีการแสดงออกร่วมกัน โดยอาศัยข้อมูลไมโครอาร์เรย์แบบ time-series ภายใต้สภาวะอุณหภูมิสูง และทำการจัดกลุ่มเพื่อค้นหายีนที่มีรูปแบบการแสดงออกเหมือนกันภายใต้เวลา จากนั้นได้มีการหาพื้นที่ที่สามารถถูกควบคุมด้วย Regulator ตัวเดียวกัน โดยการวิเคราะห์หา cis-regulatory element ร่วมกันของกลุ่มยีนที่มีรูปแบบการแสดงออกเหมือนกัน การวิเคราะห์ดังกล่าวทำให้เราเข้าใจกลไกการตอบสนองต่ออุณหภูมิสูงของ *A. platensis* C1 มากขึ้น ดังรูปที่ 3.1 โดยมีรายละเอียดการดำเนินงานวิจัยดังนี้

3.1 ข้อมูลที่ใช้ในงานวิจัย

ข้อมูลจีโนมของ *A. platensis* C1 ได้มีการเก็บรวบรวมไว้ในฐานข้อมูลที่มีชื่อว่า ArthroBase (<http://sbi.pdti.kmutt.ac.th/arthrobase/>) ซึ่งถูกสร้างขึ้นมาโดยกลุ่มวิจัยชีววิทยาระบบและชีวสารสนเทศ *A. platensis* C1 มียีนที่แสดงออกเป็น mRNA (Protein-coding Gene) ทั้งหมด 6,108 ยีน และยีนที่แสดงออกเป็น RNA (RNA Gene) 45 ยีน จากยีนทั้งหมดที่พบมีเพียง 3,757 ยีนที่สามารถระบุหน้าที่ได้ และจำแนกตามกลุ่ม COGs ได้ 3,459 ยีน โดยจำแนกเป็นยีนที่ทำหน้าที่เป็นเอนไซม์ 952 ยีน และยีนที่ทำหน้าที่เป็น Transporter 345 ยีน (Cheevadhanarak et al., 2012) และพบว่าประมาณ 217 ยีนเป็นยีนที่เกี่ยวข้องกับการควบคุมการแสดงออกของยีน เช่น ยีนในกลุ่ม Sigma Factor, Transcriptional Regulator และ Response Regulator เป็นต้น (<http://sbi.pdti.kmutt.ac.th/arthrobase/>)

3.2 การวิเคราะห์ข้อมูลไมโครอาร์เรย์ของ *A. platensis* C1 ภายใต้สภาวะอุณหภูมิสูง

ข้อมูลไมโครอาร์เรย์ที่ใช้ในงานวิจัยนี้เป็นข้อมูลไมโครอาร์เรย์แบบ Two-channel โดยสาหร่ายสไปรูลีนาถูกเลี้ยงภายใต้สภาวะอุณหภูมิสูงที่ 42 องศาเซลเซียส แบบ Time-series เวลา 20 60 120 240 และ 480 นาทีตามลำดับ และวิเคราะห์หาพื้นที่การแสดงออกแตกต่างกัน อย่างมีนัยสำคัญ ซึ่งเป็นยีนที่ตอบสนองต่อสภาวะอุณหภูมิสูง ด้วยเครื่องมือในการวิเคราะห์ที่ชื่อว่า Limma (Smyth, 2004) โดยตัด Cutoff ที่ Fold Change (FC) > 2.0 หรือ FC < 1/2 (Log of FC > 1.0 หรือ Log of FC < -1.0) และ P-value < 0.01 ซึ่งหมายถึงยีนมีการแสดงออกมากกว่าสภาวะปกติ (เวลาที่ 0 นาที) สองเท่าหรือน้อยกว่าสองเท่า ตามลำดับ

3.3 การจัดกลุ่มยีนที่มีการแสดงออกร่วมกัน

ยีนที่มีการแสดงออกแตกต่างกันอย่างมีนัยสำคัญ ด้วยวิธีการจัดกลุ่ม K-means Clustering โดยมีการประเมินค่า K ที่เหมาะสมด้วยค่า Within Cluster Variance ที่น้อยที่สุด ซึ่ง K=5 ทำให้ค่า Within Cluster Variance มีค่าน้อยที่สุด และเลือกใช้มาตรวัด Cosine Distance ซึ่งสามารถวัดได้ทั้งรูปแบบและปริมาณการแสดงออกที่เหมือนกัน

3.4 การทำนาย Operon ของ *A. platensis* C1

ยีนที่มีการจัดเรียงตัวเป็น Operon จะมีการแสดงออกพร้อมกัน โดยจะมี Regulator เข้าควบคุมเฉพาะที่บริเวณหน้า Operon หรือหน้ายีนแรกของ Operon เท่านั้น ดังนั้นในการวิเคราะห์หา Motif ร่วมกันของกลุ่มยีนจึงจำเป็นต้องพิจารณาการจัดเรียงตัวเป็น Operon ในการทำนาย Operon ถ้ากลุ่มยีนใดตรงกับเงื่อนไขดังนี้ (1) ยีนต้องเรียงตัวบนสาย DNA (Strand) ข้างเดียวกัน (2) ยีนที่อยู่ติดกันต้องห่างกันไม่เกิน 100 เบส และ (3) ยีนที่

แสดงออกบน Transcript เส้นเดียวกัน (โดยอาศัยข้อมูล RNA-seq ของห้องปฏิบัติการสาหร่ายเทคโนโลยี ของสถาบันพัฒนาและฝึกอบรมโรงงานต้นแบบ มจร. บางขุนเทียน ซึ่งเป็นข้อมูลที่ยังไม่ได้ตีพิมพ์) แล้วถือว่ายีนกลุ่มนี้ถูกทำนายว่าเป็น Operon

3.5 การวิเคราะห์ความสัมพันธ์ของยีนจากการมี *cis-regulatory element* ร่วมกัน

เพื่อค้นหายีนที่มี *cis-regulatory element* ร่วมกัน เราเริ่มจากการค้นหากลุ่มยีนที่แสดงออกร่วมกัน โดยใช้วิธีการจัดกลุ่มแบบ K-means และแบ่งยีนออกเป็นกลุ่มย่อย จากนั้นตัดหน้ายีนมา 250 เบสจากตำแหน่ง Translation Initiation Site โดยที่ยีนที่อยู่บน Operon เดียวกันจะพิจารณาเฉพาะยีนตัวที่อยู่ซ้ายสุดบนสาย Sense และขวาสุดบนสาย Antisense และใช้เครื่องมือ MEME (Bailey et al., 2009) ในการค้นหา Motif ที่มีร่วมกันบนกลุ่มยีน ด้วยการพิจารณา 3 Motif แรกที่มีคะแนนสูงสุดมาพิจารณาก่อน ถ้า Motif มีความเหมือนกันมากภายในกลุ่มยีน และปรากฏอยู่บนยีนจำนวนมากในกลุ่ม Motif ดังกล่าวก็จะถูกพิจารณาเป็น *cis-regulatory element* ของกลุ่มยีนนั้น เพื่อใช้ในการทำนาย Regulator ที่สามารถเข้าจับได้ในขั้นตอนต่อไป

3.6 การทำนาย Regulator ที่สามารถจับกับ *cis-regulatory element*

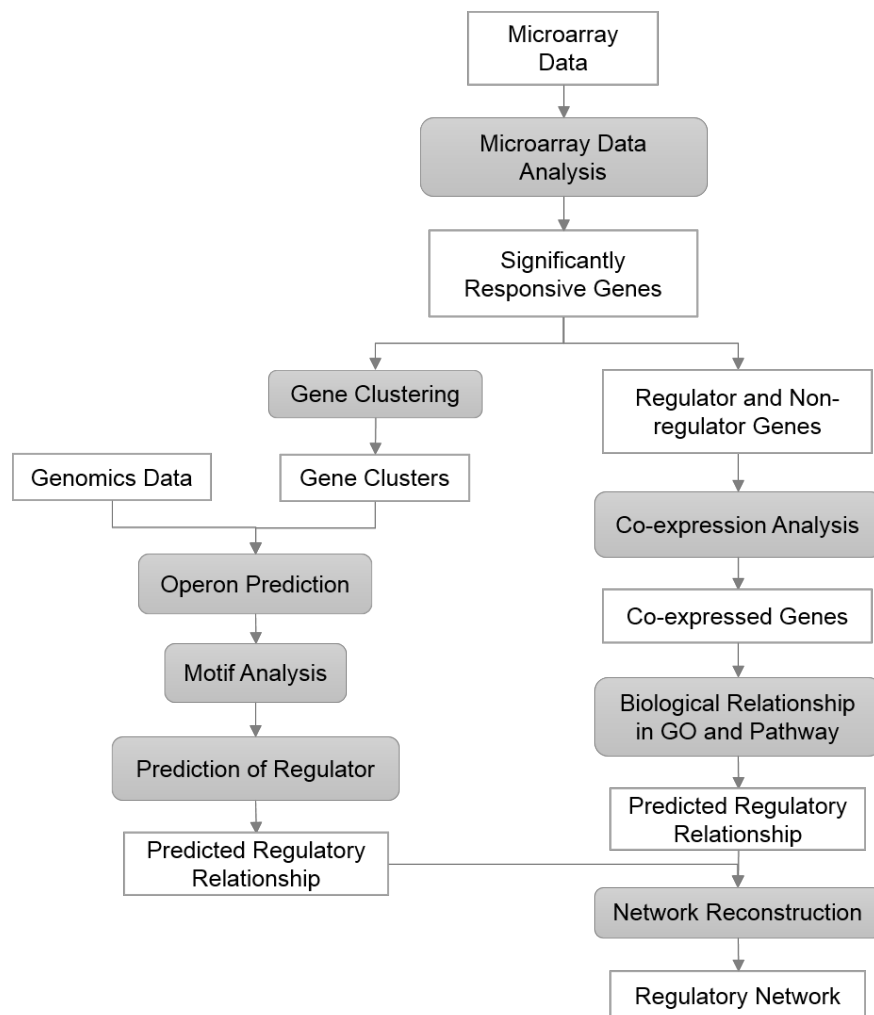
เพื่อที่จะทำนาย Regulator ที่สามารถจับกับ *cis-regulatory element* เราได้ใช้ TOMTOM (Gupta et al., 2007) ซึ่งทำการเทียบ Motif ที่ทำนายได้กับฐานข้อมูล Prodoic เวอร์ชัน 8.9 ซึ่งเป็นฐานข้อมูลที่เก็บรวบรวม *cis-regulatory element* ของสิ่งมีชีวิตแบบโปรคาริโอตไว้ รวมถึงข้อมูลของ Transcription Factor (TF) ที่สามารถจับกับ *cis-regulatory element* ดังกล่าว TOMTOM จะทำนาย *cis-regulatory element* ที่มีความคล้ายคลึงกับ Motif มากที่สุดรวมทั้ง TF ที่สามารถจับกับ *cis-regulatory element* ดังกล่าวได้

3.7 การวิเคราะห์ความสัมพันธ์ของยีนที่แสดงออกร่วมกัน

เพื่อที่จะหาความสัมพันธ์ของยีนที่แสดงออกร่วมกัน จากยีนที่ตอบสนองต่อสภาวะอุณหภูมิสูงอย่างมีนัยสำคัญ เราใช้ Pearson's Correlation Coefficient (PCC) ในการคำนวณความสัมพันธ์ระหว่างโปรไฟล์ การแสดงออกของยีนสองยีนจากข้อมูลไมโครอาร์เรย์ดังกล่าว โดยเลือก Cutoff ที่ $PCC > 0.90$ ซึ่งเป็นความสัมพันธ์เชิงบวก

3.8 การอนุมานความสัมพันธ์เชิงเครือข่ายควบคุมการทำงานของยีน

ในการอนุมานความสัมพันธ์เชิงเครือข่ายควบคุมการทำงานของยีน มีการนิยามความสัมพันธ์ 2 ชนิด (1) ความสัมพันธ์ระหว่างยีนกับยีนและ (2) ความสัมพันธ์ระหว่างยีนกับ Regulator ความสัมพันธ์ระหว่างยีนกับยีน ได้จากการที่ยีนสองยีนใดๆ มีการแสดงออกร่วมกัน โดยอาศัยข้อมูลไมโครอาร์เรย์ สำหรับความสัมพันธ์ระหว่างยีนกับ Regulator ได้จากการทำนาย Motif ของยีนที่แสดงออกร่วมกันและค้นหา *cis-regulatory element* ที่มีความคล้ายคลึงกับ Motif มากที่สุด และทำนาย Regulator ที่สามารถจับกับ *cis-regulatory element* ได้ และเพื่อที่จะให้เห็นความสัมพันธ์ที่ชัดเจน จึงนำความสัมพันธ์ทั้งหมดมาวาดเป็นเครือข่าย ด้วยเครื่องมือ Cytoscape (Shannon et al., 2003)



รูปที่ 3.1 แผนภาพแสดงภาพรวมการทำงานวิจัย (1) วิเคราะห์ข้อมูลไมโครอาร์เรย์ (2) จัดกลุ่มยีน (3) ทำนาย Operon (4) วิเคราะห์ Motif ร่วมกันของกลุ่มยีน (5) ทำนาย Regulator ที่สามารถจับกับ *cis*-regulatory element ได้ (6) วิเคราะห์การแสดงออกร่วมกัน (7) ค้นหาความสัมพันธ์ระหว่างยีนด้วย GO term และ Pathway และ (8) สร้างเครือข่ายเชิงควบคุมยีน

บทที่ 4 ผลและการอภิปรายผล

ในการศึกษากลไกการควบคุมการแสดงออกของยีนใน *A. platensis* C1 ภายใต้สภาวะอุณหภูมิสูง ด้วยข้อมูลไมโครอาร์เรย์แบบลำดับเวลา (Time-series Microarray Data) และข้อมูลลำดับเบสจีโนม ทำให้เราสามารถทำนายความสัมพันธ์เชิงควบคุม ด้วยสมมติฐานว่า ยีนที่มีการตอบสนองต่อสภาวะกดดันเหมือนกัน มีแนวโน้มที่จะถูกควบคุมพร้อมกัน และยีนที่มี Motif ที่บริเวณโปรโมเตอร์คล้ายกัน มีแนวโน้มที่จะถูกควบคุมด้วย Regulator ชนิดเดียวกัน ดังนั้น เราจึงใช้เหตุการณ์ทั้งสองนี้ในการค้นหาความสัมพันธ์เชิงควบคุมของ *A. platensis* C1 ภายใต้สภาวะอุณหภูมิสูง โดยค้นหากลุ่มยีนที่มีการตอบสนองคล้ายกันในสภาวะอุณหภูมิสูงและค้นหา Motif ที่มีส่วนร่วมของยีนเหล่านี้

4.1 ยีนที่แสดงออกแตกต่างกันอย่างมีนัยสำคัญภายใต้สภาวะอุณหภูมิสูง

การวิเคราะห์ข้อมูลไมโครอาร์เรย์ของสาหร่ายสไปรูลิน่า *A. platensis* C1 ด้วยแพ็คเกจ Limma ซึ่งอยู่ใน Bioconductor ของ R programming จากการคัดเลือกยีนที่มีการแสดงออกแตกต่างกันตามเวลาอย่างมีนัยสำคัญ พบว่ามี 250 ยีนที่มีความแตกต่างอย่างมีนัยสำคัญที่ระดับ $P\text{-value} < 0.01$ และ $\text{Fold Change} > 2$ ($\text{Log(FC)} > 1.0$ และ $\text{Log(FC)} < -1.0$) ซึ่งแสดงไว้ในเอกสารเพิ่มเติม 1 ในจำนวนนี้มีเพียงหนึ่งยีนที่ทำหน้าที่เป็น Transcriptional Regulator ซึ่งมีก็คือ SPLC1_S100480 จัดอยู่ในกลุ่มของ XRE (Xenobiotic Response Element) family transcriptional regulator ซึ่งเป็นกลุ่มโปรตีน (Protein Family) ที่มีขนาดใหญ่ helix-turn-helix DNA-binding motif ที่คล้ายคลึงกับของ CI repressor และ Cro proteins ของ λ bacteriophage โปรตีนดังกล่าวทำให้เกิด DNA สมผัสที่โดเมนด้าน N-terminator แต่ความสามารถในการจับกับ DNA ถูกควบคุมโดยโดเมนด้าน C-terminator (Durante-Rodriguez et al., 2010)

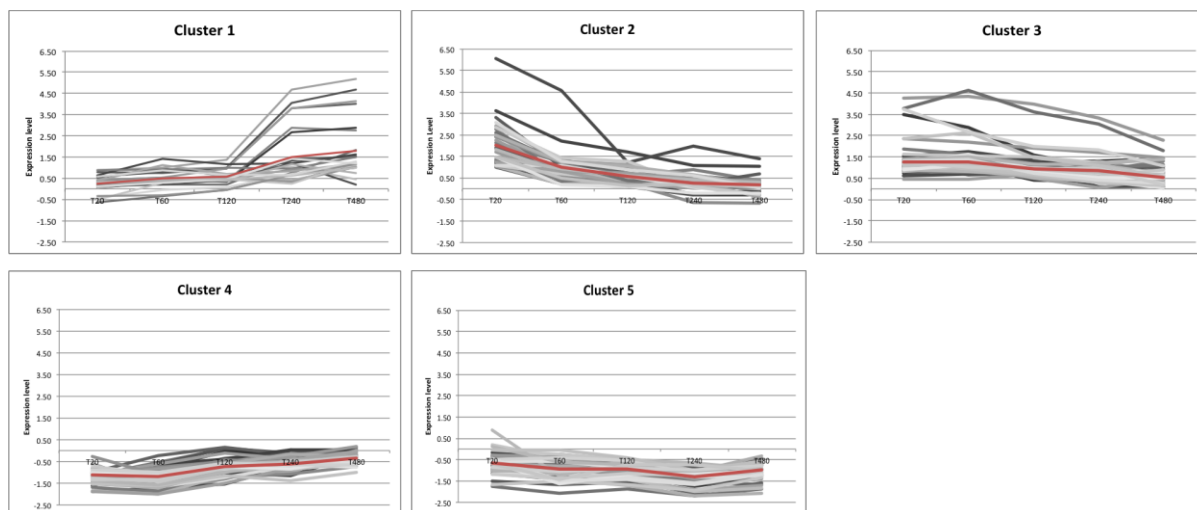
4.2 กลุ่มยีนที่มีการแสดงออกร่วมกันภายใต้สภาวะอุณหภูมิสูง

จากการจัดกลุ่มยีนด้วยอัลกอริทึม K-means Clustering ด้วยมาตรวัดแบบ Cosine Similarity และค่า $K=5$ ทำให้ได้ยีนที่มีการแสดงออกร่วมกัน 5 กลุ่ม โดยยีนมีการกระจายในกลุ่มต่างๆ ดังตารางที่ 4.1 เราได้มีการสำรวจรูปแบบการแสดงออกของยีนในแต่ละกลุ่ม โดยใช้กราฟเส้นแสดงรูปแบบการแสดงออกของยีน ดังรูปที่ 4.1 พบว่ากลุ่มที่ 1, 2 และ 3 มีการแสดงออกแบบเพิ่มขึ้นเมื่อเทียบกับสภาวะปกติ (Over Expression) โดยที่กลุ่มที่ 1 มีการเปลี่ยนแปลงน้อยในช่วง Immediate Response (20-120 นาที) และเพิ่มขึ้นในช่วง Long-term Response (120-480 นาที) ยีนที่น่าสนใจในกลุ่มนี้ได้แก่ SPLC1_S100480; Transcriptional regulator XRE family, SPLC1_S300580; AAA ATPase central domain protein, SPLC1_S310230; 2-alkenal reductase, SPLC1_S53161; methyltransferase type 11 และ SPLC1_S541980; Sulfotransferase ซึ่ง SPLC1_S100480 ทำหน้าที่เป็น Transcriptional Regulator ตัวเดียวเท่านั้นที่แสดงออกในสภาวะอุณหภูมิสูง กลุ่มที่ 2 มีการแสดงออกมากในช่วง Immediate Response และลดลงในช่วง Longterm Response ยีนที่เกี่ยวข้องกับการตอบสนองต่ออุณหภูมิสูงหลายยีน ได้แก่ SPLC1_S010870; molecular chaperone DnaK, SPLC1_S170980; Chaperone DNAj domain protein, SPLC1_S171000; Small heat shock protein molecular chaperone, SPLC1_S310010; ATP-dependent chaperone ClpB, และ SPLC1_S630200; Histidine kinase ซึ่งอยู่ใน Two-component regulatory system ทำหน้าที่ในการรับสัญญาณจากสิ่งแวดล้อมภายนอกชนิดต่างๆ และตอบสนองต่อแวดล้อมที่เปลี่ยนแปลงไป ซึ่งจะทำงานร่วมกับ Response Regulator อีกตัวหนึ่ง กลุ่มที่ 3 มีการแสดงออกใกล้เคียงกันตลอดช่วงเวลา (20-480 นาที) ยกเว้นยีน SPLC1_S170410; Chaperonin Cpn10 และ SPLC1_S170420; Chaperonin GroEL มีการแสดงออกที่สูงมากในช่วง Immediate Response เพราะยีนทั้งสองเกี่ยวข้องกับการปรับตัวต่อการเกิดสภาวะอุณหภูมิสูง และกลุ่มที่ 4 และ 5 มีการแสดงออกแบบ (Under Expression) ตลอดช่วงเวลา (20-480 นาที) โดยที่ กลุ่มที่ 4 ค่อยๆ เพิ่มค่าการแสดงออกขึ้นทีละ

น้อย ในกลุ่มนี้มียีนที่เกี่ยวข้องกับ nitrate ABC transporter ที่แสดงออกลดลงอย่างมีนัยสำคัญถึง 4 ยีน ซึ่งอยู่ใน Operon เดียวกัน ได้แก่ SPLC1_S412180; ABC-type nitrate/nitrite transport system substrate-binding protein, SPLC1_S412190; nitrate ABC transporter inner membrane subunit, SPLC1_S412200; nitrate ABC transporter ATPase subunits C and D, และ SPLC1_S412210; nitrate ABC transporter ATPase subunits C and D ในขณะที่กลุ่มที่ 5 ค่อนข้างน้อย ค่าการแสดงออกของยีนที่ลดลง ยีนที่น่าสนใจในกลุ่มที่ 5 คือ SPLC1_S240060; Ribulose-bisphosphate carboxylase, SPLC1_S240070; Chaperonin family protein RbcX, SPLC1_S240080; Ribulose-bisphosphate carboxylase ซึ่งอยู่ใน Operon เดียวกัน

ตารางที่ 4.1 ผลการจัดกลุ่มยีนด้วยวิธีการจัดกลุ่มแบบ K-means

กลุ่ม	จำนวนยีน
1	26
2	32
3	83
4	57
5	52



รูปที่ 4.1 รูปแบบการแสดงออกของยีนในแต่ละกลุ่ม เส้นสีเทาคือรูปแบบการแสดงออกของยีน และเส้นสีแดงคือค่าเฉลี่ยของการแสดงออกสำหรับทุกยีนในแต่ละกลุ่ม

4.3 Operon ของ *A. platensis* C1 สำหรับยีนที่ตอบสนองต่อสภาวะอุณหภูมิสูง

สิ่งมีชีวิตแบบโพรคาริโอตมีการอาศัยอยู่ในสิ่งแวดล้อมที่มีการเปลี่ยนแปลงอยู่เสมอ ซึ่งปริมาณอาหารมักเพิ่มขึ้นหรือลดลงในบางครั้ง สิ่งมีชีวิตเหล่านี้จำเป็นต้องตอบสนองต่อการเปลี่ยนแปลงของสิ่งแวดล้อมด้วยการเปลี่ยนระดับการแสดงออกของยีน ซึ่งการเปลี่ยนแปลงระดับการแสดงออกของยีนจำเป็นต้องใช้แหล่งคาร์บอนและสารอาหารอื่นๆ ที่จำเป็นต่อการแสดงออกของยีน การสังเคราะห์โปรตีนใหม่จึงถือเป็นการสิ้นเปลืองพลังงานธรรมชาติจึงออกแบบให้ยีนจำนวนหนึ่งในสิ่งมีชีวิตแบบโพรคาริโอตมีการจัดเรียงตัวแบบเป็นกลุ่มหรือที่เรียกว่า

Operon ซึ่งยีนที่อยู่ใน Operon เดียวกันมักจะมียีนที่เกี่ยวข้องกัน จึงควรที่จะแสดงออกมาพร้อมกัน เช่น Lac Operon (Ralston A. 2008)

A. platensis C1 นับเป็นสิ่งมีชีวิตแบบโพรคาริโอตชนิดหนึ่ง ยีนจึงมีการจัดเรียงตัวแบบ Operon เพื่อการใช้ทรัพยากรอย่างคุ้มค่า ในการวิเคราะห์หา Motif บริเวณโปรโมเตอร์ของ *A. platensis* C1 เราจึงต้องศึกษาโครงสร้างของยีนก่อน หากยีนมีการจัดเรียงตัวแบบ Operon ยีนที่อยู่หน้าสุดเท่านั้นจึงจะถูกนำมาวิเคราะห์หา Motif เพราะ Operon จะถูกควบคุมที่ยีนตัวหน้าสุดเท่านั้น จากยีนที่แสดงออกแตกต่างกันอย่างมีนัยสำคัญภายใต้สภาวะอุณหภูมิสูง จำนวน 250 ยีน มียีนที่มีโครงสร้างแบบ Operon ทั้งหมด 14 Operon ซึ่งมีจำนวน 48 ยีน ดังตารางที่ 4.2 โดยยีนที่จะมีโครงสร้างเป็นแบบ Operon ต้องผ่านเงื่อนไขดังนี้

- (1) กลุ่มยีนที่เป็น Operon จะต้องมียะห่างระหว่างยีนไม่เกิน 100 เบส
- (2) กลุ่มยีนที่เป็น Operon เป็นยีนมีหน้าที่เกี่ยวข้องกัน
- (3) กลุ่มยีนที่เป็น Operon อยู่บนสาย Transcript เดียวกัน โดยสังเกตจากข้อมูล RNA-seq

ตารางที่ 4.2 ผลการทำนาย Operon ของ *A. platensis* C1 สำหรับยีนที่ตอบสนองต่อสภาวะอุณหภูมิสูง

Operon	Gene	Co-transcription	Strand
1	SPLC1_S100410 (hypothetical protein) SPLC1_S100420 (hypothetical protein) SPLC1_S100430 (hypothetical protein) SPLC1_S100450 (hypothetical protein) SPLC1_S100460 (hypothetical protein) SPLC1_S100470 (hypothetical protein)	NA NA NA NA NA NA	-
2	SPLC1_S170920 (putative short-chain dehydrogenase/reductase SDR) SPLC1_S170940 (hypothetical protein) SPLC1_S170960 (peptidase S1 and S6 chymotrypsin/Hap) SPLC1_S170970 (hypothetical protein) SPLC1_S170980 (chaperone DnaJ domain protein)	NA NA Co-transcription Co-transcription Co-transcription	-
3	SPLC1_S360590 (sulfate ABC transporter periplasmic sulfate-binding protein) SPLC1_S360600 (putative magnesium transporter) SPLC1_S360610 (hypothetical protein) SPLC1_S360620 (MgtC/SapB transporter)	Co-transcription Co-transcription Co-transcription Co-transcription	+
4	SPLC1_S170410 (chaperonin Cpn10) SPLC1_S170420 (chaperonin GroEL)	Co-transcription Co-transcription	+
5	SPLC1_S271220 (peptidase M42 family protein) SPLC1_S271230 (hypothetical protein)	Co-transcription Co-transcription	-
6	SPLC1_S410610 (pentapeptide repeat protein) SPLC1_S410620 (hypothetical protein)	NA NA	-

7	SPLC1_S532900 (hypothetical protein) SPLC1_S532910 (serine/threonine protein kinase)	NA NA	-
8	SPLC1_S012010 (cyanate lyase) SPLC1_S012020 (hypothetical protein) SPLC1_S012030 (hypothetical protein) SPLC1_S012040 (nitrate ABC transporter ATPase subunits C and D)	Co-transcription Co-transcription Co-transcription Co-transcription	-
9	SPLC1_S060120 (tRNA pseudouridine synthase A) SPLC1_S060130 (ribosomal protein L17)	Co-transcription Co-transcription	-
10	SPLC1_S412180 (ABC-type nitrate/nitrite transport system substrate-binding protein) SPLC1_S412190 (nitrate ABC transporter inner membrane subunit) SPLC1_S412200 (nitrate ABC transporter ATPase subunits C and D) SPLC1_S412210 (nitrate ABC transporter ATPase subunits C and D) SPLC1_S412220 (hypothetical protein) SPLC1_S412240 (hypothetical protein) SPLC1_S412250 (hypothetical protein) SPLC1_S412260 (hypothetical protein)	Co-transcription Co-transcription Co-transcription Co-transcription Co-transcription Co-transcription Co-transcription	+
11	SPLC1_S240060 (ribulose-bisphosphate carboxylase) SPLC1_S240070 (chaperonin family protein RbcX) SPLC1_S240080 (ribulose-bisphosphate carboxylase)	NA NA NA	+
12	SPLC1_S420050 (hypothetical protein) SPLC1_S420060 (hypothetical protein) SPLC1_S420070 (subtilisin-like protein)	NA NA NA	-
13	SPLC1_S510240 (translation elongation factor Ts) SPLC1_S510250 (ribosomal protein S2)	NA NA	-
14	SPLC1_S540710 (glucosylglycerol-phosphate synthase) SPLC1_S540720 (FAD dependent oxidoreductase)	Co-transcription Co-transcription	+

4.4 การวิเคราะห์การแสดงออกร่วมกัน

เมื่อพิจารณาการวัดความสัมพันธ์ของยีนด้วย Pearson's Correlation Coefficient (PCC) โดยใช้ $PCC > 0.90$ พบว่ามียีนทั้งหมด 247 ยีนด้วยความสัมพันธ์เชิงบวก 2,839 ความสัมพันธ์ และพบว่ายีนที่มีความสัมพันธ์หรือมีการแสดงออกร่วมกับยีนอื่นมากที่สุด คือ SPLC1_S030430 ซึ่งทำหน้าที่เป็น Nitrite Transporter โดยยีนเหล่านี้มีเพียงยีนเดียวเท่านั้นที่ทำหน้าที่เป็น Transcriptional Regulator คือ SPLC1_S100480

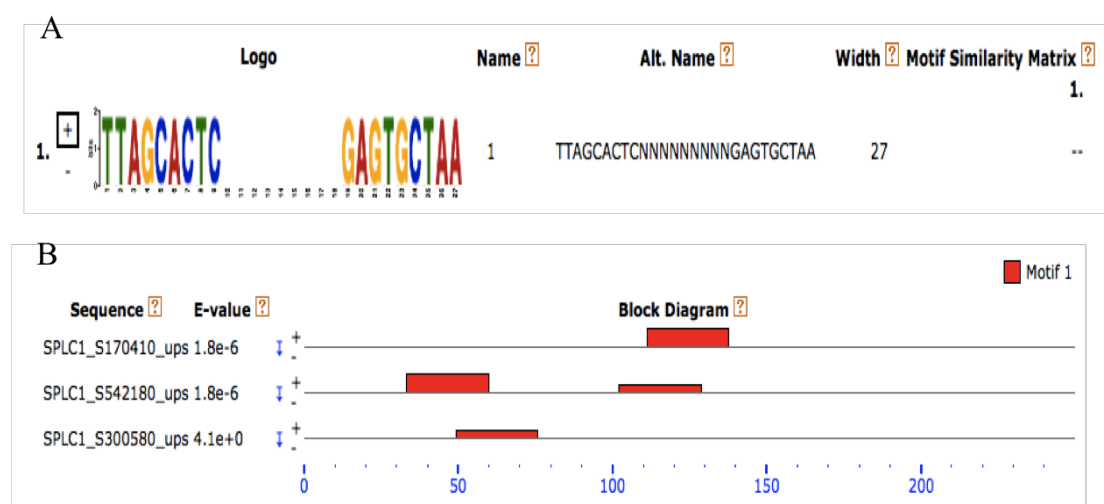
4.5 การค้นหา *cis*-regulatory element บริเวณหน้ายีนที่แสดงออกอย่างมีนัยสำคัญภายใต้สภาวะอุณหภูมิสูง

การค้นหา *cis*-regulatory element แบ่งออก 2 ส่วน คือ (1) การค้นหา known *cis*-regulatory element และ (2) การค้นหา novel *cis*-regulatory element

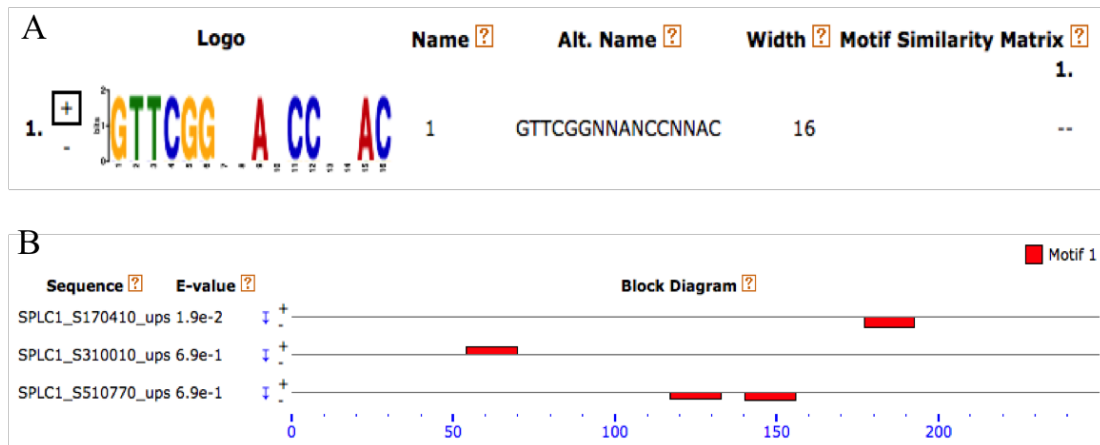
4.5.1 การค้นหา Known Motif บริเวณหน้ายีนที่แสดงออกอย่างมีนัยสำคัญภายใต้สภาวะอุณหภูมิสูง

จากการศึกษาก่อนหน้านี้พบว่า CIRCE, K-box, N-box และ H-box เป็นโมทีฟที่เกี่ยวข้องกับสภาวะอุณหภูมิสูงโดยที่ CIRCE มีรูปแบบ TTAGCACTC-N9-GAGTGCTAA ซึ่งกระตุ้น Operon *groESL* และยีน *cpn60* (*groEL-2*), K-box มีรูปแบบ GTTCGG-NNAN-CCNNAC ซึ่งกระตุ้นยีน *groEL-1*, *groEL-2*, *dnaK* และ Operon *groESL*, N-box กระตุ้น *groEL-1* และ *groEL-2* และ H-box กระตุ้น Operon *groESL* (Rajaram et al., 2014)

จากการค้นหา Motif ด้วยเครื่องมือ MAST (Motif Alignment Search Tool) ซึ่งอยู่ใน MEME Suite พบว่า CIRCE ถูกพบบริเวณหน้ายีน 250 เบสของยีน SPLC1_S170410 (*cpn10*), SPLC1_S542180 (*groEL-2*) และ SPLC1_S300580 (AAA ATPase central domain protein) ดังรูปที่ 4.2 เมื่อค้นหา K-box บริเวณหน้ายีนของยีนที่มีการแสดงออกอย่างมีนัยสำคัญ พบว่ามี 3 ยีน ได้แก่ SPLC1_S170410 (*cpn10*), SPLC1_S310010 (*clpB*) และ SPLC1_S510770 (phage shock protein C PspC) ดังรูปที่ 4.3 และเมื่อค้นหา N-box บริเวณหน้ายีน 250 ยีนที่มีการแสดงออกอย่างมีนัยสำคัญ ไม่พบ N-box ปรากฏบนยีนใดเลย



รูปที่ 4.2 ผลการทำนาย CIRCE element บริเวณหน้ายีนที่แสดงออกอย่างมีนัยสำคัญภายใต้สภาวะอุณหภูมิสูง (A) Sequence Logo ของ CIRCE element (B) ตำแหน่งของ CIRCE element ที่พบบนยีนที่มีการแสดงออกอย่างมีนัยสำคัญภายใต้สภาวะอุณหภูมิสูง



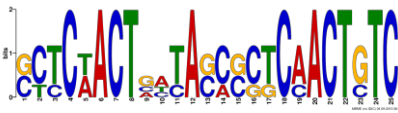
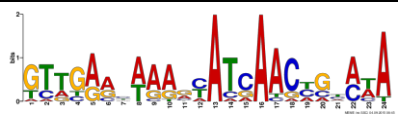
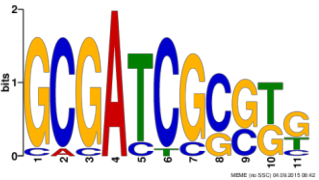
รูปที่ 4.3 ผลการทำนาย K-box element บริเวณหน้ายีนที่แสดงออกอย่างมีนัยสำคัญภายใต้สภาวะอุณหภูมิสูง (A) Sequence Logo ของ K-box element (B) ตำแหน่งของ K-box element ที่พบบนยีนที่มีการแสดงออกอย่างมีนัยสำคัญภายใต้สภาวะอุณหภูมิสูง

4.5.2 Novel *cis*-regulatory element ที่ทำนายได้ซึ่งเกี่ยวข้องกับการควบคุมยีนในสภาวะอุณหภูมิสูง

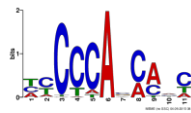
การวิเคราะห์หา Motif ด้วยเครื่องมือค้นหาชื่อ MEME Suite (Bailey et al., 2009) โดยค้นหา Motif ที่มีขนาด 6-25 เบส และเลือก Motif ที่มีคะแนนดีที่สุดในกลุ่มละ 3 ชนิด และใช้เครื่องมือชื่อ TOMTOM (Gupta et al., 2007) ในการทำนาย Regulator ที่สามารถจับกับ Motif เหล่านี้ได้ โดยที่ TOMTOM (Gupta et al., 2007) จะทำการเปรียบเทียบ Motif กับ *cis*-regulatory element ซึ่งอยู่ในฐานข้อมูลที่รวบรวม *cis*-regulatory element ของโปรคาริโอต ในงานนี้ได้ใช้ฐานข้อมูล Prodoric เวอร์ชัน 8.9 และ TOMTOM (Gupta et al., 2007) จะเรียงลำดับ Motif และสร้าง Alignment ของ Motif ที่ทำนายได้นี้ อาจจะเกี่ยวข้องกับการควบคุมยีนในสภาวะอุณหภูมิสูง

จากการค้นหาโมทีฟที่มีการพบลำดับเบสแบบ Palindrome เป็นจำนวนมาก ซึ่งมีลำดับเบสดังนี้ GCGATCGC ซึ่งเรียกว่า Highly Iterated Palindrome (HIP1) พบมากในไฮยาโนแบคทีเรีย ซึ่งการศึกษาก่อนหน้านี้พบว่า ลำดับเบสนี้มีความจำเพาะต่อเอนไซม์ DNA Methyltransferase (Elhai, 2015) ตารางที่ 4.4 แสดงผลการทำนาย Regulator ที่สามารถจับกับ Motif ที่ทำนายได้ PrrA เป็น Regulator ที่พบมากที่สุด ซึ่งทำหน้าที่เป็น Putative response regulator ซึ่งเกี่ยวข้องกับการควบคุมอีกซิเจนในการแสดงออกของยีนในกระบวนการสังเคราะห์แสงของ (Eraso and Kaplan, 1993)

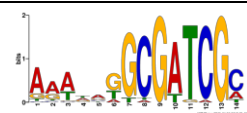
ตารางที่ 4.3 ผลการทำนายโมทีฟของกลุ่มยีนและรูปแบบของโมทีฟ

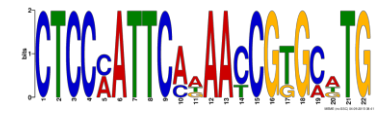
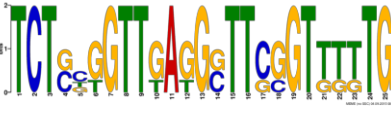
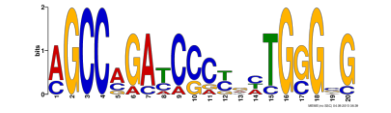
Motif	Gene	Pattern
Motif 1 E-value = 8.9e+001 Sites = 6 Width = 25	SPLC1_S100480 transcriptional regulator XRE family SPLC1_S100470 hypothetical protein SPLC1_S170170 hypothetical protein SPLC1_S300580 AAA ATPase central domain protein SPLC1_S202720 hypothetical protein SPLC1_S531610 methyltransferase type 11	
Motif 2 E-value = 1.1e+002 Sites = 4 Width = 25	SPLC1_S100480 transcriptional regulator XRE family SPLC1_S100470 hypothetical protein SPLC1_S541660 hypothetical protein SPLC1_S100710 hypothetical protein	
Motif 3 E-value = 1.5e+002 Sites = 13 Width = 24	SPLC1_S100480 transcriptional regulator XRE family SPLC1_S100470 hypothetical protein SPLC1_S170160 hypothetical protein SPLC1_S033870 hypothetical protein SPLC1_S100440 hypothetical protein SPLC1_S260450 hypothetical protein SPLC1_S300490 hypothetical protein SPLC1_S541640 hypothetical protein SPLC1_S310230 2-alkenal reductase SPLC1_S207690 hypothetical protein SPLC1_S541980 sulfotransferase SPLC1_S300600 hypothetical protein SPLC1_S170170 hypothetical protein	
Motif 4 E-value = 3.2e-007 Sites = 14 Width = 11	SPLC1_S521030 hypothetical protein SPLC1_S360230 hypothetical protein SPLC1_S260670 hypothetical protein SPLC1_S103160 protein of unknown function DUF1825 SPLC1_S031870 hypothetical protein SPLC1_S061460 hypothetical protein SPLC1_S080750 FAD dependent oxidoreductase SPLC1_S208890 hypothetical protein SPLC1_S032910 hypothetical protein SPLC1_S360590 sulfate ABC transporter periplasmic sulfate-binding protein SPLC1_S360940 hypothetical protein SPLC1_S030970 FeS assembly protein SufB SPLC1_S171000 small heat shock protein molecular chaperone	

	SPLC1_S260330 SCP-like extracellular SPLC1_S010880 hypothetical protein SPLC1_S208890 hypothetical protein SPLC1_S500630 hypothetical protein SPLC1_S260670 hypothetical protein SPLC1_S310330 SCP-like extracellular SPLC1_S360940 hypothetical protein SPLC1_S360590 sulfate ABC transporter periplasmic sulfate-binding protein SPLC1_S010890 hypothetical protein SPLC1_S061460 hypothetical protein	
Motif 5 E-value = 9.2e-002 Sites = 21 Width = 11	SPLC1_S010880 hypothetical protein SPLC1_S208890 hypothetical protein SPLC1_S500630 hypothetical protein SPLC1_S260670 hypothetical protein SPLC1_S310330 SCP-like extracellular SPLC1_S360940 hypothetical protein SPLC1_S360590 sulfate ABC transporter periplasmic sulfate-binding protein SPLC1_S010890 hypothetical protein SPLC1_S061460 hypothetical protein	
Motif 6 E-value = 4.7e+001 Sites = 4 Width = 25	SPLC1_S061460 hypothetical protein SPLC1_S310020 hypothetical protein SPLC1_S310010 ATP-dependent chaperone ClpB SPLC1_S310330 SCP-like extracellular	
Motif 7 E-value = 9.4e-024 Sites = 26 Width = 11	SPLC1_S560190 ATP-binding region ATPase domain protein SPLC1_S501000 DNA polymerase III delta prime subunit SPLC1_S190200 hypothetical protein SPLC1_S430180 hypothetical protein SPLC1_S032780 hypothetical protein SPLC1_S530170 hypothetical protein SPLC1_S203590 hypothetical protein SPLC1_S080980 thioredoxin domain protein SPLC1_S350020 hypothetical protein SPLC1_S180030 signal peptidase II SPLC1_S033820 hypothetical protein SPLC1_S531220 hypothetical protein SPLC1_S170410 chaperonin Cpn10 SPLC1_S061580 hypothetical protein SPLC1_S130960 hypothetical protein SPLC1_S370210 hypothetical protein SPLC1_S100850 hypothetical protein SPLC1_S300310 2OG-Fe(II) oxygenase SPLC1_S170950 ion transport protein	

	SPLC1_S360330 PHP domain protein SPLC1_S101290 hypothetical protein SPLC1_S411530 CRISPR-associated protein Cas2 SPLC1_S271340 Rho termination factor domain protein SPLC1_S020030 zinc finger SWIM domain protein SPLC1_S206920 KaiA family protein SPLC1_S420330 hypothetical protein	
Motif 8 E-value = 1.6e-003 Sites = 79 Width = 11	SPLC1_S205990 type I restriction modificatio SPLC1_S020030 zinc finger SWIM domain protein SPLC1_S061580 hypothetical protein SPLC1_S530170 hypothetical protein SPLC1_S202930 hypothetical protein SPLC1_S350020 hypothetical protein SPLC1_S205630 helix-hairpin-helix motif protein SPLC1_S380100 hypothetical protein SPLC1_S102560 hypothetical protein SPLC1_S190200 hypothetical protein SPLC1_S411530 CRISPR-associated protein Cas2 SPLC1_S203590 hypothetical protein SPLC1_S271230 hypothetical protein SPLC1_S203860 peptidase S1 and S6 chymotrypsin/Hap SPLC1_S542180 chaperonin GroEL SPLC1_S100100 hypothetical protein SPLC1_S501000 DNA polymerase III delta prime subunit SPLC1_S370210 hypothetical protein SPLC1_S060050 hypothetical protein SPLC1_S011220 hypothetical protein SPLC1_S530280 hypothetical protein SPLC1_S203930 putative peptidase C14 caspase catalytic subunit p20 SPLC1_S360330 PHP domain protein SPLC1_S206920 KaiA family protein SPLC1_S032780 hypothetical protein SPLC1_S032770 hypothetical protein SPLC1_S180030 signal peptidase II SPLC1_S100850 hypothetical protein SPLC1_S080980 thioredoxin domain protein SPLC1_S500430 hypothetical protein SPLC1_S260460 hypothetical protein SPLC1_S560190 ATP-binding region ATPase domain protein SPLC1_S203440 hypothetical protein	

	SPLC1_S430180 hypothetical protein SPLC1_S010050 pentapeptide repeat protein SPLC1_S480030 hypothetical protein SPLC1_S430350 hypothetical protein SPLC1_S300780 hybrid cluster protein SPLC1_S271340 Rho termination factor domain protein SPLC1_S270730 phospholipid/glycerol acyltransferase SPLC1_S260350 hypothetical protein SPLC1_S300740 hypothetical protein SPLC1_S011870 hypothetical protein SPLC1_S531220 hypothetical protein SPLC1_S208490 cobalamin biosynthesis CbiX protein SPLC1_S041040 hypothetical protein SPLC1_S532550 MORN repeat-containing protein SPLC1_S207680 hypothetical protein SPLC1_S541200 S-(hydroxymethyl)glutathione dehydrogenase/class III alcoholdehydrogenase SPLC1_S532910 serine/threonine protein kinase SPLC1_S532860 serine/threonine protein kinase SPLC1_S081950 alanine racemase SPLC1_S520310 hypothetical protein SPLC1_S410620 hypothetical protein SPLC1_S171060 hypothetical protein SPLC1_S170410 chaperonin Cpn10 SPLC1_S160040 hypothetical protein SPLC1_S511030 hypothetical protein SPLC1_S102540 hypothetical protein SPLC1_S061630 CheW protein SPLC1_S130960 hypothetical protein SPLC1_S270010 hypothetical protein SPLC1_S207110 amino acid permease-associated region SPLC1_S206560 phosphoenolpyruvate synthase SPLC1_S300180 type III restriction protein res subunit SPLC1_S208530 hypothetical protein SPLC1_S100690 hypothetical protein SPLC1_S033820 hypothetical protein SPLC1_S102990 ATP-dependent Clp protease adaptor protein ClpS SPLC1_S220190 hypothetical protein SPLC1_S103050 AAA ATPase central region SPLC1_S300310 2OG-Fe(II) oxygenase	
--	---	--

	SPLC1_S082630 hypothetical protein SPLC1_S051190 glutaminase SPLC1_S230980 rhodanese domain protein SPLC1_S210020 zinc finger SWIM domain protein SPLC1_S170950 ion transport protein SPLC1_S420330 hypothetical protein SPLC1_S101290 hypothetical protein	
Motif 9 E-value = 2.4e+000 Sites = 5 Width = 25	SPLC1_S207110 amino acid permease-associated region SPLC1_S206560 phosphoenolpyruvate synthase SPLC1_S171060 hypothetical protein SPLC1_S041040 hypothetical protein SPLC1_S500430 hypothetical protein	
Motif 10 E-value = 7.1e-019 Sites = 27 Width = 14	SPLC1_S030450 hypothetical protein SPLC1_S540730 NAD-dependent glycerol-3-phosphate dehydrogenase domain protein SPLC1_S370770 hypothetical protein SPLC1_S260970 hypothetical protein SPLC1_S361030 putative RNA methylase SPLC1_S130480 hypothetical protein SPLC1_S410830 twin-arginine translocation protein TatA/E family subunit SPLC1_S540700 hypothetical protein SPLC1_S207280 hypothetical protein SPLC1_S010400 dTDP-4-dehydrorhamnose reductase SPLC1_S100290 urea ABC transporter ATP-binding protein UrtD SPLC1_S207080 hypothetical protein SPLC1_S101570 2-alkenal reductase SPLC1_S542730 histidyl-tRNA synthetase SPLC1_S541060 ABC transporter related protein SPLC1_S080430 hypothetical protein SPLC1_S060350 ribosomal protein L2 SPLC1_S207140 hypothetical protein SPLC1_S103260 adenylosuccinate synthase SPLC1_S540640 hypothetical protein SPLC1_S201370 4-alpha-glucanotransferase SPLC1_S541290 putative GTP cyclohydrolase I SPLC1_S080420 hypothetical protein SPLC1_S060130 ribosomal protein L17 SPLC1_S033220 PBS lyase HEAT domain protein repeat-containing protein SPLC1_S080680 30S ribosomal protein S6 SPLC1_S541170 hypothetical protein	

Motif 11 E-value = 7.3e-006 Sites = 7 Width = 24	SPLC1_S204920 hypothetical protein SPLC1_S040090 group II intron maturase-specific domain protein SPLC1_S541170 hypothetical protein SPLC1_S080680 30S ribosomal protein S6 SPLC1_S540640 hypothetical protein SPLC1_S542730 histidyl-tRNA synthetase SPLC1_S010400 dTDP-4-dehydrorhamnose reductase	
Motif 12 E-value = 1.2e+001 Sites = 4 Width = 22	SPLC1_S040090 group II intron maturase-specific domain protein SPLC1_S080680 30S ribosomal protein S6 SPLC1_S204920 hypothetical protein SPLC1_S540730 NAD-dependent glycerol-3-phosphate dehydrogenase domain protein	
Motif 13 E-value = 3.2e+000 Sites = 4 Width = 25	SPLC1_S033770 hypothetical protein SPLC1_S033760 hypothetical protein SPLC1_S200530 hypothetical protein SPLC1_S180040 BioY protein	
Motif 14 E-value = 3.3e+001 Sites = 4 Width = 25	SPLC1_S543170 hypothetical protein SPLC1_S011760 HNH endonuclease SPLC1_S080380 preprotein translocase SecG subunit SPLC1_S550010 group II intron maturase-specific domain protein	
Motif 15 E-value = 4.4e+001 Sites = 7 Width = 20	SPLC1_S033770 hypothetical protein SPLC1_S033760 hypothetical protein SPLC1_S102800 hypothetical protein SPLC1_S180040 BioY protein SPLC1_S240070 chaperonin family protein RbcX SPLC1_S510770 phage shock protein C PspC SPLC1_S200700 50S ribosomal protein L10	

ตารางที่ 4.4 ผลการทำนาย Regulator ที่อาจจะจับกับ Motif

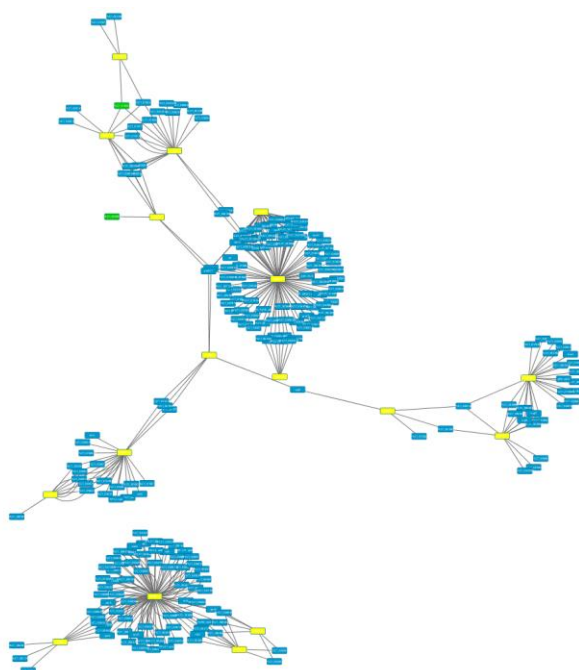
Motif	Predicted Regulators
Motif 1	MX000208 (PrrA), MX000021 (Spo0A), MX000097 (Mlc), MX000046 (ComA), MX000008 (GlpR), MX000033 (AlgU), MX000011 (RhIR), MX000186 (AlgZ), MX000183 (TrpR), MX000109 (TrpI)
Motif 2	MX000002 (Anr), MX000102 (Dnr), MX000101 (AlgR), MX000137 (LacI), MX000190 (FNRL), MX000155 (OxyR), MX000152 (GlnG), MX000133 (GntR), MX000107 (OxyR), MX000022 (CRE)
Motif 3	MX000029 (Fis), MX000031 (PhoP), MX000058 (PucR), MX000178 (TorR),

	MX000193 (KdgR), MX000043 (AhrC), MX000009 (ResD), MX000116 (ArgR), MX000088 (Xre), MX000089 (XylR)
Motif 4	MX000208 (PrrA), MX000107 (OxyR), MX000171 (RhaS), MX000184 (YiaJ), MX000101 (AlgR), MX000022 (CRE), MX000097 (Mlc), MX000159 (Crp)
Motif 5	MX000197 (NatA), MX000176 (YdiH), MX000074 (SigB), MX000139 (MalT), MX000009 (ResD), MX000117 (CspA), MX000019 (Spo0A), MX000039 (RofA), MX000167 (DegU), MX000190 (FNRL)
Motif 6	MX000088 (Xre), MX000143 (GcvA), MX000139 (MalT), MX000059 (PurR), MX000076 (SigH), MX000179 (XapR), MX000166 (RcsB), MX000178 (TorR)
Motif 7	MX000208 (PrrA), MX000184 (YiaJ), MX000186 (AlgZ), MX000063 (SenS), MX000107 (OxyR), MX000192 (NnrR), MX000171 (RhaS), MX000185 (XylR)
Motif 8	MX000039 (RofA), MX000001 (MexR), MX000068 (SigE), MX000186 (AlgZ), MX000197 (NatA), MX000034 (AlgU), MX000152 (GlnG)
Motif 9	MX000208 (PrrA), MX000175 (SdiA), MX000186 (AlgZ), MX000115 (AraC), MX000046 (ComA), MX000177 (TreR), MX000197 (NatA), MX000139 (MalT), MX000164 (Lrp), MX000021 (Spo0A)
Motif 10	MX000107 (OxyR), MX000184 (YiaJ), MX000185 (XylR), MX000114 (Ada), MX000047 (GerE), MX000077 (SigH), MX000110 (Vfr), MX000208 (PrrA), MX000168 (DegU), MX000186 (AlgZ)
Motif 11	MX000111 (RpoN), MX000036 (AlgU), MX000186 (AlgZ), MX000162 (Mlc), MX000104 (FleQ), MX000143 (GcvA), MX000033 (AlgU), MX000208 (PrrA), MX000042 (RegR), MX000194 (SigI)
Motif 12	MX000117 (CspA), MX000082 (SigW), MX000178 (TorR), MX000074 (SigB), MX000043 (AhrC), MX000105 (GlpR), MX000163 (Lrp), MX000038 (CovR), MX000172 (RhaR)
Motif 13	MX000031 (PhoP), MX000014 (OmpR), MX000103 (ExsA), MX000185 (XylR), MX000166 (RcsB), MX000164 (Lrp), MX000180 (HydG), MX000172 (RhaR), MX000171 (RhaS)
Motif 14	MX000118 (CynR), MX000080 (SigW), MX000083 (SigX), MX000164 (Lrp), MX000123 (DeoR), MX000035 (AlgU), MX000184 (YiaJ), MX000194 (SigI)
Motif 15	MX000111 (RpoN), MX000180 (HydG), MX000208 (PrrA), MX000042 (RegR), MX000007 (ArgR), MX000152 (GlnG), MX000055 (Mta)

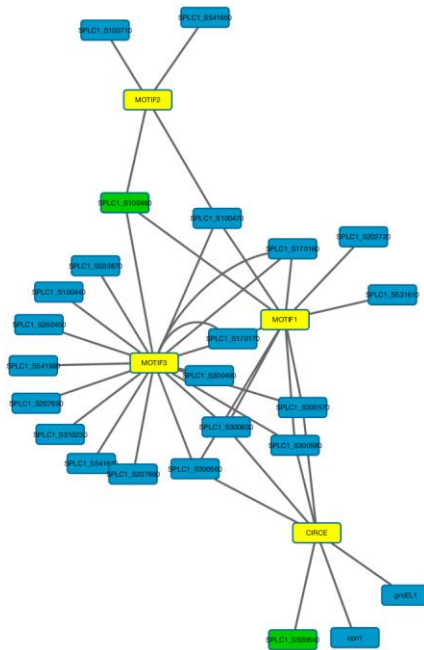
4.6 เครือข่ายควบคุมยีนของ *A. platensis* C1 ภายใต้สภาวะอุณหภูมิสูง

เครือข่ายการควบคุมการแสดงออกของยีน (Gene Regulatory Network) ประกอบด้วยยีนและความสัมพันธ์ของยีน ซึ่งสามารถใช้สัญลักษณ์แทนด้วย node และ edge ตามลำดับ โดยยีนที่มีความสัมพันธ์กันจะมี edge เชื่อมต่อกัน กลไกการควบคุมยีนเกิดขึ้นได้ในหลายขั้นตอน ของการแสดงออกของยีน (Gene Expression) ได้แก่ การถอดรหัส (Transcription) การเปลี่ยนแปลงหลังการถอดรหัส (Post-transcriptional Modification) การขนส่ง RNA (RNA Transport) การแปลรหัส (Translation) การย่อยสลายสาย mRNA (mRNA Degradation) และการเปลี่ยนแปลงหลังการแปลรหัส (Post-translation Modification)

จากการทำนาย novel และ known *cis*-regulatory element บนยีนต่างๆ ที่ตอบสนองต่อสภาวะอุณหภูมิสูง เราสามารถนำข้อมูลดังกล่าวมาสร้างเป็นเครือข่ายการควบคุมยีน ดังรูปที่ 4.4 โดยในเครือข่ายนี้มี Regulator สองชนิดคือ SPLC1_S100480 ซึ่งทำหน้าที่เกี่ยวกับ XRE Transcriptional Regulator และ SPLC1_S500640 ซึ่งทำหน้าที่เป็น heat-inducible transcription repressor HrcA อย่างไรก็ตาม SPLC1_S100480 มีการเปลี่ยนแปลงอย่างมีนัยสำคัญตามเวลาภายใต้สภาวะอุณหภูมิสูง ในขณะที่ HrcA ไม่มีการเปลี่ยนแปลงอย่างมีนัยสำคัญ พบว่า SPLC1_S100480 อาจจะควบคุมยีนหลายชนิดด้วย Motif สามแบบด้วยกัน ได้แก่ Motif1, Motif2 และ Motif3 ดังตารางที่ 4.3 ซึ่งอาจจะควบคุมยีนหลายชนิด สำหรับ HrcA สามารถควบคุมโอเปอรอน *cpn10* และ *groEL1* ผ่าน CIRCE element นอกจากนี้ HrcA ยังสามารถควบคุมโอเปอรอน SPLC1_S300560, SPLC1_S300570, SPLC1_S300580 และ SPLC1_S300600 ซึ่งเกี่ยวข้องกับการ encode ให้เอนไซม์ AAA ATPase ผ่าน CIRCE element เช่นกัน โดยเมื่อสังเกตการแสดงออกของยีน HrcA พบว่ามีการแสดงออกลดลงเล็กน้อย ซึ่ง HrcA ทำหน้าที่ในการยับยั้งการแสดงออกของยีน เมื่อ HrcA มีการแสดงออกลดลงจึงทำให้ยีนที่ถูกควบคุมด้วย HrcA มีการแสดงออกเพิ่มขึ้นได้



รูปที่ 4.4 เครือข่ายควบคุมยีนของ *A. platensis* C1 ภายใต้สภาวะอุณหภูมิสูง โดยที่ Node สีฟ้า สีเขียวและสีเหลือง คือ gene, regulator และ cis-regulatory element ตามลำดับ ซึ่งประกอบด้วย 250 Node และ 371 ความสัมพันธ์



รูปที่ 4.5 กลไกการทำงานของ Regulator สองชนิดได้แก่ SPLC1_S100480 (XRE Transcriptional Regulator) และ SPLC1_S500640 (HrcA Transcriptional Repressor)

บทที่ 5 บทสรุป

การศึกษากลไกการแสดงออกของยีนภายใต้สภาวะอุณหภูมิสูงของ *A. platensis* C1 ทำให้เราสามารถเพิ่มผลผลิตน้ำหนักรวมของ *A. platensis* C1 และหากเราเข้าใจกลไกการแสดงออกของยีนอย่างลึกซึ้ง ก็จะสามารถใช้ *A. platensis* C1 เป็นแหล่งผลิตสารสำคัญ เช่น แคโรทีนอยด์ ไฟโคไซยานิน และแกรมนมาลินโเลอิกแอซิด (GLA) ในการศึกษาครั้งนี้ เราสามารถทำนายความสัมพันธ์ของยีนจากการแสดงออกร่วมกัน และการมี Motif ร่วมกัน ด้วยสมมติฐานที่ว่า ยีนที่มีการแสดงออกร่วมกัน อาจจะถูกควบคุมด้วย Regulator ตัวเดียวกันและบริเวณโปรโมเตอร์ของยีนเหล่านี้มี Motif ที่คล้ายคลึงกัน ก็มีโอกาสมากขึ้นที่จะถูกควบคุมด้วย Regulator ชนิดเดียวกัน การทำนาย Motif ที่มีรายงานว่าเกี่ยวข้องกับสภาวะอุณหภูมิสูง ซึ่งได้แก่ CIRCE, K-box, N-box และ H-box แสดงให้เห็นว่า cis-regulatory element เหล่านี้อยู่บนยีนที่มีการตอบสนองต่อสภาวะอุณหภูมิสูง และการทำนาย Motif ชนิดใหม่ ทำให้ได้รับ cis-regulatory element ซึ่งอาจใช้สำหรับการทดลองต่อไปใน *A. platensis* C1

นอกจากนี้ยังมีการทำนาย Regulator ที่สามารถจับกับโมทีฟที่ทำนายไว้ โดยอาศัยข้อมูลของสิ่งมีชีวิตแบบโปรคาริโอตชนิดอื่นๆ เราสามารถสร้างเครือข่ายการทำงานของยีนภายใต้สภาวะอุณหภูมิสูงจาก cis-regulatory element ที่ทำนายได้และความสัมพันธ์ของยีนที่เกิดจากการแสดงออกร่วมกันภายใต้สภาวะอุณหภูมิสูง อย่างไรก็ตาม เครือข่ายที่สร้างขึ้นมาเป็นเครือข่ายแบบ Static กล่าวคือ ในสภาวะที่เปลี่ยนแปลงไป การทำงานของ ยีนและ Regulator อาจเปลี่ยนแปลงไปในรูปแบบอื่นได้ และเครือข่ายที่ทำนายได้นี้ จะสามารถนำไปใช้ในการศึกษาความสัมพันธ์และการทำงานร่วมกันของยีน โดยต้องมีการทำการทดลองเพิ่มเพื่อยืนยันความสัมพันธ์ของยีนดังกล่าว ก่อนนำไปประยุกต์ใช้ในงานทางด้านการตัดต่อพันธุกรรมหรือทางด้านเทคโนโลยีชีวภาพต่อไป

เอกสารอ้างอิง

- Arun P, Subhashini M, Santhosh CH, Krishna P, Prakash J. (2013) Prediction of cis regulatory elements in the genome of *Synechococcus Elongatus* PCC 6301. *Computational Systems Biology*. 13: 369-373.
- Bailey TL, Bodén M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, and Noble WS. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research*. 37: W202-W208.
- Bar-Joseph Z, Gerber GK, Lee TI, Rinaldi NJ, Yoo JY, Robert F, Gordon DB, Fraenkel E, Jaakkola TS, Young RA, Gifford DK. (2003) Computational discovery of gene modules and regulatory networks. *Nature Biotechnology*. 21(11): 1337-42.
- Beer MA. and Tavazoie S. (2004) Predicting gene expression from sequence. *Cell*. 117: 185-198.
- Belay A. (2002) The potential application of *Spirulina* (Arthrospira) as a nutritional and therapeutic supplement in health management. *Journal of the American Nutraceutical Association*. 5: 27-48.
- Boutanaev AM, Kalmykova AI, Shevelyov YY, Nurminsky DI. (2002) Large clusters of co-expressed genes in the *Drosophila* genome. *Nature* 420: 666-669.
- Ciferri O. (1983) *Spirulina*, the edible micro-organism. *Microbiology Reviews*. 47: 551-578.
- Ciferri O. and Tiboni O. (1985) The biochemistry and industrial potential of *Spirulina*. *Annual Reviews of Microbiology*. 39: 503-526.
- Cheevadhanarak S, Paithoonrangsarid K, Prommeenate P, Kaewngam W, Musigkain A, Tragoonrung S, Tabata S, Kaneko T, Chaijaruwanich J, Sangsrakru D, Tangphatsomruang S, Chanprasert J, Tongsimma S, Kusonmano K, Jeamton W, Dulsawat S, Klanchui A, Vorapreeda T, Chumchua V, Khannapho C, Thammarongtham C, Plengvidhya V, Subudhi S, Hongsthong A, Ruengjitchatchawalya M, Meechai A, Senachak J, and Tantichareon M. (2012) Draft genome sequence of *Arthrospira platensis* C1 (PCC9438). *Standard in Genomic Sci*. 6(1): 43-53.
- Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, Hannett NM, Harbison CR, Thompson CM, Simon I, Zeitlinger J, Jennings EG, Murray HL, Gordon DB, Ren B, Wyrick JJ, Tagne J, Volkert TL, Fraenkel E, Gifford DK, Young RA. (2002) Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science*. 298: 799-804.
- Dai Z, Dai X. (2011) Nuclear colocalization of transcription factor target genes strengthens coregulation in yeast. *Nucleic Acids Res*.

Dottorini T, Persampieri T, Palladino P, Baker DA, Spaccapelo R, et al. (2013) Regulation of *Anopheles gambiae* male accessory gland genes influences postmating response in female. *FASEB J.* 27: 86-97.

Durante-Rodriguez G, Valderrama JA, Mancheno JM, Rivas G, Alfonso C, Arias-Palomo E, Llorca O, Garcia JL, Carmona M. (2010) Biochemical characterization of the transcriptional regulator BzdR from *Azoarcus* sp. CIB. *The Journal of Biological Chemistry.* 285: 35694-35705.

Eisen MB, Spellman PT, Brown PO, Botstein D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA.* 95: 14863-14868.

Elhai J. (2015) Highly iterated palindromic sequences (HIPs) and their relationship to DNA methyltransferases. *Life* 5(1): 921-948.

Eraso JM, Kaplan S. (1994) prrA, a putative response regulator involved in oxygen regulation of photosynthesis gene expression in *Rhodobacter sphaeroides*. *Journal of Bacteriology.* 32-43.

Fujisawa T, Narikawa R, Okamoto S, Ehira S, Yoshimura H, Suzuki I, Masuda T, Mochimaru M, Takaichi S, Awai K, Sekine M, Horikawa H, Yashiro I, Omata S, Takarada H, Katano Y, Kosugi H, Tanikawa S, Ohmori K, Sato N, Ikeuchi M, Fujita N, and Ohmori M. (2010) Genomic structure of an economically important cyanobacterium, *Arthrospira* (Spirulina) *platensis* NIES-39. *DNA Research* 17:85-103

Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS (2007) Quantifying similarity between motifs. *Genome Biol* 8: R24.

Hershberg R, Yeger-Lotern E, Margalit H. (2005) Chromosomal organization is shaped by transcription regulatory network. *Trends Genet.* 21: 138-142.

Heyer LJ, Kruglyak S, Yooseph S. (1999) Exploring expression data: identification and analysis of coexpressed genes. *Genome Res.* 9: 1106-1115.

Hurst LD, Pál C, Lercher MJ. (2004) The evolutionary dynamics of eukaryotic gene order. *Nat Rev Genet.* 5: 299-310.

Janssen PJ, Morin N, Mergeay M, Leroy B, Wattiez R, Vallaeys T, Waleron K, Waleron M, Wilmotte A, Quillardet P, Tandeau de Marsac N, Talla E, Zhang C, and Leys N (2010) Genome sequence of the edible cyanobacterium *Arthrospira* sp. PCC 8005. *Journal of bacteriology* 192(9): 2465-2466.

Lefort F, Calmin G, Crovadore J, Falquet J, Hurni J, Osteras M, Haldemann F, and Farinelli L. (2014) Whole-genome shotgun sequence of *Arthrospira platensis* strain Paraca, a cultivated and edible cyanobacterium. *Journal ASM*

Novichkov PS, Rodionov DA, Stavrovskaya ED, Novichkova ES, Kazakov AE, Gelfand MS, Arkin AP, Mironov AA, and Dubchak I. (2010) RegPredict: an integrated system for regulon inference in prokaryotes by comparative genomics approach. *Nucleic Acids Research*. 38: W299-W307.

Panyakampol J, Cheevadhanarak S, Sutheworapong S, Chaijaruwanich J, Senachak J, Siangdung W, Jeamton W, Tanticharoen M, and Paithoonrangsarit K. (2014) Physiological and Transcriptional Responses to High Temperature in *Arthrospira* (Spirulina) *platensis* C1. *Plant Cell Physiol*. 56(3): 481-496.

Pilpel Y, Sudarsanam P. and Church G. (2001) Identifying regulatory networks by combinatorial analysis of promoter elements. *Nat. Genet*. 29: 153-159.

Rajaram H, Chaurasia AK, and Apte SK. (2014) Cyanobacterial heat-shock response: role and regulation of molecular chaperones. *Microbiology* 160(4): 647-658.

Ralston A (2008) Operons and prokaryotic gene regulation. *Nature Education* 1(1): 216.

Segal E, Yelensky R. and Koller D. (2003) Genome-wide discovery of transcriptional modules from dna sequence and gene expression. *Bioinformatics* 19: i273-i282.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* 13(11): 2498-504.

Smyth, G. K. (2004). Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, Vol. 3, No. 1, Article 3.

Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, et al. (1998) Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol Biol Cell*. 9: 3273-3297.

Spolaore P, Joannis-Cassan C, Duran E, and Isambert A (2006) Commercial applications of microalgae. *Journal of Bioscience and Bioengineering* 101:87-96.

Wang W, Cherry JM, Botstein D. and Li H. (2002) A systematic approach to reconstructing transcription networks in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci*. 99: 16893-98.

Wasserman WW, Krivan W. (2003) In silico identification of metazoan transcriptional regulatory regions. *Naturwissenschaften*. 90: 156-166.

Xing B. and van der Laan MJ. (2005) A statistical method for constructing transcriptional regulatory networks using gene expression and sequence data. *Journal of Computational Biology*. 12: 229-246.

Zorina AA, Mironov KS, Stepanchenko NS, Sinetova MA, Koroban NV, Zinchenko VV, Kupriyanova EV, Allahverdiev SI, and Los DA. (2011) Regulation Systems for Stress Responses in Cyanobacteria. Russian Journal of Plant Physiology. 58(5): 749-767.

เอกสารเพิ่มเติม 1

Locus_tag	Gene	COGs	kegg_cat	Description	T20	T60	T120	T240	T480
SPLC1_S010050		[S] Function unknown	#N/A	pentapeptide repeat protein	1.02	1.01	0.81	1.06	0.09
SPLC1_S010400	<i>rfbD</i>	[M] Cell wall/membrane/envelope biogenesis	Biosynthesis of Other Secondary Metabolites	dTDP-4-dehydrorhamnose reductase	-1.03	-1.42	-1.08	-0.94	-0.42
SPLC1_S010870	<i>dnaK</i>	[O] Posttranslational modification, protein turnover, chaperones	Folding, Sorting and Degradation	molecular chaperone DnaK	3.64	2.22	1.70	1.07	1.04
SPLC1_S010880		#N/A	#N/A	hypothetical protein	2.12	1.26	0.65	0.88	0.38
SPLC1_S010890		#N/A	#N/A	hypothetical protein	1.90	0.95	0.23	0.60	0.33
SPLC1_S011220		[S] Function unknown	#N/A	hypothetical protein	1.18	1.06	0.84	1.01	0.23
SPLC1_S011370	<i>nirA</i>	[P] Inorganic ion transport and metabolism	#N/A	nitrite and sulphite reductase 4Fe-4S region	-1.16	-0.84	-0.95	-1.01	-0.97
SPLC1_S011760		[V] Defense mechanisms	#N/A	HNH endonuclease	-0.81	-1.15	-1.19	-1.59	-1.21
SPLC1_S011870		#N/A	#N/A	hypothetical protein	1.07	1.13	0.80	0.96	0.41
SPLC1_S012010	<i>cynS</i>	[P] Inorganic ion transport and metabolism	Energy Metabolism	cyanate lyase	-1.16	-0.55	0.08	-0.33	-0.40
SPLC1_S012020		#N/A	#N/A	hypothetical protein	-1.13	-0.61	0.11	-0.10	-0.19
SPLC1_S012030		[S] Function unknown	#N/A	hypothetical protein	-1.62	-0.83	-0.04	-0.17	-0.24
SPLC1_S012040		[P] Inorganic ion transport and metabolism	#N/A	nitrate ABC transporter ATPase subunits C and D	-1.01	-0.22	0.15	-0.21	-0.21
SPLC1_S020030		[S] Function unknown	#N/A	zinc finger SWIM domain protein	0.74	1.12	0.40	0.31	0.28
SPLC1_S030420	<i>narB</i>	[C] Energy production and conversion	Energy Metabolism	molybdopterin oxidoreductase	-1.39	-1.82	-1.30	-0.58	-0.25
SPLC1_S030430	<i>nrtP</i>	[P] Inorganic ion transport and metabolism	#N/A	nitrite transporter	-1.89	-1.74	-1.52	-0.70	-0.28
SPLC1_S030440		#N/A	#N/A	hypothetical protein	-0.98	-1.16	-0.37	-0.14	-0.43
SPLC1_S030450		#N/A	#N/A	hypothetical protein	-1.31	-1.31	-0.73	0.03	-0.25
SPLC1_S030970		[O] Posttranslational modification, protein turnover, chaperones	#N/A	FeS assembly protein SufB	1.18	0.75	0.50	-0.28	-0.28
SPLC1_S031870		[S] Function unknown	#N/A	hypothetical protein	2.42	1.38	0.75	0.39	0.14
SPLC1_S031980		#N/A	#N/A	hypothetical protein	-1.49	-1.68	-1.54	-1.60	-1.22
SPLC1_S032770		#N/A	#N/A	hypothetical protein	0.87	1.45	1.39	0.97	0.94
SPLC1_S032780		#N/A	#N/A	hypothetical protein	1.14	1.05	0.80	0.63	0.29
SPLC1_S032910		#N/A	#N/A	hypothetical protein	1.70	0.59	0.44	0.16	0.21
SPLC1_S033220	<i>cpcE</i>	[C] Energy production and conversion	Energy Metabolism	PBS lyase HEAT domain protein repeat-containing protein	-1.20	-0.91	-0.80	-0.83	-0.06
SPLC1_S033470		[L] Replication, recombination and repair	#N/A	retron-type reverse transcriptase	-1.01	-1.30	-1.25	-1.74	-1.32
SPLC1_S033480		[L] Replication, recombination and repair	#N/A	group II intron maturase-specific domain protein	-1.13	-1.44	-1.48	-2.07	-1.58
SPLC1_S033490		[J] Translation, ribosomal structure and biogenesis	#N/A	sigma 54 modulation protein/ribosomal protein S30EA	-0.64	-1.22	-0.97	-1.47	-1.04
SPLC1_S033760		#N/A	#N/A	hypothetical protein	-0.66	-0.67	-0.81	-1.04	-0.67
SPLC1_S033770		#N/A	#N/A	hypothetical protein	-0.66	-0.67	-0.71	-1.03	-0.64
SPLC1_S033820		[S] Function unknown	#N/A	hypothetical protein	1.02	0.98	0.74	0.66	0.60
SPLC1_S033870		#N/A	#N/A	hypothetical protein	0.53	0.94	0.61	1.32	1.54

Locus_tag	Gene	COGs	kegg_cat	Description	T20	T60	T120	T240	T480
SPLC1_S010050		[S] Function unknown	#N/A	pentapeptide repeat protein	1.02	1.01	0.81	1.06	0.09
SPLC1_S040090		[L] Replication, recombination and repair	#N/A	group II intron maturase-specific domain protein	-1.03	-1.26	-0.62	-0.51	-0.32
SPLC1_S040110		[L] Replication, recombination and repair	#N/A	putative reverse transcriptase	-0.84	-1.05	-0.68	-0.37	-0.26
SPLC1_S040140		#N/A	#N/A	hypothetical protein	-0.30	-1.32	-1.14	-1.57	-1.66
SPLC1_S040590	<i>nbIB</i>	[C] Energy production and conversion	#N/A	phycocyanin alpha phycocyanobilin lyase related protein	-0.63	-0.84	-1.07	-1.44	-0.78
SPLC1_S041040		#N/A	#N/A	hypothetical protein	1.42	1.27	1.15	0.74	0.90
SPLC1_S051190	<i>glcA, GLS</i>	[E] Amino acid transport and metabolism	Amino Acid Metabolism	glutaminase	1.19	1.08	1.06	1.22	0.19
SPLC1_S060050		#N/A	#N/A	hypothetical protein	1.52	1.49	1.21	1.29	1.41
SPLC1_S060120		[J] Translation, ribosomal structure and biogenesis	#N/A	tRNA pseudouridine synthase A	-0.91	-1.08	-0.72	-0.37	-0.17
SPLC1_S060130	<i>RP-L17, rplQ</i>	[J] Translation, ribosomal structure and biogenesis	Translation	ribosomal protein L17	-0.83	-1.01	-0.55	-0.19	-0.05
SPLC1_S060180		[J] Translation, ribosomal structure and biogenesis	#N/A	translation initiation factor IF-1	-1.09	-1.04	-0.67	-0.48	0.17
SPLC1_S060350	<i>RP-L2, rplB</i>	[J] Translation, ribosomal structure and biogenesis	Translation	ribosomal protein L2	-1.12	-0.66	-0.40	0.04	0.03
SPLC1_S061460		#N/A	#N/A	hypothetical protein	1.02	0.21	0.12	-0.30	-0.06
SPLC1_S061580		#N/A	#N/A	hypothetical protein	0.91	1.12	0.98	1.04	0.41
SPLC1_S061630	<i>pill</i>	#N/A	Signal Transduction	CheW protein	1.03	0.98	0.61	0.35	0.23
SPLC1_S080370		[L] Replication, recombination and repair	#N/A	reverse transcriptase	-1.76	-2.09	-1.86	-2.19	-1.89
SPLC1_S080380	<i>secG</i>	[U] Intracellular trafficking, secretion, and vesicular transport	Folding, Sorting and Degradation	preprotein translocase SecG subunit	-0.77	-0.90	-1.02	-1.23	-0.89
SPLC1_S080420		#N/A	#N/A	hypothetical protein	-1.68	-1.89	-1.10	-0.95	0.05
SPLC1_S080430		#N/A	#N/A	hypothetical protein	-1.86	-2.00	-1.15	-1.00	-0.29
SPLC1_S080560		[J] Translation, ribosomal structure and biogenesis	#N/A	translation initiation factor IF-2	-0.56	-0.36	-0.64	-1.05	-0.46
SPLC1_S080680		#N/A	#N/A	30S ribosomal protein S6	-0.97	-1.13	-0.57	-0.61	-0.32
SPLC1_S080740		[O] Posttranslational modification, protein turnover, chaperones	#N/A	FAD-dependent pyridine nucleotide-disulfide oxidoreductase	1.48	0.98	0.49	0.13	0.26
SPLC1_S080750	<i>glpA, glpD</i>	[R] General function prediction only	Lipid Metabolism	FAD dependent oxidoreductase	1.26	0.81	0.59	0.13	0.34
SPLC1_S080980		[O] Posttranslational modification, protein turnover, chaperones	#N/A	thioredoxin domain protein	1.07	1.14	0.91	0.10	0.14
SPLC1_S081950	<i>alr</i>	[M] Cell wall/membrane/envelope biogenesis	Metabolism	alanine racemase	1.21	1.06	0.97	0.93	0.12
SPLC1_S082630		#N/A	#N/A	hypothetical protein	1.11	0.95	0.85	1.16	0.19
SPLC1_S100100		#N/A	#N/A	hypothetical protein	3.50	2.87	1.60	1.01	0.79
SPLC1_S100290	<i>urtD</i>	[R] General function prediction only	Membrane Transport	urea ABC transporter ATP-binding protein UrtD	-1.11	-0.61	-0.11	-0.37	-0.33

Locus_tag	Gene	COGs	kegg_cat	Description	T20	T60	T120	T240	T480
SPLC1_S010050		[S] Function unknown	#N/A	pentapeptide repeat protein	1.02	1.01	0.81	1.06	0.09
SPLC1_S100410		#N/A	#N/A	hypothetical protein	0.28	0.30	0.36	2.88	2.76
SPLC1_S100420		[D] Cell cycle control, cell division, chromosome partitioning	#N/A	hypothetical protein	0.78	0.73	0.95	3.81	4.01
SPLC1_S100430		[K] Transcription	#N/A	hypothetical protein	0.31	0.44	0.33	2.65	2.87
SPLC1_S100440		#N/A	#N/A	hypothetical protein	0.18	0.21	0.19	1.41	1.48
SPLC1_S100450		#N/A	#N/A	hypothetical protein	0.89	1.01	1.37	4.68	5.18
SPLC1_S100460		#N/A	#N/A	hypothetical protein	0.59	0.78	0.99	4.05	4.70
SPLC1_S100470		#N/A	#N/A	hypothetical protein	0.39	0.52	0.95	3.79	4.12
SPLC1_S100480		[S] Function unknown	#N/A	transcriptional regulator XRE family	-0.65	-0.35	-0.04	0.76	1.84
SPLC1_S100690		#N/A	#N/A	hypothetical protein	0.93	1.46	0.83	0.85	0.57
SPLC1_S100710		#N/A	#N/A	hypothetical protein	0.64	1.40	1.18	1.22	1.61
SPLC1_S100730		#N/A	#N/A	hypothetical protein	0.05	0.38	0.36	0.64	1.01
SPLC1_S100850		#N/A	#N/A	hypothetical protein	1.02	1.05	0.77	0.43	0.26
SPLC1_S101110		[T] Signal transduction mechanisms	#N/A	putative two component sensor histidine kinase	-1.47	-1.53	-0.81	-0.78	-0.34
SPLC1_S101140		#N/A	#N/A	hypothetical protein	0.31	0.41	0.66	1.05	0.46
SPLC1_S101290		#N/A	#N/A	hypothetical protein	1.15	1.06	0.76	0.90	0.61
SPLC1_S101570		#N/A	#N/A	2-alkenal reductase	-1.26	-1.30	-1.05	-1.16	-0.20
SPLC1_S102190		#N/A	#N/A	hypothetical protein	-1.04	-0.96	-0.52	-0.76	-0.50
SPLC1_S102540		#N/A	#N/A	hypothetical protein	1.51	1.20	0.74	1.29	0.73
SPLC1_S102560		[R] General function prediction only	#N/A	hypothetical protein	1.07	0.96	0.49	0.64	0.28
SPLC1_S102800		#N/A	#N/A	hypothetical protein	-0.42	-0.71	-0.74	-1.27	-0.67
SPLC1_S102990		[S] Function unknown	#N/A	ATP-dependent Clp protease adaptor protein ClpS	1.07	0.93	0.56	0.81	0.49
SPLC1_S103050		[O] Posttranslational modification, protein turnover, chaperones	#N/A	AAA ATPase central region	1.07	0.98	0.85	0.89	0.11
SPLC1_S103160		#N/A	#N/A	protein of unknown function DUF1825	1.07	0.25	0.24	-0.24	-0.11
SPLC1_S103250	<i>RP-L25, rplY</i>	[J] Translation, ribosomal structure and biogenesis	Translation	ribosomal 5S rRNA E-loop binding protein Ctc/L25/TL5	-0.62	-0.69	-0.65	-1.18	-0.49
SPLC1_S103260	<i>purA</i>	[F] Nucleotide transport and metabolism	Amino Acid Metabolism	adenylosuccinate synthase	-1.06	-0.93	-0.78	-0.90	-0.55
SPLC1_S130330		#N/A	#N/A	hypothetical protein	-0.19	-0.43	-0.60	-1.03	-0.76
SPLC1_S130390		[R] General function prediction only	#N/A	tail sheath protein	-0.38	-0.82	-0.77	-1.23	-1.15
SPLC1_S130480		[T] Signal transduction mechanisms	#N/A	hypothetical protein	-1.20	-0.94	-0.57	-0.90	-0.47
SPLC1_S130960		#N/A	#N/A	hypothetical protein	0.98	1.05	0.75	0.63	0.32
SPLC1_S160040		[S] Function unknown	#N/A	hypothetical protein	1.00	0.93	0.83	0.78	0.05
SPLC1_S170160		#N/A	#N/A	hypothetical protein	0.03	0.33	0.54	0.36	1.13
SPLC1_S170170		#N/A	#N/A	hypothetical protein	0.03	0.37	0.57	0.39	1.02
SPLC1_S170410	<i>cpn10</i>	[O] Posttranslational modification, protein turnover, chaperones	#N/A	chaperonin Cpn10	4.27	4.32	3.95	3.33	2.28

Locus_tag	Gene	COGs	kegg_cat	Description	T20	T60	T120	T240	T480
SPLC1_S010050		[S] Function unknown	#N/A	pentapeptide repeat protein	1.02	1.01	0.81	1.06	0.09
SPLC1_S170420	<i>groEL1</i>	[O] Posttranslational modification, protein turnover, chaperones	Folding, Sorting and Degradation	chaperonin GroEL	3.77	4.64	3.62	3.03	1.80
SPLC1_S170920		#N/A	#N/A	putative short-chain dehydrogenase/reductase SDR	1.23	0.41	0.21	0.13	0.20
SPLC1_S170940		#N/A	#N/A	hypothetical protein	1.04	0.21	0.30	0.17	-0.01
SPLC1_S170950		#N/A	#N/A	ion transport protein	1.35	1.10	0.99	0.89	0.21
SPLC1_S170960		[O] Posttranslational modification, protein turnover, chaperones	#N/A	peptidase S1 and S6 chymotrypsin/Hap	3.30	0.99	1.02	0.21	0.69
SPLC1_S170970		#N/A	#N/A	hypothetical protein	2.51	0.60	0.27	0.29	0.42
SPLC1_S170980	<i>dnaJ</i>	[O] Posttranslational modification, protein turnover, chaperones	#N/A	chaperone DnaJ domain protein	2.77	0.55	0.26	0.03	0.08
SPLC1_S171000	<i>hspA, HSP20</i>	[O] Posttranslational modification, protein turnover, chaperones	Folding, Sorting and Degradation	small heat shock protein molecular chaperone	6.07	4.55	1.22	1.97	1.41
SPLC1_S171060		#N/A	#N/A	hypothetical protein	1.10	0.85	0.83	0.97	0.11
SPLC1_S180030	<i>lspA</i>	#N/A	Folding, Sorting and Degradation	signal peptidase II	1.29	1.16	1.02	1.08	0.12
SPLC1_S180040		[R] General function prediction only	#N/A	BioY protein	-0.41	-0.28	-0.68	-1.18	-0.70
SPLC1_S190200		[S] Function unknown	#N/A	hypothetical protein	0.92	0.78	0.70	1.03	1.04
SPLC1_S200530		#N/A	#N/A	hypothetical protein	-0.69	-0.26	-0.53	-0.89	-1.11
SPLC1_S200690	<i>RP-L7, rplL</i>	[J] Translation, ribosomal structure and biogenesis	Translation	ribosomal protein L7/L12	-0.44	-1.63	-1.42	-1.92	-1.70
SPLC1_S200700	<i>RP-L10, rplJ</i>	[J] Translation, ribosomal structure and biogenesis	Translation	50S ribosomal protein L10	-0.56	-1.37	-0.99	-1.24	-1.27
SPLC1_S201370	<i>malQ</i>	[G] Carbohydrate transport and metabolism	Carbohydrate Metabolism	4-alpha-glucanotransferase	-1.16	-1.02	-0.83	-1.00	-0.02
SPLC1_S202720		[S] Function unknown	#N/A	hypothetical protein	0.88	0.91	1.01	0.94	1.50
SPLC1_S202930		#N/A	#N/A	hypothetical protein	1.01	1.01	0.48	0.06	0.49
SPLC1_S203440		#N/A	#N/A	hypothetical protein	1.27	0.65	0.74	0.51	0.49
SPLC1_S203590		#N/A	#N/A	hypothetical protein	1.54	1.73	1.33	0.35	0.09
SPLC1_S203860		[O] Posttranslational modification, protein turnover, chaperones	#N/A	peptidase S1 and S6 chymotrypsin/Hap	1.00	0.95	0.65	0.40	0.51
SPLC1_S203930		#N/A	#N/A	putative peptidase C14 caspase catalytic subunit p20	1.14	1.07	0.95	1.01	0.04
SPLC1_S204920		[S] Function unknown	#N/A	hypothetical protein	-1.22	-1.41	-0.82	-0.71	-0.74
SPLC1_S204940	<i>ftsW, spoVE</i>	[D] Cell cycle control, cell division, chromosome partitioning	Cell Growth and Death	cell cycle protein	-1.05	-1.20	-0.69	-0.79	-0.63
SPLC1_S205570	<i>ATPF0B, atpF</i>	[C] Energy production and conversion	Energy Metabolism	ATP synthase F0 B subunit	-0.79	-0.90	-0.99	-1.20	-0.95
SPLC1_S205630		[I] Lipid transport and metabolism	#N/A	helix-hairpin-helix motif protein	1.27	1.14	0.98	0.89	0.20

Locus_tag	Gene	COGs	kegg_cat	Description	T20	T60	T120	T240	T480
SPLC1_S010050		[S] Function unknown	#N/A	pentapeptide repeat protein	1.02	1.01	0.81	1.06	0.09
SPLC1_S205650	<i>iscS, NFS1</i>	[E] Amino acid transport and metabolism	Folding, Sorting and Degradation	aminotransferase class V	-0.28	-0.33	-0.49	-1.12	-0.71
SPLC1_S205990		#N/A	#N/A	type I restriction modification	1.11	1.08	0.78	0.44	0.37
SPLC1_S206560	<i>pps, ppsA</i>	[G] Carbohydrate transport and metabolism	Carbohydrate Metabolism	phosphoenolpyruvate synthase	1.88	1.65	1.20	0.72	0.40
SPLC1_S206830	<i>chlB</i>	[C] Energy production and conversion	Metabolism	light-independent protochlorophyllide reductase B subunit	0.19	-0.41	-0.60	-1.30	-0.61
SPLC1_S206860		[L] Replication, recombination and repair	#N/A	hypothetical protein	1.03	0.64	0.27	-0.66	-0.67
SPLC1_S206920		#N/A	#N/A	KaiA family protein	0.61	0.68	0.76	1.13	0.75
SPLC1_S206940	<i>kaiC</i>	[T] Signal transduction mechanisms	#N/A	putative circadian clock protein KaiC	-0.95	-0.82	-1.06	-1.17	-0.71
SPLC1_S207080	<i>RP-L21, rplU</i>	[J] Translation, ribosomal structure and biogenesis	Translation	hypothetical protein	-1.14	-0.78	-0.51	-0.41	-0.30
SPLC1_S207110		#N/A	#N/A	amino acid permease-associated region	1.13	1.44	1.11	0.61	0.81
SPLC1_S207140		#N/A	#N/A	hypothetical protein	-0.78	-1.09	-0.71	-0.58	-0.16
SPLC1_S207280		#N/A	#N/A	hypothetical protein	-0.61	-1.58	-0.79	-0.64	-0.62
SPLC1_S207680		#N/A	#N/A	hypothetical protein	0.82	1.19	0.98	0.91	0.48
SPLC1_S207690		#N/A	#N/A	hypothetical protein	0.13	0.20	0.42	1.19	0.19
SPLC1_S208490		[S] Function unknown	#N/A	cobalamin biosynthesis CbiX protein	1.20	1.30	0.95	0.98	1.07
SPLC1_S208530		#N/A	#N/A	hypothetical protein	0.83	1.28	0.67	0.59	0.72
SPLC1_S208890		#N/A	#N/A	hypothetical protein	1.09	0.68	0.42	0.13	0.37
SPLC1_S210020		#N/A	#N/A	zinc finger SWIM domain protein	0.86	1.10	0.82	0.87	1.30
SPLC1_S220190	<i>speD</i>	[E] Amino acid transport and metabolism	#N/A	hypothetical protein	0.62	1.11	0.79	0.72	0.39
SPLC1_S230640	<i>atoB</i>	[I] Lipid transport and metabolism	Amino Acid Metabolism	putative acetyl coenzyme A acetyltransferase	-1.01	-1.41	-1.70	-2.21	-2.09
SPLC1_S230980	<i>TST, MPST, sseA</i>	[P] Inorganic ion transport and metabolism	Amino Acid Metabolism	rhodanese domain protein	1.02	0.95	0.76	0.59	0.27
SPLC1_S240060	<i>rbcL</i>	[G] Carbohydrate transport and metabolism	Carbohydrate Metabolism	ribulose-bisphosphate carboxylase	-0.45	-0.71	-0.60	-1.13	-0.66
SPLC1_S240070		#N/A	#N/A	chaperonin family protein RbcX	-0.75	-1.11	-0.87	-1.40	-0.94
SPLC1_S240080	<i>rbcS</i>	[C] Energy production and conversion	#N/A	ribulose-bisphosphate carboxylase	-0.57	-0.95	-0.71	-1.26	-0.73
SPLC1_S260330		[S] Function unknown	#N/A	SCP-like extracellular	1.36	1.05	0.56	0.35	0.19
SPLC1_S260350		#N/A	#N/A	hypothetical protein	1.02	0.89	0.82	0.75	0.16
SPLC1_S260450		#N/A	#N/A	hypothetical protein	0.41	0.52	0.44	0.91	1.04
SPLC1_S260460		#N/A	#N/A	hypothetical protein	0.83	0.78	0.73	0.96	1.13
SPLC1_S260670		[S] Function unknown	#N/A	hypothetical protein	1.83	0.80	0.43	0.23	0.09
SPLC1_S260970		[R] General function prediction only	#N/A	hypothetical protein	-1.19	-1.23	-0.84	-0.62	-0.32
SPLC1_S270010		#N/A	#N/A	hypothetical protein	0.84	1.10	0.63	0.51	0.22
SPLC1_S270730		[I] Lipid transport and metabolism	#N/A	phospholipid/glycerol acyltransferase	1.08	1.06	0.78	0.76	0.68
SPLC1_S271220		[G] Carbohydrate transport and metabolism	#N/A	peptidase M42 family protein	1.10	0.94	0.81	1.05	0.14
SPLC1_S271230		#N/A	#N/A	hypothetical protein	0.45	0.45	0.71	1.07	0.40

Locus_tag	Gene	COGs	kegg_cat	Description	T20	T60	T120	T240	T480
SPLC1_S010050		[S] Function unknown	#N/A	pentapeptide repeat protein	1.02	1.01	0.81	1.06	0.09
SPLC1_S271340		[S] Function unknown	#N/A	Rho termination factor domain protein	1.01	1.22	0.87	0.31	0.34
SPLC1_S271390		[G] Carbohydrate transport and metabolism	Energy Metabolism	phosphoketolase	-0.80	-1.17	-0.78	-0.86	-0.53
SPLC1_S300180		#N/A	#N/A	type III restriction protein res subunit	1.06	0.89	0.80	0.94	0.15
SPLC1_S300310		[O] Posttranslational modification, protein turnover, chaperones	#N/A	2OG-Fe(II) oxygenase	0.94	1.01	0.62	0.53	0.63
SPLC1_S300490		#N/A	#N/A	hypothetical protein	-0.44	-0.06	0.09	0.70	1.13
SPLC1_S300580		[O] Posttranslational modification, protein turnover, chaperones	#N/A	AAA ATPase central domain protein	0.13	0.40	0.55	0.35	1.06
SPLC1_S300600		#N/A	#N/A	hypothetical protein	-0.51	0.29	0.44	0.23	1.24
SPLC1_S300740		[C] Energy production and conversion	#N/A	hypothetical protein	1.20	0.88	0.84	0.90	0.38
SPLC1_S300780		[C] Energy production and conversion	#N/A	hybrid cluster protein	1.34	1.07	0.61	0.35	0.50
SPLC1_S300940		[V] Defense mechanisms	#N/A	reverse transcriptase	-0.90	-1.22	-0.64	-0.46	-0.45
SPLC1_S310010	<i>clpB</i>	[O] Posttranslational modification, protein turnover, chaperones	#N/A	ATP-dependent chaperone ClpB	2.41	1.20	0.36	0.38	0.43
SPLC1_S310020		#N/A	#N/A	hypothetical protein	2.59	1.21	0.38	0.34	0.24
SPLC1_S310230		[O] Posttranslational modification, protein turnover, chaperones	#N/A	2-alkenal reductase	0.38	1.12	0.72	0.81	1.19
SPLC1_S310330		[S] Function unknown	#N/A	SCP-like extracellular	1.23	0.81	0.23	0.04	0.23
SPLC1_S310420	<i>glgA</i>	[G] Carbohydrate transport and metabolism	Carbohydrate Metabolism	glycogen/starch synthase ADP-glucose type	-0.64	-0.96	-0.81	-1.04	-0.80
SPLC1_S350020		#N/A	#N/A	hypothetical protein	1.02	0.86	0.67	0.47	0.38
SPLC1_S360230		#N/A	#N/A	hypothetical protein	1.08	1.09	0.59	0.06	-0.25
SPLC1_S360330		[R] General function prediction only	#N/A	PHP domain protein	0.89	1.04	0.68	0.76	0.69
SPLC1_S360590		#N/A	#N/A	sulfate ABC transporter periplasmic sulfate-binding protein	2.90	1.28	1.13	0.39	-0.40
SPLC1_S360600	<i>mgtE</i>	[P] Inorganic ion transport and metabolism	#N/A	putative magnesium transporter	2.83	1.44	1.26	0.09	-0.27
SPLC1_S360610		#N/A	#N/A	hypothetical protein	2.48	1.26	1.02	0.64	0.16
SPLC1_S360620	<i>mgtC</i>	[S] Function unknown	#N/A	MgtC/SapB transporter	2.25	0.82	0.72	0.44	0.04
SPLC1_S360940		#N/A	#N/A	hypothetical protein	1.48	0.94	0.62	0.41	0.30
SPLC1_S361030		[L] Replication, recombination and repair	#N/A	putative RNA methylase	-1.04	-0.85	-0.56	-0.29	0.18
SPLC1_S370210		#N/A	#N/A	hypothetical protein	1.06	0.92	0.63	1.06	0.56
SPLC1_S370770		#N/A	#N/A	hypothetical protein	-0.75	-1.09	-0.61	-0.52	-0.40
SPLC1_S370960		#N/A	#N/A	hypothetical protein	-0.29	-1.14	-0.44	-0.54	-0.43
SPLC1_S380100		[S] Function unknown	#N/A	hypothetical protein	1.63	1.57	1.50	1.21	1.43
SPLC1_S410610		[S] Function unknown	#N/A	pentapeptide repeat protein	1.11	0.99	0.84	0.84	0.28
SPLC1_S410620		#N/A	#N/A	hypothetical protein	2.36	2.21	1.91	1.72	1.46

Locus_tag	Gene	COGs	kegg_cat	Description	T20	T60	T120	T240	T480
SPLC1_S010050		[S] Function unknown	#N/A	pentapeptide repeat protein	1.02	1.01	0.81	1.06	0.09
SPLC1_S410830	<i>tatA</i>	[U] Intracellular trafficking, secretion, and vesicular transport	Folding, Sorting and Degradation	twin-arginine translocation protein TatA/E family subunit	-0.83	-1.01	-0.77	-0.69	-0.28
SPLC1_S411440		[R] General function prediction only	#N/A	hypothetical protein	-0.33	-1.03	-0.74	-1.02	-0.77
SPLC1_S411530		[L] Replication, recombination and repair	#N/A	CRISPR-associated protein Cas2	1.16	1.02	0.68	1.14	0.69
SPLC1_S412180		[P] Inorganic ion transport and metabolism	#N/A	ABC-type nitrate/nitrite transport system substrate-binding protein	-1.35	-1.20	-0.95	-0.65	-0.09
SPLC1_S412190	<i>nrtB</i>	[P] Inorganic ion transport and metabolism	#N/A	nitrate ABC transporter inner membrane subunit	-1.53	-1.67	-1.24	-1.06	-0.74
SPLC1_S412200	<i>nrtC</i>	[P] Inorganic ion transport and metabolism	Membrane Transport	nitrate ABC transporter ATPase subunits C and D	-0.85	-1.06	-0.62	-0.15	-0.25
SPLC1_S412210	<i>nrtD</i>	[P] Inorganic ion transport and metabolism	Membrane Transport	nitrate ABC transporter ATPase subunits C and D	-1.50	-1.69	-1.25	-0.75	-0.34
SPLC1_S412220		#N/A	#N/A	hypothetical protein	-1.14	-1.34	-1.01	-0.41	-0.06
SPLC1_S412240		#N/A	#N/A	hypothetical protein	-1.26	-1.43	-0.80	-0.46	-0.46
SPLC1_S412250		#N/A	#N/A	hypothetical protein	-1.49	-1.74	-0.92	-0.59	-0.46
SPLC1_S412260		[R] General function prediction only	#N/A	hypothetical protein	-1.87	-2.02	-1.49	-0.80	-0.21
SPLC1_S420050		#N/A	#N/A	hypothetical protein	-0.68	-1.15	-1.42	-2.00	-1.85
SPLC1_S420060		#N/A	#N/A	hypothetical protein	-0.81	-1.23	-1.03	-1.35	-1.19
SPLC1_S420070		[O] Posttranslational modification, protein turnover, chaperones	#N/A	subtilisin-like protein	-0.74	-1.24	-1.28	-1.44	-1.37
SPLC1_S420330		[R] General function prediction only	#N/A	hypothetical protein	1.19	1.63	0.92	0.36	0.03
SPLC1_S430100	<i>UGDH, ugd</i>	[M] Cell wall/membrane/envelope biogenesis	Carbohydrate Metabolism	UDP-glucose/GDP-mannose dehydrogenase	-0.66	-0.96	-0.86	-1.06	-0.62
SPLC1_S430160	<i>RARS, argS</i>	[J] Translation, ribosomal structure and biogenesis	Translation	arginyl-tRNA synthetase	-1.16	-1.16	-1.29	-1.59	-1.08
SPLC1_S430180		[R] General function prediction only	#N/A	hypothetical protein	1.14	1.37	0.86	0.45	0.41
SPLC1_S430350		#N/A	#N/A	hypothetical protein	0.97	1.24	0.75	0.55	0.39
SPLC1_S480030		#N/A	#N/A	hypothetical protein	0.93	1.07	0.52	0.35	0.20
SPLC1_S500430		#N/A	#N/A	hypothetical protein	1.26	1.22	0.90	1.04	0.41
SPLC1_S500630		[S] Function unknown	#N/A	hypothetical protein	1.47	0.17	0.09	-0.21	-0.20
SPLC1_S500800		[J] Translation, ribosomal structure and biogenesis	#N/A	putative elongation factor P	-0.58	-0.32	-0.60	-1.05	-0.63
SPLC1_S501000	<i>DPO3D2, hoIb</i>	[L] Replication, recombination and repair	Metabolism	DNA polymerase III delta prime subunit	1.12	0.97	0.93	0.78	0.14
SPLC1_S510080		[H] Coenzyme transport and metabolism	#N/A	C-methyltransferase	-0.06	-0.26	-0.51	-1.02	-0.57
SPLC1_S510240		[J] Translation, ribosomal structure and biogenesis	#N/A	translation elongation factor Ts	-0.64	-0.81	-0.66	-1.08	-0.31

Locus_tag	Gene	COGs	kegg_cat	Description	T20	T60	T120	T240	T480
SPLC1_S010050		[S] Function unknown	#N/A	pentapeptide repeat protein	1.02	1.01	0.81	1.06	0.09
SPLC1_S510250	<i>RP-S2, rpsB</i>	[J] Translation, ribosomal structure and biogenesis	Translation	ribosomal protein S2	-0.29	-0.46	-0.82	-1.41	-0.96
SPLC1_S510770		#N/A	#N/A	phage shock protein C PspC	0.90	-1.27	-0.47	-0.64	-0.93
SPLC1_S511030		#N/A	#N/A	hypothetical protein	0.93	1.00	0.98	0.85	0.81
SPLC1_S511080		#N/A	#N/A	hypothetical protein	-1.01	-1.04	-0.61	-0.51	-0.31
SPLC1_S520310		#N/A	#N/A	hypothetical protein	0.89	0.80	0.63	1.03	0.17
SPLC1_S520320		[C] Energy production and conversion	#N/A	geranylgeranyl reductase	-0.59	-0.55	-0.90	-1.32	-1.24
SPLC1_S521030		#N/A	#N/A	hypothetical protein	1.48	0.51	0.16	0.24	0.05
SPLC1_S521200		#N/A	#N/A	hypothetical protein	0.11	-0.63	-0.73	-1.21	-0.96
SPLC1_S530170		[S] Function unknown	#N/A	hypothetical protein	1.15	1.02	0.74	0.43	0.64
SPLC1_S530280		#N/A	#N/A	hypothetical protein	1.14	0.95	0.65	0.28	0.25
SPLC1_S531220		#N/A	#N/A	hypothetical protein	1.37	1.52	1.16	1.17	1.37
SPLC1_S531610		[H] Coenzyme transport and metabolism	#N/A	methyltransferase type 11	-0.34	-0.31	-0.05	0.68	1.18
SPLC1_S531750		#N/A	#N/A	hemolysin-type calcium-binding region	-0.48	-0.42	-0.55	-1.06	-0.75
SPLC1_S531760		#N/A	#N/A	hypothetical protein	-0.04	-0.03	-0.42	-1.00	-0.57
SPLC1_S532550		[S] Function unknown	#N/A	MORN repeat-containing protein	1.16	1.07	0.94	0.57	0.91
SPLC1_S532860		[S] Function unknown	#N/A	serine/threonine protein kinase	0.89	1.13	0.77	0.52	0.48
SPLC1_S532900		#N/A	#N/A	hypothetical protein	1.18	1.13	0.68	0.30	0.17
SPLC1_S532910		[S] Function unknown	#N/A	serine/threonine protein kinase	1.31	1.39	0.74	0.42	0.24
SPLC1_S533550		#N/A	#N/A	hypothetical protein	-0.98	-1.06	-0.67	-0.66	-0.77
SPLC1_S540630		[P] Inorganic ion transport and metabolism	#N/A	hypothetical protein	-0.95	-1.09	-0.63	-0.74	-0.54
SPLC1_S540640		[P] Inorganic ion transport and metabolism	#N/A	hypothetical protein	-1.02	-1.10	-0.87	-0.73	-0.52
SPLC1_S540700		[G] Carbohydrate transport and metabolism	#N/A	hypothetical protein	-1.52	-1.62	-1.16	-1.40	-1.01
SPLC1_S540710		[G] Carbohydrate transport and metabolism	#N/A	glucosylglycerol-phosphate synthase	-0.96	-1.14	-0.83	-1.12	-0.74
SPLC1_S540720	<i>glpA, glpD</i>	[C] Energy production and conversion	Lipid Metabolism	FAD dependent oxidoreductase	-1.66	-1.50	-1.67	-1.95	-1.36
SPLC1_S540730	<i>gpsA</i>	[C] Energy production and conversion	#N/A	NAD-dependent glycerol-3-phosphate dehydrogenase domain protein	-1.37	-1.59	-0.99	-0.63	-0.30
SPLC1_S541050		#N/A	#N/A	hypothetical protein	-0.63	-1.64	-1.17	-1.20	-1.01
SPLC1_S541060		[V] Defense mechanisms	#N/A	ABC transporter related protein	-0.65	-1.45	-0.81	-0.85	-0.61
SPLC1_S541170		#N/A	#N/A	hypothetical protein	-1.01	-1.03	-0.73	-0.85	-0.34
SPLC1_S541200	<i>rmA, ADH5, adhE</i>	[C] Energy production and conversion	Amino Acid Metabolism	S-(hydroxymethyl)glutathione dehydrogenase/class III alcoholdehydrogenase	1.27	1.04	0.96	0.98	0.18
SPLC1_S541290	<i>folE</i>	[H] Coenzyme transport and metabolism	Metabolism	putative GTP cyclohydrolase I	-1.35	-1.32	-0.77	-0.53	-0.27
SPLC1_S541640		#N/A	#N/A	hypothetical protein	0.26	0.40	0.37	1.14	0.74
SPLC1_S541660		#N/A	#N/A	hypothetical protein	0.11	0.27	0.44	1.40	1.22
SPLC1_S541980		#N/A	#N/A	sulfotransferase	0.26	0.51	0.44	0.63	1.02

Locus_tag	Gene	COGs	kegg_cat	Description	T20	T60	T120	T240	T480
SPLC1_S010050		[S] Function unknown	#N/A	pentapeptide repeat protein	1.02	1.01	0.81	1.06	0.09
SPLC1_S542180	<i>groEL2</i>	[O] Posttranslational modification, protein turnover, chaperones	Folding, Sorting and Degradation	chaperonin GroEL	2.37	2.65	1.45	1.12	0.56
SPLC1_S542730	<i>HARS, hisS</i>	[J] Translation, ribosomal structure and biogenesis	Translation	histidyl-tRNA synthetase	-0.88	-1.06	-0.57	-0.71	-0.54
SPLC1_S543170		[V] Defense mechanisms	#N/A	hypothetical protein	-1.18	-1.38	-1.12	-1.28	-1.11
SPLC1_S550010		[V] Defense mechanisms	#N/A	group II intron maturase-specific domain protein	-0.47	-1.45	-1.43	-1.69	-1.16
SPLC1_S550520		#N/A	#N/A	hypothetical protein	-0.90	-1.01	-0.58	-0.64	-0.65
SPLC1_S560040	<i>trpA</i>	[E] Amino acid transport and metabolism	Amino Acid Metabolism	tryptophan synthase alpha subunit	-0.87	-1.06	-0.88	-0.98	-0.79
SPLC1_S560130		#N/A	#N/A	hypothetical protein	-0.68	-1.44	-0.72	-0.74	-1.01
SPLC1_S560190	<i>htpG</i>	[O] Posttranslational modification, protein turnover, chaperones	Cancers	ATP-binding region ATPase domain protein	3.73	2.69	2.01	1.84	1.08
SPLC1_S590700		#N/A	#N/A	hypothetical protein	0.57	0.89	0.68	0.68	1.28
SPLC1_S630200	<i>hik34</i>	[T] Signal transduction mechanisms	#N/A	histidine kinase	3.08	1.28	0.90	0.61	0.04