



รายงานการวิจัยฉบับสมบูรณ์

ชีววิทยาเชิงระบบสำหรับสืบค้นกลไกการตอบสนองต่อการอักเสบในเซลล์แมคโครฟาจของหนู

Systems Biology for Investigating the Mechanism of Inflammatory Response
in Murine Macrophage

โดย

ผู้ช่วยศาสตราจารย์ ดร. พิทักษ์ สุตรอนันต์

รองศาสตราจารย์ ดร. กล่าวขวัญ ศรีสุข

ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

ได้รับเงินอุดหนุนการวิจัยงบประมาณเงินรายได้ (เงินอุดหนุนจากรัฐบาล)

งบประมาณปี 2557



รายงานการวิจัยฉบับสมบูรณ์

ชีววิทยาเชิงระบบสำหรับสืบค้นกลไกการตอบสนองต่อการอักเสบในเซลล์แมคโครฟาจของหนู

Systems Biology for Investigating the Mechanism of Inflammatory Response
in Murine Macrophage

โดย

ผู้ช่วยศาสตราจารย์ ดร. พิทักษ์ สุตรอนันต์

รองศาสตราจารย์ ดร. กล่าวขวัญ ศรีสุข

ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

ได้รับเงินอุดหนุนการวิจัยงบประมาณเงินรายได้ (เงินอุดหนุนจากรัฐบาล)

งบประมาณปี 2557

หัวข้อโครงการวิจัย ชีววิทยาเชิงระบบสำหรับสืบค้นกลไกการตอบสนองต่อการอักเสบในเซลล์แมคโคร
ฟาจของหนู

หัวหน้าโครงการวิจัย ผู้ช่วยศาสตราจารย์ ดร. พิทักษ์ สุตรอนันต์

ผู้ร่วมโครงการวิจัย รองศาสตราจารย์ ดร. กล่าวขวัญ ศรีสุข

หน่วยงาน ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

โครงการวิจัยนี้ได้รับเงินอุดหนุนการวิจัยงบประมาณเงินรายได้ (เงินอุดหนุนจากรัฐบาล) รหัส 87/2557

บทคัดย่อ

เซลล์แมคโครฟาจ RAW264.7 ที่ถูกกระตุ้นด้วย LPS เป็นโมเดลที่ถูกใช้ในการศึกษากลไกการอักเสบ ซึ่งมีหลายกลุ่มวิจัยได้นำมาใช้เป็นโมเดลในการศึกษาสารออกฤทธิ์ต้านการอักเสบ โดยอาศัยเทคโนโลยีไมโครอาร์เรย์ที่สามารถนำมาใช้ในการตรวจวัดการตอบสนองของยีนทั้งหมดภายในเซลล์ในช่วงเวลาเดียวกัน ซึ่งข้อมูลที่ได้จากการตรวจวัดระดับของ mRNA ภายในข้อมูลไมโครอาร์เรย์ที่มีการศึกษากันเป็นจำนวนมากนั้นถูกรวบรวมไว้ในฐานข้อมูลสาธารณะ ข้อมูลไมโครอาร์เรย์เหล่านี้สามารถนำมาใช้ในการศึกษาด้วยเครื่องมือทางชีวสารสนเทศหรือการศึกษาทางด้านชีววิทยาเชิงระบบ เพื่อค้นหากลไกการตอบสนองการอักเสบของเซลล์แมคโครฟาจอย่างเป็นระบบได้ เครื่องมือที่สำคัญที่นำมาใช้ในการวิจัยครั้งนี้ได้แก่ เครื่องมือทางชีวสารสนเทศที่เกี่ยวข้องกับการคัดเลือกยีน การสร้างเครือข่ายปฏิสัมพันธ์ของการแสดงออกร่วม การสร้างเครือข่ายการควบคุมการแสดงออกของยีน การบูรณาการข้อมูลทางชีวภาพ รวมถึงการแปลผลทางชีวภาพ เพื่อทำการค้นหาเครือข่ายย่อยและชุดยีนที่น่าสนใจซึ่งมีการตอบสนองต่อการอักเสบ จากผลการคัดเลือกทำให้ได้ยีน *SOCS3* ที่ผ่านยืนยันผลในระดับห้องปฏิบัติการโดยการตรวจวัดระดับการแสดงออกด้วยเทคนิค real-time PCR ซึ่งยีน *SOCS3* นี้เป็นยีนที่จะแสดงออกเป็นโปรตีนที่เกี่ยวข้องกับการควบคุมในเชิงลบของการตอบสนองในระบบภูมิคุ้มกัน กล่าวโดยสรุปนั้นชีววิทยาเชิงระบบสามารถนำมาใช้ในการค้นหาค้นหาองค์ความรู้ใหม่ภายใต้การวิเคราะห์ข้อมูลขนาดใหญ่ที่ไม่สามารถวิเคราะห์ได้ภายใต้วิธีการพื้นฐานทั่วไป ซึ่งนำไปสู่การค้นพบกลุ่มยีนใหม่จากชุดข้อมูลไมโครอาร์เรย์ที่ทำให้สามารถนำมาสร้างโมเดลเพื่อทำความเข้าใจกลไกในระดับโมเลกุลของการตอบสนองการอักเสบได้

Research Project Systems Biology for Investigating the Mechanism of Inflammatory Response in Murine Macrophage

Head of Project Asst. Prof. Dr. Pitak Sootanan

Co-Researcher Assoc. Prof. Dr. Klaokwan Srisook

Affiliation Department of Biochemistry, Faculty of Science, Burapha University

This research was supported by Burapha University through its grant number 87/2557.

Abstract

Macrophage RAW264.7 stimulated with LPS is a model to learn about the mechanism of inflammation. It was used by several research groups for testing anti-inflammatory compounds by using microarray technology. This technique can simultaneously detect the response on whole cell in interesting condition by measurement of mRNA levels. There was a huge of submitted microarray datasets of LPS-stimulated macrophage in public database. These data can be used for analysis with several bioinformatics tools or system biology approach to investigate the mechanism of inflammatory response in murine macrophage. Gene selection, co-expression network, gene regulatory network, biological data integration, and biological interpretation techniques were used in this study for finding some interested modules and their responsive genes. One selected inflammatory responsive gene, *SOCS3*, has already validated with real-time PCR to measure its expression level on time-series. *SOCS3* is gene that encoded for suppressor of cytokine signaling protein which is negative regulator of immune responses. In conclusion, system biology can be used to discovered more knowledge which are concealed under a huge of data that cannot be found by conventional methods. This leads to find a novel responsive gene from the traditional public microarray data that make a new model for understanding the mechanism of inflammatory response.

ประกาศคุณูปการ

โครงการวิจัยฉบับนี้สามารถสำเร็จลุล่วงได้ด้วยดีจากการช่วยเหลือและร่วมกันทำงานของนิสิตระดับปริญญาตรีที่ทำโครงการวิจัยระดับปริญญาตรีสาขาชีวเคมีร่วมกัน ได้แก่ นายวันเฉลิม จินอู่, นายเทพพนม อรัญญาณ, นายจรัญ รุ่งเรือง, นางสาวพศิกา ไบยา, นางสาว วิไลพร ศรีตะบุตร, นายกิตติพันธ์ ธาระเชษฐ์, นางสาวอรอนงค์ จันทุมมี, และนางสาวปราณี พรหมคล้าย รวมถึงนิสิตระดับปริญญาโท ได้แก่ นางสาวฐิติรัตน์ นุชศิลา, นางสาวจิตาภา สอนศิริ, และนางสาวเหมือนฝัน โวหารกล้า โดยเฉพาะนางสาวเหมือนฝัน โวหารกล้า ที่ร่วมกันฝึกฝนอุปกรณ์และทำงานมาด้วยกันตั้งแต่ทำโครงการวิจัยระดับปริญญาตรี สาขาชีวเคมี จนกระทั่งศึกษาต่อในระดับปริญญาโท สาขาวิทยาศาสตร์ชีวภาพที่ถือว่าเป็นหัวแรงหลักของการทำโครงการวิจัยครั้งนี้ เนื่องจากต้องพบอุปสรรคมากมายโดยเฉพาะเรื่องของเครื่องมือและอุปกรณ์การวิจัย และขอขอบพระคุณ รองศาสตราจารย์ ดร. กล่าวขวัญ ศรีสุข ที่เป็นให้เกียรติเป็นผู้ร่วมโครงการวิจัย พร้อมทั้งลูกศิษย์ทั้งระดับปริญญาตรี ปริญญาโท และปริญญาเอก ในห้องปฏิบัติการเพาะเลี้ยงเซลล์ ที่ได้ช่วยเหลือในขั้นตอนของการทดลองในระดับห้องปฏิบัติการ ที่ต้องอาศัยความเชี่ยวชาญและทักษะในการเพาะเลี้ยงเซลล์และการทำปฏิบัติการที่ต้องอาศัยเทคนิคทางชีววิทยาระดับโมเลกุล ทำให้การทำวิจัยในห้องปฏิบัติการประสบผลสำเร็จลงได้ และขอขอบคุณผู้มีส่วนช่วยเหลืออื่นที่ไม่ได้เอ่ยนาม โดยเฉพาะครอบครัวที่มีส่วนสำคัญ และเป็นแรงใจจนสามารถทำเล่มรายงานเพื่อปิดโครงการวิจัยนี้ลงได้

ผู้ช่วยศาสตราจารย์ ดร. พิทักษ์ สุตรอนันต์

หัวหน้าโครงการวิจัย

31 สิงหาคม พ.ศ. 2559

สารบัญเรื่อง

	หน้า
ปกใน	ก
บทคัดย่อภาษาไทย	ข
บทคัดย่อภาษาอังกฤษ	ค
ประกาศคุณูปการ	ง
สารบัญเรื่อง	จ
สารบัญตาราง	ช
สารบัญภาพ	ณ
บทที่ 1 บทนำ	1
1.1 ความสำคัญและที่มาของปัญหาที่ทำการวิจัย	1
1.2 วัตถุประสงค์ของโครงการวิจัย	2
1.3 ขอบเขตของโครงการวิจัย	2
1.4 ประโยชน์ที่คาดว่าจะได้รับ	3
1.5 แผนการถ่ายทอดเทคโนโลยีหรือผลการวิจัยสู่กลุ่มเป้าหมาย	3
บทที่ 2 ทฤษฎี และงานวิจัยที่เกี่ยวข้อง	5
2.1 ทฤษฎี สมมติฐาน (ถ้ามี) และกรอบแนวความคิดของโครงการวิจัย	5
2.2 การทบทวนวรรณกรรม/สารสนเทศ (information) ที่เกี่ยวข้อง	6
2.2.1 การอักเสบ (Inflammation)	6
2.2.2 การวิเคราะห์ไมโครอาร์เรย์ของการแสดงออกของยีน	7
2.2.3 ข้อมูลการศึกษาไมโครอาร์เรย์ของการแสดงออกของยีนในเซลล์แมคโครฟาจหนู และการบูรณาการข้อมูลโดยอาศัยหลักการชีววิทยาเชิงระบบ	8
บทที่ 3 วัสดุ อุปกรณ์ และวิธีการทดลอง	9
3.1 วัสดุ และอุปกรณ์	9
3.1.1 เครื่องมือและโปรแกรมทางชีวสารสนเทศ	9

สารบัญเรื่อง (ต่อ)

	หน้า
3.1.2 วัสดุและอุปกรณ์ทางห้องปฏิบัติการ	10
3.2 สารเคมี	11
3.3 วิธีการทดลอง	11
3.3.1 การรวบรวมและการเตรียมข้อมูลไมโครอาร์เรย์	12
3.3.2 การค้นหาชุดยีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบ	13
3.3.3 การศึกษาและออกแบบแพลตฟอร์มจากข้อมูลไมโครอาร์เรย์	14
3.3.4 การบูรณาการข้อมูลแพลตฟอร์มที่ได้จากข้อมูลไมโครอาร์เรย์	14
3.3.5 การตรวจสอบและยืนยันผลยีนที่ถูกคัดเลือกที่มีการตอบสนองต่อการอักเสบ เมื่อมีการกระตุ้นด้วย LPS ตามช่วงเวลา ในระดับอาร์เอ็นเอ	15
บทที่ 4 ผลการทดลอง	20
4.1 การค้นหากระบวนการทางชีวสารสนเทศที่เหมาะสม	20
4.2 การเตรียมข้อมูลไมโครอาร์เรย์ และการคัดเลือกกลุ่มยีนที่เกี่ยวข้องกับการตอบสนอง ต่อการอักเสบ	21
4.3 การออกแบบและบูรณาการแพลตฟอร์มจากข้อมูลไมโครอาร์เรย์	24
4.4 การวิเคราะห์กลุ่มยีนที่มีแนวโน้มการแสดงออกร่วมกัน และการวิเคราะห์ร่วมกับ เครือข่ายปฏิสัมพันธ์ระหว่างโปรตีน	27
4.5 การวิเคราะห์เครือข่ายปฏิสัมพันธ์ และการแปลผลทางชีวภาพ	29
4.6 การคัดเลือกยีนและการตรวจสอบการแสดงออกของยีนระดับ mRNA ใน ห้องปฏิบัติการโดยใช้เทคนิค real time PCR	31
บทที่ 5 อภิปราย สรุปผลการทดลอง และข้อเสนอแนะ	33
5.1 อภิปรายผลการทดลอง	33
5.2 สรุปผลการทดลอง	37
5.3 ข้อเสนอแนะ	37
เอกสารอ้างอิง	38
ภาคผนวก	43

สารบัญเรื่อง (ต่อ)

	หน้า
ภาคผนวก ก นำเสนอรายงานสืบเนื่องจากงานประชุมทางวิชาการระดับชาติ ปี 2559	44
ภาคผนวก ข นำเสนอรายงานสืบเนื่องจากงานประชุมทางวิชาการระดับนานาชาติ ปี 2558	57
ภาคผนวก ค นำเสนอรายงานสืบเนื่องจากงานประชุมทางวิชาการระดับชาติ ปี 2558	64
ภาคผนวก ง นำเสนอบทความย่อ งานประชุมทางวิชาการระดับนานาชาติ ปี 2558	71
ภาคผนวก จ นำเสนอรายงานสืบเนื่องจากงานประชุมทางวิชาการระดับชาติ ปี 2557	73
ภาคผนวก ฉ บทความวิจัย วารสารระดับชาติ ปี 2557	81
ภาคผนวก ช นำเสนอรายงานสืบเนื่องจากงานประชุมทางวิชาการระดับชาติ ปี 2557	94
ภาคผนวก ซ นำเสนอบทความย่อ งานประชุมทางวิชาการระดับนานาชาติ ปี 2557	101
ภาคผนวก ฌ นำเสนอบทความย่อ งานประชุมทางวิชาการระดับชาติ ปี 2556	103

สารบัญตาราง

ตารางที่		หน้า
3-1	แสดงรายละเอียดของชุดข้อมูลไมโครอาร์เรย์	13
3-2	ลำดับนิวคลีโอไทด์ของไพรเมอร์ที่ใช้ในการทดลอง	17
3-3	สารและปริมาตรที่ใช้ในการทำปฏิกิริยา real-time PCR	18
3-4	สภาวะที่ใช้ในปฏิกิริยา real-time PCR	18
4-1	จำนวนโพรบที่ถูกคัดเลือกโดยวิธี SAM, eBayes , PLS-CV และโพรบที่ถูกคัดเลือกอย่างน้อย 2 วิธี จาก 3 วิธี	22
4-2	ยีนที่พบในแต่ละชุดข้อมูลไมโครอาร์เรย์จากโพรบที่ถูกคัดเลือกอย่างน้อยสองวิธี	23
4-3	รายละเอียดของยีนในแต่ละเครือข่ายย่อยทั้งสองเครือข่าย	26
4-4	การจัดกลุ่มยีนที่มีการแสดงออกพร้อมกันเมื่อมีการกระตุ้นด้วย LPS ที่เวลา 3 6 8 และ 18 ชั่วโมง โดยวิธี SOM (self-organizing map)	28
4-5	Biological process (BP) ของยีนในทั้งสองเครือข่ายย่อยโดยใช้ยีนออนโทโลยี	29
4-6	Cellular component (CC) ของยีนในทั้งสองเครือข่ายย่อยโดยใช้ยีนออนโทโลยี	30
4-7	Molecular function ของยีนในทั้งสองเครือข่ายย่อยโดยใช้ยีนออนโทโลยี	30
4-8	หน้าที่และกระบวนการทางชีวภาพของยีนในทั้งสองเครือข่ายย่อยฐานข้อมูล Kyoto Encyclopedia of Genomes pathway enrichment analysis (KEGG)	30

สารบัญภาพ

ภาพที่		หน้า
2-1	การบูรณาการข้อมูลระดับจีโนมเพื่อศึกษาการติดเชื้อไวรัสเอชไอวีซึ่งเป็นระบบโรคเชิงซ้อนโดยใช้หลักการทางชีววิทยาระบบ (Tan, 2007)	6
3-1	ผังงานของโครงการชีววิทยาเชิงระบบสำหรับสืบค้นกลไกการตอบสนองต่อการอักเสบในเซลล์แมคโครฟาจของหนู	12
4-1	ข้อมูลไมโครอาร์เรย์ทั้ง 3 ชุดข้อมูลก่อนและหลังผ่านกระบวนการปรับมาตรฐานข้อมูล Box plot raw data (A), log ₁₀ -transformed data (B) และ normalized data (C)	21
4-2	ไดอะแกรมแสดงจำนวนโพรบที่ถูกคัดเลือกโดยวิธี SAM, eBayes , PLS-CV และโพรบที่ถูกคัดเลือกอย่างน้อย 2 วิธี	23
4-3	เครือข่ายการแสดงออกร่วมของยีนที่ได้จากการคัดเลือกจากโปรแกรม ArrayMining และยีนที่ได้จากการสืบค้นจากฐานข้อมูล PolySearch	25
4-4	เครือข่ายย่อยที่ถูกคัดเลือกจากเครือข่ายการแสดงออกร่วมของยีน	25
4-5	Heat map รูปแบบการแสดงออกของยีนทั้งสามกลุ่ม (A) และกราฟแสดงแนวโน้มรูปแบบการแสดงออกของยีน (B) เมื่อถูกกระตุ้นด้วย LPS ที่เวลา 3, 6, 8 และ 18 ชั่วโมง จากการจัดกลุ่มตามรูปแบบการแสดงออกของยีนโดยวิธี SOM	27
4-6	ปฏิสัมพันธ์ระหว่างโปรตีนของยีนในเครือข่ายย่อยทั้งสองเครือข่ายในฐานข้อมูล STRING	28
4-7	วิธีการส่งสัญญาณ Jak-STAT signaling pathway	31
4-8	ระดับการแสดงออกของ mRNA ของยีน iNOS ในเซลล์แมคโครฟาจ RAW264.7 ภายใต้สภาวะที่มีการกระตุ้นด้วย LPS ตามช่วงเวลา 3, 6, 8, 18, และ 24 ชั่วโมง	32
4-9	ระดับการแสดงออกของ mRNA ของยีน SOCS3 ในเซลล์แมคโครฟาจ RAW264.7 ภายใต้สภาวะที่มีการกระตุ้นด้วย LPS ตามช่วงเวลา 3, 6, 8, 18, และ 24 ชั่วโมง	32

บทที่ 1

บทนำ

1.1 ความสำคัญและที่มาของปัญหาที่ทำการวิจัย

ชีววิทยาเชิงระบบ (systems biology) เป็นการศึกษาเพื่อทำการเข้าใจระบบและกลไกทางชีวภาพในแบบสหวิทยาการ ในการพิจารณาปฏิสัมพันธ์ระหว่างชีวโมเลกุล/สารเคมีภายในระบบทางชีวภาพ โดยอาศัยการติดตามและตรวจวัดชีวโมเลกุล/สารเคมีทั้งระบบที่มีการแสดงออกภายใต้เซลล์หรือสภาวะที่ศึกษา ซึ่งการตรวจวัดดังกล่าวจะทำให้ได้ข้อมูลการแสดงออกของชีวโมเลกุล/สารเคมีเป็นจำนวนมากในการตรวจวัดพร้อมกันเพียงครั้งเดียว (Kitano, 2002) จากการศึกษาข้อมูลรหัสพันธุกรรมระดับจีโนมที่มีเพิ่มมากขึ้นในสิ่งมีชีวิตชนิดต่างๆ โดยเฉพาะในมนุษย์และสัตว์เลี้ยงลูกด้วยนม ทำให้มีการนำข้อมูลรหัสพันธุกรรมระดับจีโนมนี้ไปใช้ในการศึกษาเชิงหน้าที่ระดับจีโนม (functional genomics) ที่มีความนิยมกันอย่างแพร่หลายก็คือ การนำข้อมูลรหัสพันธุกรรมระดับจีโนมมาออกแบบไมโครอาร์เรย์เพื่อศึกษาการแสดงออกของยีนทั้งหมดภายในเซลล์หรือที่เรียกว่าทรานสคริปโตมิกส์ (transcriptomics) โดยเฉพาะการศึกษาในสัตว์เลี้ยงลูกด้วยนมที่มีความใกล้ชิดกับมนุษย์คือหนู (murine) ทำให้ในปัจจุบันมีข้อมูลที่ได้จากการศึกษาไมโครอาร์เรย์ของการแสดงออกของยีน โดยเฉพาะที่เกี่ยวข้องกับเซลล์แมคโครฟาจ ของหนูสำหรับศึกษากลไกการตอบสนองต่อสัญญาณจากสิ่งแปลกปลอมของเซลล์จุลินทรีย์โดยเฉพาะไลโปโพลีแซคคาไรด์ (lipopolysaccharide, LPS) นี้เป็นองค์ประกอบสำคัญของผนังเซลล์ชั้นนอกของแบคทีเรียแกรมลบ (gram-negative bacteria) (Poltorak, 1998) ซึ่งก่อให้เกิดโรคที่เกี่ยวข้องกับการเกิดการอักเสบในร่างกาย (inflammatory response) ได้ การรวบรวมข้อมูลที่ได้จากการศึกษาไมโครอาร์เรย์ของการแสดงออกของยีนในประเด็นข้างต้นจากหลากหลายกลุ่มวิจัย จะทำให้เราสามารถสืบค้นกลไกที่สัมพันธ์กับการตอบสนองของเซลล์แมคโครฟาจของหนูต่อ LPS หรือกลไกที่เกี่ยวข้องกับการอักเสบได้ดีมากยิ่งขึ้น

จากข้อมูลไมโครอาร์เรย์การแสดงออกของยีน ที่เป็นการตรวจวัดชนิดและปริมาณอาร์เอ็นเอที่ถูกสร้างขึ้นมาของสิ่งมีชีวิตภายใต้สภาวะที่ศึกษาในขณะช่วงเวลาใดเวลาหนึ่ง ซึ่งการศึกษาไมโครอาร์เรย์นี้ขึ้นอยู่กับความเข้าใจข้อมูลพันธุกรรมระดับจีโนมของสิ่งมีชีวิตนั้นๆ รวมถึงการออกแบบไมโครอาร์เรย์ที่ใช้ในการศึกษาว่าจะครอบคลุมยีนหรือรหัสพันธุกรรมของสิ่งมีชีวิตที่สนใจมากน้อยเพียงไร (Lipshutz, 1995 และ Schena, 1995) นอกจากนั้นการใช้เพียงข้อมูลไมโครอาร์เรย์ของการแสดงออกของยีนอาจทำให้ได้ข้อมูลที่ไม่ครบถ้วนสมบูรณ์พอที่จะนำมาอธิบายกลไกการตอบสนองต่อ LPS ที่เกิดขึ้นภายในเซลล์แมคโครฟาจของหนูอย่างเป็นระบบได้ เนื่องจากการใช้ข้อมูลเพียงระดับเดียวคือปริมาณการแสดงออกของยีนในระดับปริมาณของอาร์เอ็นเอเท่านั้น หากมีการใช้ข้อมูลการตรวจวัดชีวโมเลกุลชนิดอื่นภายในสิ่งมีชีวิต ที่สำคัญคือการใช้ข้อมูลการแสดงออกระดับโปรตีน หรือการตรวจวัดชนิดและปริมาณโปรตีนที่ถูกสร้างขึ้นมาในเซลล์ภายใต้สภาวะและช่วงเวลาที่สนใจ ก็จะทำให้การอธิบายกลไกการตอบสนองของเซลล์เป็นระบบที่มีความครบถ้วนสมบูรณ์เพิ่มมากยิ่งขึ้น (Ge, 2003) จากการศึกษาโดย Chantimal (2012) ที่ได้ทำการศึกษาการแสดงออกระดับโปรตีน (proteomics) ด้วยเทคนิค nano-LC-MS/MS ในเซลล์แมคโครฟาจของหนูในการตอบสนองต่อ LPS รวมทั้งได้พิจารณาผลที่เกิดขึ้นของการใช้ EAMA (Ethyl 2-Acetyl-3-(4-hydroxy-3-Methoxyphenyl) Acrylate) ที่มีผลในการยับยั้งการแสดงออกของโปรตีนที่เกี่ยวข้องการผลิต ไนตริกออกไซด์ (nitric oxide, NO) ซึ่งเป็นต้นตอของการอักเสบที่เกิดขึ้นได้ การศึกษาโดย Chantimal (2012) นี้ นอกจากผลที่ได้ข้างต้น ยังทำให้เรามี

ข้อมูลการแสดงออกของโปรตีนทั้งหมดภายใต้สภาวะการตอบสนองต่อ LPS ของเซลล์แมคโครฟาจที่เราสามารถนำมาใช้เพื่อช่วยอธิบายกลไกการตอบสนองได้อย่างสมบูรณ์มากยิ่งขึ้นนอกจากการใช้ข้อมูลการแสดงออกของยีนระดับอาร์เอ็นเอจากข้อมูลไมโครอาร์เรย์

การรวบรวมข้อมูลไมโครอาร์เรย์ของการแสดงออกของยีนที่ศึกษาเกี่ยวกับการตอบสนองต่อ LPS ในเซลล์แมคโครฟาจของหนู บนฐานข้อมูลสาธารณะ เช่น GEO (Edgar, 2002, Gene Expression Omnibus, <http://www.ncbi.nlm.nih.gov/geo/>) จากหลากหลายกลุ่มวิจัย ช่วยทำให้เรามีข้อมูลที่น่าเชื่อถือมากยิ่งขึ้นเมื่อนำมาวิเคราะห์เพื่อค้นหายีนที่มีความสำคัญและสามารถอธิบายกลไกที่มีส่วนเกี่ยวข้องกับการตอบสนองต่อ LPS ของเซลล์แมคโครฟาจของหนูได้ดีมากยิ่งขึ้นกว่าการใช้ข้อมูลจากกลุ่มวิจัยเพียงกลุ่มเดียว นอกจากนี้ผลที่ได้จากการวิเคราะห์ด้วยกระบวนการทางชีววิทยาเชิงระบบของข้อมูลไมโครอาร์เรย์ก็จะเป็นข้อมูลที่สำคัญที่จะนำมาใช้ในการสร้างแพลตฟอร์มของรูปแบบการถ่ายทอดสัญญาณของการแสดงออกในระดับอาร์เอ็นเอเพื่อนำมาใช้บูรณาการร่วมกับข้อมูลการศึกษาการแสดงออกระดับโปรตีน ซึ่งเป็นข้อมูลที่ได้จากการศึกษาโดยทีมงานผู้วิจัย ก็จะทำให้เราสามารถหาหรือโปรตีนเป้าหมายเพื่อนำมาตรวจสอบและยืนยันผลเพิ่มเติมทางห้องปฏิบัติการ เพื่อทำการตรวจวัดปริมาณการแสดงออกของยีนและโปรตีนเป้าหมายเหล่านั้น ซึ่งจะทำให้เราสามารถทำความเข้าใจและอธิบายกลไกการตอบสนองต่อ LPS ในเซลล์แมคโครฟาจของหนูที่สมบูรณ์มากยิ่งขึ้น ซึ่งมีส่วนสำคัญต่อการพัฒนาแพลตฟอร์มสำหรับใช้ช่วยในการตัดสินใจสำหรับการศึกษาผลของสารสกัดจากพืชสมุนไพร หรือสารสังเคราะห์ที่ออกฤทธิ์ยับยั้งกระบวนการอักเสบ ที่มีสำคัญต่อการเกิดการอักเสบในหนูซึ่งเป็นโมเดลที่สำคัญต่อการพัฒนาการรักษาโรคที่เกี่ยวข้องในมนุษย์ต่อไป

1.2 วัตถุประสงค์ของโครงการวิจัย

1. เพื่อตรวจสอบข้อมูลไมโครอาร์เรย์ของการแสดงออกของยีนจากหลากหลายกลุ่มวิจัยบนฐานข้อมูลสาธารณะที่ศึกษาเกี่ยวกับการตอบสนองต่อ LPS ในเซลล์แมคโครฟาจของหนู สำหรับนำมาใช้ค้นหากลุ่มของยีนเป้าหมาย พร้อมสร้างแพลตฟอร์มของการถ่ายทอดสัญญาณการแสดงออกระดับอาร์เอ็นเอที่เกิดจากการบูรณาการกับข้อมูลทางชีวภาพระดับอื่นๆ ได้
2. เพื่อพัฒนาแพลตฟอร์มของกลไกการตอบสนองต่อ LPS ในเซลล์แมคโครฟาจของหนู สำหรับนำมาเป็นโมเดลเพื่อใช้ในการศึกษาผลการออกฤทธิ์ของสารสกัดจากพืชสมุนไพร หรือสารสังเคราะห์ในการยับยั้งการอักเสบ

1.3 ขอบเขตของโครงการวิจัย

1. ค้นหากลุ่มยีนเป้าหมายโดยอาศัยการบูรณาการร่วมกันของชุดข้อมูลไมโครอาร์เรย์ที่ได้จากฐานข้อมูลสาธารณะ การใช้ข้อมูลการแสดงออกระดับโปรตีน และการใช้ข้อมูลทางชีวภาพระดับอื่นๆ มาพัฒนาแพลตฟอร์มที่เกี่ยวข้องกับการตอบสนองต่อ LPS ของเซลล์แมคโครฟาจหนู
2. ทำการตรวจสอบแพลตฟอร์มของกลไกการตอบสนองที่สร้างขึ้นมา โดยการยืนยันผลการแสดงออกของยีนด้วยการตรวจวัดปริมาณอาร์เอ็นเอของยีนเป้าหมายจากกลไกที่ถูกคัดเลือก ภายใต้การทดลองในห้องปฏิบัติการ

3. นำแพลตฟอร์มที่ได้ตรวจสอบและปรับปรุงแล้ว มาใช้ในการอธิบายกลไกการตอบสนองต่อ LPS ของเซลล์แมคโครฟาจ เพื่อนำมาใช้ช่วยตัดสินใจในการศึกษาผลของสารสกัดจากพืชสมุนไพร หรือสารสังเคราะห์ที่มีต่อการยับยั้งการอักเสบต่อไปได้

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้กลุ่มยีนหรือโปรตีนเป้าหมายเพื่อนำมาใช้อธิบายกลไกที่เกี่ยวข้องกับการตอบสนองต่อ LPS ในเซลล์แมคโครฟาจของหนูได้
2. การสังเคราะห์เทคโนโลยีและองค์ความรู้ทางชีวสารสนเทศทางการแพทย์ จะช่วยเพิ่มศักยภาพในการเผยแพร่เทคโนโลยี องค์ความรู้และบริการต่อสาธารณะ
3. เผยแพร่องค์ความรู้ และบทความวิจัย ในรูปแบบสิ่งตีพิมพ์ และงานประชุมวิชาการในระดับชาติและ/หรือนานาชาติ

1.5 แผนการถ่ายทอดเทคโนโลยีหรือผลการวิจัยสู่กลุ่มเป้าหมาย

การนำเสนอเป็นโปสเตอร์ บทความย่อ รายงานสืบเนื่องการประชุม ผ่านการประชุมวิชาการระดับชาติ และนานาชาติ และบทความวิจัยในวารสารระดับชาติ ดังนี้

1. อรอนงค์ จันทุมมี ปราณิ พรหมคล้าย จิตาภา สอนศิริ เหมือนฝัน โวหารกล้า และพิทักษ์ สุตรอนันต์. (2559). การคัดเลือกกลุ่มยีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบด้วยการจัดกลุ่มยีนและการวิเคราะห์เครือข่ายการแสดงออกของยีน. รายงานสืบเนื่องจากการประชุมวิชาการระดับชาติ “วิทยาศาสตร์วิจัย” ครั้งที่ 8, วันที่ 30 – 31 พฤษภาคม 2559, มหาวิทยาลัยพะเยา จังหวัดพะเยา, หน้า 241 – 246. (ภาคผนวก ก)
2. กิตติพันธ์ ธาระเขตร์ ฐิติรัตน์ นุชศิลา เหมือนฝัน โวหารกล้า และพิทักษ์ สุตรอนันต์. (2559). การค้นหายีนที่เกี่ยวข้องกับการตอบสนองการอักเสบจากข้อมูลไมโครอาร์เรย์โดยการคัดเลือกและการวิเคราะห์เครือข่ายการแสดงออกร่วม. รายงานสืบเนื่องจากการประชุมวิชาการระดับชาติ “วิทยาศาสตร์วิจัย” ครั้งที่ 8, วันที่ 30 – 31 พฤษภาคม 2559, มหาวิทยาลัยพะเยา จังหวัดพะเยา, หน้า 235 – 240. (ภาคผนวก ก)
3. Wohankla, M., Srisook, K. and Soonanan, P. (2015). Identification of low differentially expressed responsive genes in microarray data of LPS-stimulated macrophage RAW264.7. The 27th Annual Meeting of the Thai Society for Biotechnology and International Conference (TSB2015). November 17-20, 2015 Mandarin Hotel Bangkok by Centre Point, Bangkok, Thailand. 752 – 757. (ภาคผนวก ข)
4. เทพพนม อารัญญาญุ จรรย์ รุ่งเรือง เหมือนฝัน โวหารกล้า และพิทักษ์ สุตรอนันต์. (2558). เครือข่ายความสัมพันธ์การแสดงออกร่วมของกลุ่มยีนในเซลล์แมคโครฟาจ RAW 264.7 เมื่อถูกกระตุ้นด้วย LPS. รายงานสืบเนื่องจากการประชุมวิชาการระดับชาติ “วิทยาศาสตร์วิจัย” ครั้งที่ 7, วันที่ 30-31 มีนาคม 2558, มหาวิทยาลัยนเรศวร จังหวัดพิษณุโลก, BI-P-048, 6 หน้า. (ภาคผนวก ค)
5. Wohankla, M. and Sootanan, P. 2015. “Inflammatory network inferred from microarray time series data of LPS-stimulated macrophage”. PP17. The 1st

International Conference of Pharmaceutical Sciences and Medicines, 22 – 23 January 2015, Burapha University, Thailand. (ภาคผนวก ง)

6. วันเฉลิม จินอู๋ และพิทักษ์ สุตรอนันต์. (2557). กลไกการตอบสนองการอักเสบจากการวิเคราะห์ข้อมูลไมโครอาร์เรย์. รายงานการประชุมวิชาการ การศึกษาทางพันธุศาสตร์และจีโนมิกส์ด้วยเครื่องวิเคราะห์ยุคใหม่, วันที่ 29–30 กรกฎาคม พ.ศ. 2557, กรุงเทพฯ, 25 – 31. (ภาคผนวก จ)
7. พศิกา ไบยา วิไลพร ศรีตะบุตร และพิทักษ์ สุตรอนันต์. (2557). เครือข่ายควบคุมการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบจากข้อมูลไมโครอาร์เรย์. วารสารวิทยาศาสตร์บูรพา ปีที่ 19 ฉบับพิเศษ การประชุมวิชาการระดับชาติ วิทยาศาสตร์วิจัยครั้งที่ 6, ISBN 0858-7612, 173–184. (ภาคผนวก ฉ)
8. เหมือนฝัน โวหารกล้า และพิทักษ์ สุตรอนันต์. (2557). การวิเคราะห์การแสดงออกของยีนในข้อมูลไมโครอาร์เรย์ของเซลล์แมคโครฟาจที่ถูกกระตุ้นด้วยไลโปโพลีแซคคาไรด์. รายงานสืบเนื่องจากการประชุมวิชาการระดับชาติ “วิทยาศาสตร์วิจัย” ครั้งที่ 6, วันที่ 20 – 21 มีนาคม 2557, มหาวิทยาลัยบูรพา จังหวัดชลบุรี, 195–200. (ภาคผนวก ช)
9. Sootanan, P. 2014. “Responsive-gene finding of LPS-stimulated macrophage via network approaches in microarray data”. FO-02. The 4th International Biochemistry and Molecular Biology Conference, April 2-3, 2014, Bangkok, Thailand. (ภาคผนวก ซ)
10. Wohankla, M. and Sootanan, P. 2556. “Comparative pathway-based analysis of gene expression profiles between two macrophages stimulated with LPS”. F_F0024. การประชุมวิชาการวิทยาศาสตร์และเทคโนโลยีแห่งประเทศไทย ครั้งที่ 39, ศูนย์นิทรรศการและการประชุมไบเทค, กรุงเทพฯ, ประเทศไทย, 21 – 23 ตุลาคม 2556, หน้า 176. (ภาคผนวก ฌ)
11. Arunphanu, T., Rungrueng, J. and Sootanan, P. 2556. “Coexpression network approach to explain the response of LPS-stimulated RAW264.7”. F_F0026. การประชุมวิชาการวิทยาศาสตร์และเทคโนโลยีแห่งประเทศไทย ครั้งที่ 39, ศูนย์นิทรรศการและการประชุมไบเทค, กรุงเทพฯ, ประเทศไทย, 21 – 23 ตุลาคม 2556, หน้า 177. (ภาคผนวก ฉ)
12. Ginou, W., Wichanprecha, K. and Sootanan, P. 2556. “Finding differentially expressed genes with data mining tools in microarray data of LPS-stimulated macrophage RAW264.7”. F_F0049. การประชุมวิชาการวิทยาศาสตร์และเทคโนโลยีแห่งประเทศไทย ครั้งที่ 39, ศูนย์นิทรรศการและการประชุมไบเทค, กรุงเทพฯ, ประเทศไทย, 21 – 23 ตุลาคม 2556, หน้า 186. (ภาคผนวก ฉ)

บทที่ 2

ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

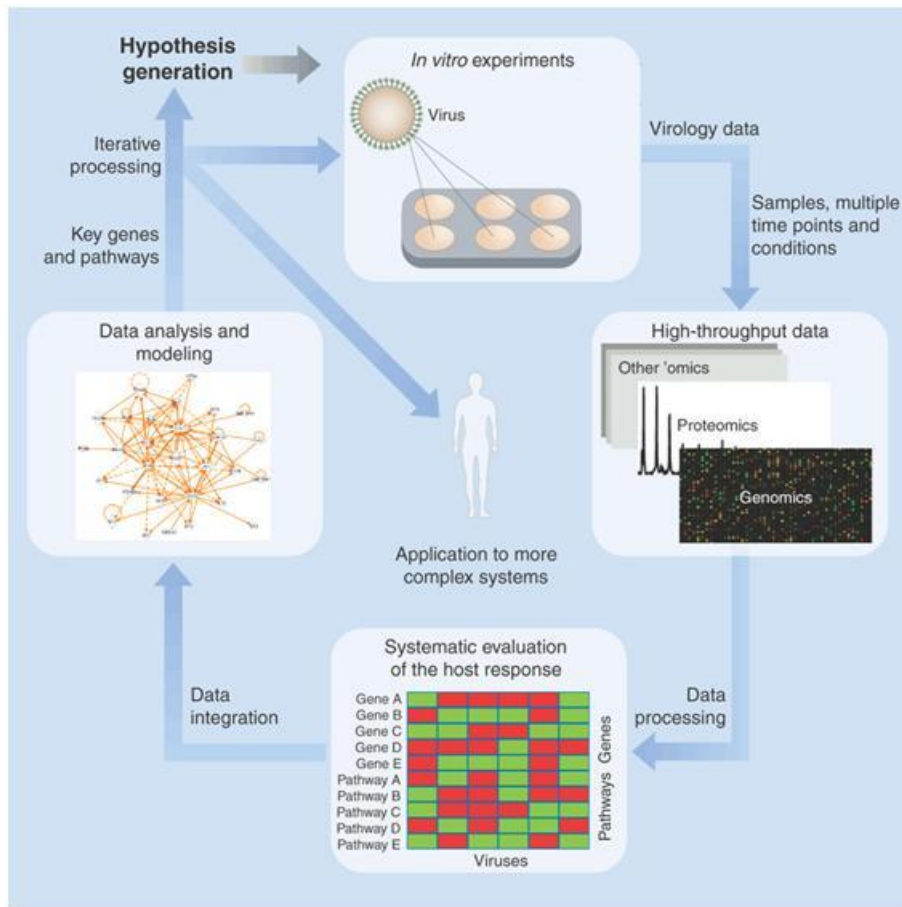
2.1 ทฤษฎี สมมุติฐาน (ถ้ามี) และกรอบแนวความคิดของโครงการวิจัย

ความรู้และเทคโนโลยีที่ใช้ในการศึกษาไกลระดับโมเลกุลภายในสิ่งมีชีวิตได้มีการพัฒนาอย่างต่อเนื่องนับตั้งแต่ได้การเริ่มต้นศึกษาข้อมูลระดับจีโนมของมนุษย์ (Lander, 2001) ซึ่งถือเป็นจุดเริ่มต้นของการศึกษาและทำความเข้าใจกลไกภายในสิ่งมีชีวิตระดับจีโนม หรือที่เรียกว่า ชีววิทยาเชิงระบบ ซึ่งอาศัยข้อมูลระดับจีโนมที่ได้ในการออกแบบการตรวจวัดชนิดและปริมาณของชีวโมเลกุลที่มีการแสดงออกในระดับต่างๆ เช่น การตรวจวัดชนิดและปริมาณการแสดงออกของยีนโดยการตรวจวัดอาร์เอ็นเอ และโปรตีน หรือการตรวจวัดชนิดและปริมาณการแสดงออกของเมตาบอไลต์ เพื่อนำมาศึกษาผลของการเปลี่ยนแปลงหรือการตอบสนองของสภาวะที่ศึกษา ซึ่งจะช่วยให้ได้ข้อมูลการแสดงออกของชีวโมเลกุลในแต่ละชนิดอย่างเป็นระบบ ทำให้เราสามารถทำความเข้าใจและอธิบายกลไก รวมถึงวิธีการส่งผ่านสัญญาณระดับโมเลกุลได้ครบถ้วน สมบูรณ์ได้ดีขึ้นกว่าการศึกษาชีวโมเลกุลในแบบดั้งเดิมที่เป็นการศึกษาผลการเปลี่ยนแปลงหรือการตอบสนองของชีวโมเลกุลเพียงชนิดเดียวหรือกลุ่มเดียว (Kitano, 2002)

เทคโนโลยีไมโครอาร์เรย์เป็นเทคโนโลยีที่ได้รับความนิยมเพื่อนำมาใช้ในการศึกษาการแสดงออกของยีนจำนวนมากเป็นหลักหมื่นยีนในช่วงเวลาเดียวกัน โดยการตรวจวัดชนิดและปริมาณของอาร์เอ็นเอของยีนทั้งหมดภายใต้สภาวะที่สนใจ ซึ่งการพัฒนาเทคโนโลยีที่มีอย่างรวดเร็วและเป็นที่ยอมรับในการใช้งาน จึงทำให้มีการสร้างข้อมูลจากการใช้งานไมโครอาร์เรย์ภายใต้สภาวะต่างๆ อยู่เป็นจำนวนมาก และมีการเก็บรวบรวมข้อมูลการแสดงออกของยีนด้วยไมโครอาร์เรย์เหล่านี้ในฐานะข้อมูลสาธารณะซึ่งเป็นฐานข้อมูลเปิดที่นักวิจัยโดยทั่วไปสามารถเข้าไปดาวน์โหลดเพื่อนำไปใช้ในการศึกษาวิจัยได้ (Edgar, 2002) ซึ่งข้อมูลการแสดงออกของยีนด้วยไมโครอาร์เรย์ที่มีอยู่เหล่านี้ก็มีประเด็นการศึกษาวิจัยที่หลากหลาย และก็มีหลายกลุ่มวิจัยที่ได้ทำการวิจัยในประเด็นที่เหมือนหรือคล้ายกัน ดังนั้นจึงทำให้เกิดชุดข้อมูลไมโครอาร์เรย์ที่ได้จากการศึกษาวิจัยประเด็นที่คล้ายคลึงกัน เช่น การศึกษาเกี่ยวกับกลไกการอักเสบ ซึ่งเป็นการศึกษาผลการตอบสนองต่อ LPS ในเซลล์แมคโครฟาจของหนู ภายใต้สภาวะการทดสอบกับสารออกฤทธิ์หรือการยับยั้งยีนที่แตกต่างกันไป (Gold, 2012 และ Kollipara, 2010) ซึ่งถือว่าเป็นการศึกษาที่เป็นโมเดลต้นแบบสำหรับศึกษาการตอบสนองต่อการอักเสบของเซลล์สัตว์เลี้ยงลูกด้วยนมที่จะมีการพัฒนาต่อเพื่อใช้ในการรักษาโรคที่เกี่ยวข้องกับการอักเสบในมนุษย์ต่อไป จากข้อมูลไมโครอาร์เรย์เหล่านี้ก็จะจะเป็นข้อมูลที่จะทำให้เราสามารถค้นหากลุ่มยีนหรือโปรตีนที่สำคัญต่อการตอบสนองต่อ LPS ในเซลล์แมคโครฟาจของหนู เพื่อนำมาใช้ในการพัฒนาแพลตฟอร์มที่มีการบูรณาการข้อมูลร่วมกับข้อมูลการแสดงออกระดับโปรตีนทั้งหมดซึ่งเป็นข้อมูลของผู้ร่วมวิจัยของทางกลุ่มวิจัยเอง ซึ่งจะช่วยให้ได้แพลตฟอร์มที่มีความครบถ้วนสมบูรณ์เพิ่มมากขึ้น เพื่อนำมาใช้ในการอธิบายกลไกระดับโมเลกุลที่เกิดขึ้น และสามารถนำมาเป็นโมเดลสำหรับช่วยตัดสินใจเพื่อการตรวจสอบทางห้องปฏิบัติการได้อย่างถูกต้องแม่นยำเพิ่มมากขึ้น

ตัวอย่างของการบูรณาการความรู้ที่ได้จากข้อมูลไมโครอาร์เรย์ของการแสดงออกระดับอาร์เอ็นเอและข้อมูลการแสดงออกระดับโปรตีนโดยใช้ชีววิทยาเชิงระบบแสดงดังภาพที่ 2-1 ซึ่งเป็นการศึกษาการตอบสนองของผู้ป่วยที่มีต่อการติดเชื้อไวรัสเอชไอวี (Tan, 2007) ภายหลังจากวิเคราะห์แบบบูรณาการของข้อมูลโอมิกส์ในแต่ละระดับแล้วโดยเฉพาะระดับอาร์เอ็นเอและโปรตีน และเมื่อผ่านการทำซ้ำของชุดข้อมูลที่หลากหลาย ก็

จะทำให้ได้กลุ่มยีนหรือพาหะที่สำคัญซึ่งเกี่ยวข้องกับลักษณะการตอบสนองต่อการติดเชื้อไวรัสที่สนใจ แล้วนำผลของการวิเคราะห์และโมเดลมาสร้างเป็นสมมติฐานใหม่เพื่อทำความเข้าใจกลไกการตอบสนองในเชิงลึกให้มากยิ่งขึ้น ก่อนที่จะมาตรวจสอบผลภายใต้การทดลองในห้องปฏิบัติการเพื่อยืนยันสมมติฐานข้างต้น แล้วนำองค์ความรู้ใหม่ที่ได้จากการศึกษาวิจัยเหล่านี้มาประยุกต์ใช้กับระบบของโรคเชิงซ้อนอื่นที่เกี่ยวข้องต่อไปได้



ภาพที่ 2-1 การบูรณาการข้อมูลระดับจีโนมเพื่อศึกษาการติดเชื้อไวรัสเอชไอวีซึ่งเป็นระบบโรคเชิงซ้อนโดยใช้หลักการทางชีววิทยาระบบ (Tan, 2007)

2.2 การทบทวนวรรณกรรม/สารสนเทศ (information) ที่เกี่ยวข้อง

2.2.1 การอักเสบ (Inflammation)

การอักเสบเป็นปฏิกิริยาที่ร่างกายตอบสนองต่อสิ่งที่ทำให้เนื้อเยื่อของร่างกายได้รับบาดเจ็บ เช่น เชื้อโรค การตายของเซลล์จากการขาดเลือดหรือขาดออกซิเจน กระบวนการอักเสบจัดเป็นกลไกสำคัญในการป้องกันสิ่งแปลกปลอม หรือสิ่งที่จะทำให้เซลล์หรือเนื้อเยื่อของร่างกายได้รับบาดเจ็บ ผลของการอักเสบจะทำให้เกิดการกำจัดสิ่งแปลกปลอมหรือเชื้อโรคออกไป นอกจากนี้ยังกำจัดเซลล์ที่ได้รับบาดเจ็บหรือตายจากการเข้ามาของสิ่งแปลกปลอมดังกล่าว (ปรีศนา โภคาพานิช, 2551 อ้างอิงจาก พีรยุทธ สิทธิไชยากุล, 2551) ซึ่งการอักเสบนั้นจะเป็นปฏิกิริยาที่ร่างกายตอบสนองแบบไม่จำเพาะต่อการติดเชื้อหรือไม่ใช่การติดเชื้อก็ได้ เมื่อเริ่มต้น ภูมิคุ้มกันของร่างกายจะทำการควบคุมและทำลายสิ่งแปลกปลอมก่อนการอักเสบในทันที รวมถึงการ

กำจัดเนื้อเยื่อที่เสียหาย โดยจะเรียกการตอบสนองเช่นนี้ว่า การอักเสบเฉียบพลัน (acute inflammation) อาการที่เกิดขึ้นคืออาการบวม แดง ร้อน และเจ็บปวดบริเวณของเนื้อเยื่อ การบวมนี้เกิดขึ้นเนื่องจากการมีของเหลวหลั่งออกมาจากหลอดเลือดฝอยรวมตัวอยู่ในเนื้อเยื่อ ทำให้เซลล์จำพวกฟาโกไซท์ (phagocyte) เช่น เซลล์แมคโครฟาจ (macrophage) ทำการสร้างไนตริกออกไซด์ (nitric oxide, NO) ที่ทำหน้าที่เป็นสารสื่อสัญญาณชีวภาพที่เป็นพิษต่อแบคทีเรียและเชื้อโรคที่บุกรุกเข้ามาในร่างกาย โดยอาศัยการทำงานของเอนไซม์ inducible nitric oxide synthase (iNOS) ในการตอบสนองของระบบภูมิคุ้มกันระหว่างการติดเชื้อ ทำให้ cyclic guanosine monophosphate (cGMP) เพิ่มขึ้น เป็นผลให้เกิดการขยายตัวของเส้นเลือด ซึ่งทำให้มีเลือดมาหล่อเลี้ยงบริเวณนั้นมากเกินไป ส่งผลให้เกิดการรื้อแดงและเนื้อเยื่อบริเวณนั้นมีการอักเสบ เป็นหนอง

รวมทั้งยังมีการหลั่งสารสื่อการอักเสบอื่น เช่น โพรสตาแกลนดิน E2 (prostaglandin E2) และไซโตไคน์ (cytokine) เป็นต้น นอกจากนี้ยังมีเซลล์ชนิดอื่น เช่น แบโซฟิล (basophil) ที่จะทำการปล่อยสารเคมีพวก ฮิสตามีน (histamine) มาทำให้หลอดเลือดขยายตัว มีเลือดไหลมาบริเวณนั้นมาก ในขณะเดียวกันก็ยอมให้ของเหลวในหลอดเลือดผ่านออกมาก ซึ่งจะทำให้เกิดการอักเสบได้ แต่ถ้าระบบภูมิคุ้มกันของร่างกายไม่สามารถควบคุมสาเหตุการอักเสบแบบเฉียบพลันได้ ก็จะทำให้เกิดการทำลายและลุกลามอย่างต่อเนื่อง โดยจะเรียกสภาวะนั้นว่าการอักเสบเรื้อรัง (chronic inflammation) ที่ส่งผลให้เซลล์หรือเนื้อเยื่อถูกทำลายมากขึ้น จนอาจทำให้กลายเป็นโรคต่างๆ เช่น โรคมะเร็ง โรคไขข้ออักเสบ โรคหลอดเลือดแดงแข็งตัว โรคเบาหวาน เป็นต้น (มัลลิกา ปาละโชติ, 2550 อ้างอิงจาก นงลักษณ์ สุวรรณพิณิจ, 2547) โดย Srisook และคณะ (2011) ได้ทำการศึกษาสารสกัดเอทิลอะซิเตท (ethyl acetate extract) จากต้นเพชรสังฆาตที่แสดงให้เห็นถึงฤทธิ์การต่อต้านการอักเสบ (anti-inflammatory activity) โดยการตรวจสอบการตอบสนองของเซลล์แมคโครฟาจต่อ LPS ซึ่งผลการทดลองแสดงให้เห็นว่าสารสกัดจากพืชสมุนไพรนี้สามารถยับยั้งการผลิตไนตริกออกไซด์ในเซลล์แมคโครฟาจ ซึ่งสารสกัดจากพืชสมุนไพรนี้จะส่งผลทำให้การทำงานของเอนไซม์ iNOS ลดลง และสารสกัดจากพืชสมุนไพรยังมีผลต่อการแสดงออกที่เพิ่มขึ้นของเอนไซม์ ฮีมออกซิจีเนส-1 (heme oxygenase-1, HO-1) ทั้งในระดับอาร์เอ็นเอและโปรตีน ซึ่งเอนไซม์ HO-1 นี้ทำหน้าที่เป็นสารต้านการออกซิเดชันซึ่งถูกจัดเป็นเอนไซม์ที่ช่วยต้านการอักเสบ จากการทดลองข้างต้นแสดงให้เห็นถึงลักษณะการศึกษาแบบจำเพาะกับเอนไซม์หรือปฏิกิริยาที่เกี่ยวข้องกับกลไกการอักเสบโดยตรง แต่ในปัจจุบันการศึกษาอย่างเป็นระบบแบบชีววิทยาเชิงระบบคือการทำความเข้าใจหรือช่วยการส่งสัญญาณที่เกี่ยวข้องกับการอักเสบทั้งระบบ ซึ่งจะทำให้เรามองเห็นภาพและเข้าใจของกลไกการอักเสบและกลไกการต่อต้านการอักเสบได้ครบถ้วนสมบูรณ์และครอบคลุมมากยิ่งขึ้นได้

2.2.2 การวิเคราะห์ไมโครอาร์เรย์ของการแสดงออกของยีน (Microarray gene expression analysis)

เทคโนโลยีดีเอ็นเอไมโครอาร์เรย์สามารถนำไปประยุกต์ใช้เพื่อตรวจสอบการแสดงออกของยีนจำนวนมากภายในเวลาเดียวกันและมีการนำไปใช้อย่างแพร่หลายในการศึกษาวิจัยทางชีวภาพและในสปีชีส์ของสิ่งมีชีวิตที่หลากหลาย (Brown and Botstein, 1999) การเพิ่มมากขึ้นของข้อมูลที่ได้จากการตรวจวัดการแสดงออกของยีนด้วยเทคโนโลยีไมโครอาร์เรย์ทำให้เกิดฐานข้อมูลสาธารณะที่บรรจุข้อมูลเหล่านี้อยู่มากมายซึ่งเราสามารถนำมาใช้งานได้ ภายในฐานข้อมูลเหล่านี้ประกอบไปด้วยข้อมูลไมโครอาร์เรย์ที่ได้จากการศึกษาภายใต้หลากหลายสภาวะจึงเป็นแหล่งข้อมูลที่มีคุณค่าสำหรับการนำไปวิเคราะห์เพื่อค้นหาความรู้ที่ซ่อนอยู่ภายใน (Gollub, 2003; and Barrett and Edgar, 2004) จุดเริ่มต้นของเทคโนโลยีดีเอ็นเอไมโครอาร์เรย์มักถูกนำมาใช้เพื่อการค้นคว้าวิจัยภายในห้องปฏิบัติการทางชีวภาพเพื่อทำความเข้าใจกระบวนการทางชีวภาพให้

ดียิ่งขึ้น จนกระทั่งภายหลังได้มีการพัฒนาต่อยอดเรื่องของการปรับแต่งเทคโนโลยีให้เหมาะสมเพื่อนำไปใช้ในปฏิบัติการเชิงคลินิกซึ่งเข้าใกล้กับการใช้ชีวิตจริงมากยิ่งขึ้น (Sandvik, 2006 และ Pusztai, 2004) ซึ่งการนำไปใช้จริงทางคลินิกเพื่อการตรวจวินิจฉัยโรคนั้นต้องการความถูกต้องแม่นยำเพื่อการรักษาที่ดี จากการศึกษาวิจัยในปัจจุบันพบว่าตรวจสอบการแสดงออกของยีนโดยเทคโนโลยีไมโครอาร์เรย์นั้นสามารถนำมาประยุกต์ใช้เพื่อช่วยเหลือในการวินิจฉัยโรคระดับคลินิกได้ เนื่องจากการแสดงออกของยีนจะมีแบบแผนที่สอดคล้องกับสภาวะที่สนใจศึกษา ซึ่งการเปลี่ยนแปลงที่เกิดขึ้นบางอย่างนั้นยังไม่สามารถตรวจพบได้หากใช้เพียงข้อมูลทางพยาธิสภาพ (Golub, 1999) การวิเคราะห์ข้อมูลไมโครอาร์เรย์ที่เกี่ยวข้องกับสภาวะที่สนใจจึงมีส่วนสำคัญที่ช่วยให้เราเข้าใจกลไกและกระบวนการเปลี่ยนแปลงทางชีวภาพระดับโมเลกุลได้ดียิ่งขึ้น และยังสามารถนำมาหาเครื่องหมายระดับโมเลกุลหรือเครื่องหมายยีนเพื่อช่วยในการจำแนกหรือวินิจฉัยโรคให้มีความถูกต้องแม่นยำมากขึ้นได้ (Basil, 2006, Yap, 2007, และ Zhang, 2007)

2.2.3 ข้อมูลการศึกษาไมโครอาร์เรย์ของการแสดงออกของยีนในเซลล์แมคโครฟาจหนู และการบูรณาการข้อมูลโดยอาศัยหลักการชีววิทยาเชิงระบบ

ปัจจุบันมีข้อมูลที่ศึกษาการตอบสนองที่เกี่ยวข้องกับการอักเสบโดยใช้เซลล์แมคโครฟาจของหนูเป็นต้นแบบอยู่เป็นจำนวนมาก ซึ่งข้อมูลไมโครอาร์เรย์ของชุดข้อมูลเหล่านี้ถูกบรรจุอยู่ในฐานข้อมูลสาธารณะ GEO (Edgar 2002, Gene Expression Omnibus, <http://www.ncbi.nlm.nih.gov/geo/>) ซึ่งเราสามารถเข้าไปดึงข้อมูลเหล่านี้มาใช้กันได้ ยกตัวอย่างเช่นข้อมูลไมโครอาร์เรย์ซีรีส์ GSE32359 ที่ใช้ไมโครอาร์เรย์แบบ Affymetrix Mouse Exon 1.0 ST Array สำหรับศึกษาการแสดงออกระดับอาร์เอ็นเอของเซลล์แมคโครฟาจจากไขกระดูกหนู ซึ่งประกอบไปด้วยตัวอย่างที่ได้รับการกระตุ้นด้วย LPS กับที่ไม่ถูกกระตุ้นอย่างละ 3 ตัวอย่าง (Gold, 2012) นอกจากนี้เรายังสามารถพบข้อมูลไมโครอาร์เรย์ที่เกี่ยวข้องกับการตอบสนองต่อ LPS ของเซลล์แมคโครฟาจหนูในชุดข้อมูลที่ปฏิบัติไมโครอาร์เรย์ภายใต้แพลตฟอร์มเดียวกันคือ Affymetrix Mouse Expression 430A Array ซึ่งสามารถตรวจพบได้ถึง 3 ชุดข้อมูลคือ GSE1459 (Murray, 2005) GSE2002 (Shell, 2005) และ GSE3565 (Hammer, 2006) ซึ่งชุดข้อมูลไมโครอาร์เรย์เหล่านี้เป็นเพียงบางตัวอย่างของชุดข้อมูลไมโครอาร์เรย์ที่สามารถตรวจพบได้จากแหล่งข้อมูล GEO (Edgar 2002) ซึ่งข้อมูลเหล่านี้จะเป็นข้อมูลที่สำคัญเพื่อนำมาใช้ในการวิเคราะห์ เพื่อทำการค้นหากลุ่มยีนที่มีความสำคัญและเกี่ยวข้องกับการตอบสนองต่อ LPS ของเซลล์แมคโครฟาจหนู เพื่อนำมาสร้างแบบแผนของการส่งผ่านสัญญาณที่เกี่ยวข้อง ซึ่งเราสามารถนำไปประยุกต์ใช้ได้ในการสร้างแพลตฟอร์มเชิงข้อมูลเพื่อนำมาช่วยเป็นต้นแบบในการตัดสินใจและพิจารณาถึงผลที่จะเกิดขึ้นจากการทดสอบสารออกฤทธิ์ที่มีต่อการอักเสบ และการวิเคราะห์โดยการรวบรวมข้อมูลการแสดงออกระดับอาร์เอ็นเอที่ได้จากข้อมูลไมโครอาร์เรย์ข้างต้นร่วมกับข้อมูลระดับโปรตีนที่ทางผู้ร่วมวิจัยของกลุ่มวิจัยได้ทำไว้ก็จะยิ่งทำให้เราสามารถทำความเข้าใจกลไกที่เกิดขึ้นได้ครบถ้วนและดีมากยิ่งขึ้น ตัวอย่างเช่นการวิเคราะห์ร่วมกันของข้อมูลระดับอาร์เอ็นเอและโปรตีนโดยใช้ชีววิทยาเชิงระบบเพื่อการศึกษาการตอบสนองของผู้ป่วยที่มีต่อการติดเชื้อไวรัสเอชไอวี (Tan, 2007) ดังรูปที่ 1 ซึ่งการบูรณาการความรู้ที่ได้จากข้อมูลในแต่ละระดับเข้าด้วยกันจะทำให้เราได้ผลลัพธ์ที่เป็นความรู้ใหม่เพิ่มมากขึ้นจากเดิม ซึ่งอาจจะแสดงถึงคุณสมบัติหรือคุณลักษณะใหม่ที่ไม่สามารถหาได้จากการวิเคราะห์ข้อมูลเพียงระดับเดียว (Joyce, 2006)

บทที่ 3

วัสดุ อุปกรณ์ และวิธีการทดลอง

3.1 วัสดุ และอุปกรณ์

3.1.1 เครื่องมือและโปรแกรมทางชีวสารสนเทศ

1. เครื่องคอมพิวเตอร์ส่วนบุคคล (Personal Computer)
 - a. Processor: Intel® Core™ 2Duo P7450 @2.13 GHz
Installed memory: (RAM) 4.00 GB
System type: 32-bit Operating System
 - b. Processor: Intel® Core™ i3-2365 CPU @1.40 GHz
Installed memory: (RAM) 4.00 GB
System type: 64-bit Operating System
 - c. Processor: Intel® Core™ i3-2350 CPU @2.30 GHz NVIDIA GeForce GT650M
Installed memory: (RAM) 8.00 GB DDR3
System type: 64-bit Operating System
 - d. Processor: Intel® Core™ i5-2450M CPU @2.40 GHz NVIDIA GeForce GT650M
Installed memory: (RAM) 8.00 GB DDR3
System type: 64-bit Operating System
 - e. Processor: Intel® Core™ i5-2450M CPU @2.50 GHz
Installed memory: (RAM) 6.00 GB
System type: 64-bit Operating System
 - f. Processor: Intel® Core™ i3 CPU M 330 @2.13 GHz
Installed memory: (RAM) 2.00 GB
System type: 32-bit Operating System
 - g. Processor: Intel® Core™ 2 Duo CPU T6400 @2.00 GHz
Installed memory: (RAM) 2.00 GB
System type: 32-bit Operating System
2. ระบบปฏิบัติการไมโครซอฟวินโดวส์เซเว่น (Windows 7)
3. โปรแกรมไมโครซอฟท์เอกซ์เซล (Microsoft Excel 2010, 2013)
4. โปรแกรม DAVID (The Database for Annotation, Visualization and Integrated Discovery, version 6.7; Huang, Sherman and Lempicki, 2009)
5. โปรแกรม Cytoscape version 2.6.3 (<http://www.cytoscape.org>; Cline et al., 2007)
6. โปรแกรม Arraymining (http://www.arraymining.net/R-php-1/ASAP/microarrayin_fobiotic.php; Glaab et al., 2009)
7. โปรแกรม String (<http://string-db.org/>; Szklarczyk et al., 2015)
8. โปรแกรม Cladist (Fukuchi et al., 2009)

9. โปรแกรม DREM 2.0 (Dynamic Regulatory Events Miner version 2.0; Schulz et al., 2012)
10. โปรแกรม MEV (Multi Experiment Viewer, version 4.8.1; Eleanor et al., 2011)
11. โปรแกรม WEKA (Waikato Environment for Knowledge Analysis, version 3.6.2; Hall et al., 2009)
12. โปรแกรม Polysearch (<http://wishart.biology.ualberta.ca/polysearch/>; Cheng et al., 2008)
13. โปรแกรม CyNetworkBMA (<http://apps.cytoscape.org/apps/cynetworkbma>; Yeung et al., 2011)
14. โปรแกรม JactiveModules (<http://apps.cytoscape.org/apps/JactiveModules>; Ideker et al., 2002)
15. โปรแกรม R version 3.1.1 (<https://www.r-project.org/>; R Core Team, 2013)
16. โปรแกรม BiNGO version 3.2.0 (<http://apps.cytoscape.org/apps/bingo>; Maere, Heymans and Kuiper, 2005)
17. ฐานข้อมูล GEO (<http://www.ncbi.nlm.nih.gov/geo/>; Barrett and Edgar, 2006)
18. ฐานข้อมูล KEGG (<http://www.genome.jp/kegg/>; Kanehisa, 2010)

3.1.2 วัสดุและอุปกรณ์ทางห้องปฏิบัติการ

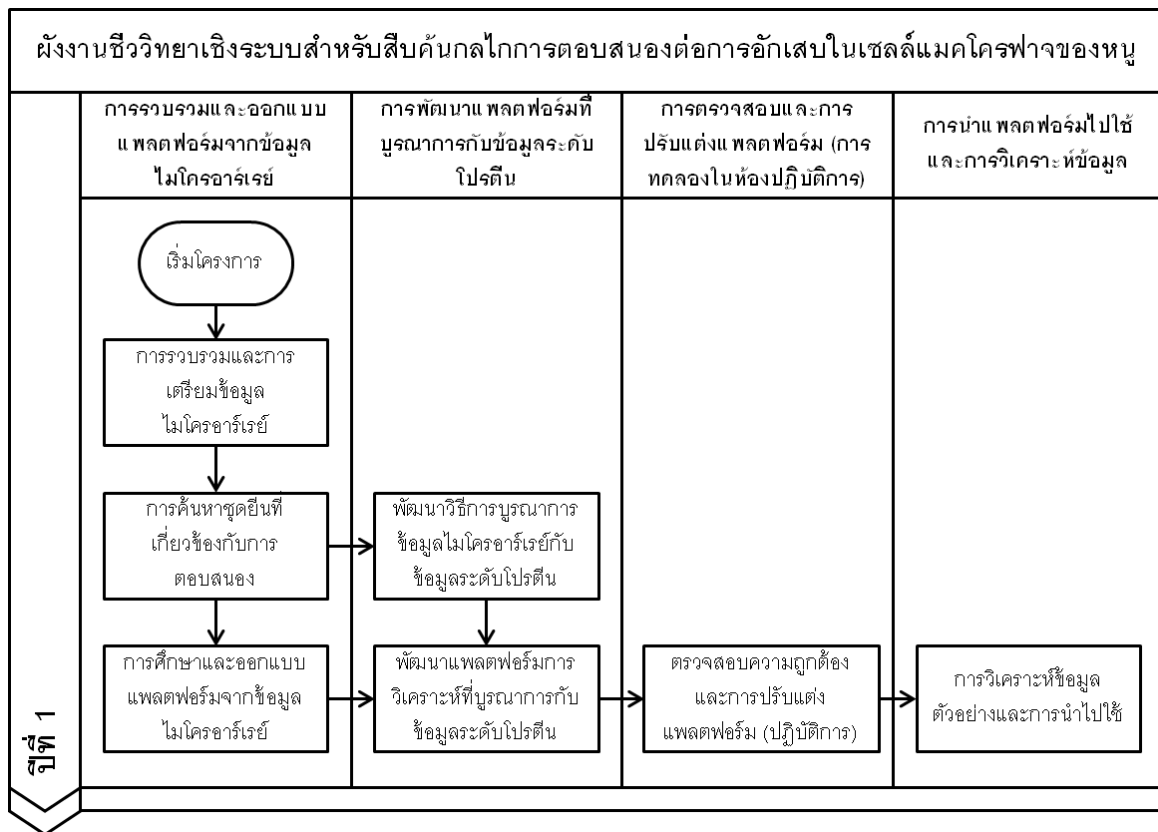
1. ตู้ปลอดเชื้อ (laminar flow class II) (NU-440, NUAIRE, ประเทศสหรัฐอเมริกา)
2. ตู้บ่มเซลล์แบบใช้คาร์บอนไดออกไซด์ (CO₂ incubator) (CB210, Binder, ประเทศสหพันธ์สาธารณรัฐเยอรมนี)
3. กล้องจุลทรรศน์แบบกลับหัว (inverted microscope) (OLYMPUS IX70, ประเทศญี่ปุ่น)
4. เครื่องปั่นเหวี่ยง (centrifuge) (K240R, Centurion Scientific, ประเทศสหราชอาณาจักร)
5. เครื่องชั่งตวงวัด 2 และ 4 ตำแหน่ง (Precsia, ประเทศสมาพันธรัฐสวิส)
6. เครื่องวัดค่าการดูดกลืนแสงแบบไมโครเพลท (VERSA MAX, ประเทศสหรัฐอเมริกา)
7. เครื่อง PCR (T-gradient, Biometra®, ประเทศสหรัฐอเมริกา)
8. เครื่อง Real-time PCR (iCycler, Bio-Rad, ประเทศสหรัฐอเมริกา)
9. เครื่อง UV/VIS Spectrophotometer (GVC CINTRA 40 , ประเทศแคนาดา)
10. จานเพาะเลี้ยงเซลล์แบบ 24 หลุม (Corning, ประเทศสหรัฐอเมริกา)
11. จานเพาะเลี้ยงเซลล์ขนาด 60 มม. (Corning, ประเทศสหรัฐอเมริกา)
12. จานเพาะเลี้ยงเซลล์ขนาด 100 มม. (Corning, ประเทศสหรัฐอเมริกา)
13. ชุดกรองปลอดเชื้อสำหรับติดกับขวดขนาดเส้นผ่านศูนย์กลางรูกรอง 0.22 ไมครอน (Corning, ประเทศสหรัฐอเมริกา)
14. ไมโครเพลทแบบ 96 หลุม (Sterilin, ประเทศสหราชอาณาจักร)

3.2 สารเคมี

1. Dimethyl sulfoxide, DMSO (Sigma-Aldrich, ประเทศสหรัฐอเมริกา)
2. Di-Sodium hydrogen phosphate , Na₂HPO₄ (Carlo Erba, ประเทศสหพันธ์สาธารณรัฐเยอรมนี)
3. D-Glucose (Sigma, ประเทศสหรัฐอเมริกา)
4. Dulbecco's Modified Eagle Medium (DMEM) ที่เติม phenol red (Gibco, ประเทศสหรัฐอเมริกา)
5. Fetal bovine serum (FBS) (Gibco, ประเทศสหรัฐอเมริกา)
6. Lipopolysaccharide (LPS) จาก Escherichia coli 0111: B4 (Sigma chemical, ประเทศสหรัฐอเมริกา)
7. Penicillin/Streptomycin (Gibco-Invitrogen, ประเทศสหรัฐอเมริกา)
8. Potassium chloride, KCl (Carlo Erba, ประเทศสหพันธ์สาธารณรัฐเยอรมนี)
9. Potassium dihydrogen orthophosphate, KH₂PO₄ (Analar, ประเทศสหราชอาณาจักร)
10. Sodium bicarbonate, NaHCO₃ (Sigma, ประเทศสหรัฐอเมริกา)
11. Sodium chloride (Carlo ERBA reagenti, ประเทศสาธารณรัฐฝรั่งเศส)
12. Thizolyl Blue Tetrazolium Bromide, approx. (MTT) (Sigma-Aldrich, ประเทศสหรัฐอเมริกา)

3.3 วิธีการทดลอง

การพัฒนาแพลตฟอร์มทางชีววิทยาระบบสำหรับการวิเคราะห์และสืบค้นกลไกการตอบสนองต่อการอักเสบในเซลล์แมโครฟาจของหนู ซึ่งเป็นเซลล์ที่นิยมใช้สำหรับเป็นโมเดลหรือต้นแบบเพื่อการศึกษาในมนุษย์ ภาพรวมของงานวิจัยนี้มีขั้นตอนในการพัฒนาดังแผนงานรูปที่ 2 ที่ประกอบไปด้วย การรวบรวมและออกแบบแพลตฟอร์มจากข้อมูลไมโครอาร์เรย์ การพัฒนาแพลตฟอร์มที่บูรณาการกับข้อมูลระดับโปรตีน การตรวจสอบความถูกต้องและการปรับแต่งแพลตฟอร์ม การตรวจสอบยืนยันผลด้วยวิธีเรียลไทม์พีซีอาร์ (real time PCR) และการนำแพลตฟอร์มไปใช้ จากภาพที่ 3-1 ลูกศรแสดงการทำงานจากขั้นตอนหนึ่งไปสู่อีกขั้นตอนต่อไป โดยมีรายละเอียดของการดำเนินการวิจัยในแต่ละขั้นตอนดังต่อไปนี้



ภาพที่ 3-1 ผังงานของโครงการชีววิทยาเชิงระบบสำหรับสืบค้นกลไกการตอบสนองต่อการอักเสบในเซลล์แมคโครฟาจของหนู

3.3.1 การรวบรวมและการเตรียมข้อมูลไมโครอาร์เรย์

การทดลองนี้จะใช้ข้อมูลการแสดงออกของยีนทั้งหมดในจีโนมที่ได้จากการใช้ไมโครอาร์เรย์ภายใต้สภาวะการศึกษาที่เกี่ยวข้องกับการตรวจสอบผลการตอบสนองต่อ LPS ในเซลล์แมคโครฟาจของหนู จากหลากหลายกลุ่มวิจัยที่สามารถดาวน์โหลดชุดข้อมูลไมโครอาร์เรย์เหล่านี้ได้จากฐานข้อมูลสาธารณะ GEO (Edgar 2002, <http://www.ncbi.nlm.nih.gov/geo/>) โดยข้อมูลหลักที่ใช้ในการทดลองนี้มีทั้งหมด 3 ชุดข้อมูล ได้แก่ GSE4712 (Comer et al., 2006) GSE21841 (Hammer et al., 2010) และ GSE2002 (Shell et al., 2005) โดยชุดข้อมูลไมโครอาร์เรย์ที่ได้นำมาวิเคราะห์นั้น จะคัดเลือกมา 2 กลุ่มการทดลองคือ กลุ่มที่ถูกกระตุ้นด้วย LPS และกลุ่มที่ไม่ถูกกระตุ้นด้วย LPS ตามช่วงเวลา ในเซลล์แมคโครฟาจเซลล์ไลน์ RAW264.7 ของหนู ดังที่ได้แสดงรายละเอียดของชุดข้อมูลในตารางที่ 3-1 ก่อนที่จะนำข้อมูลไมโครอาร์เรย์ที่สามารถดาวน์โหลดได้ในรูปแบบของข้อมูลดิบ ข้อมูลภาพ หรือข้อมูลการแสดงออก ของแต่ละชุดข้อมูลมาผ่านขั้นตอนการเตรียมข้อมูล และปรับเทียบมาตรฐานข้อมูลของแต่ละชุดข้อมูลให้อยู่ในมาตรฐานเดียวกันเพื่อนำข้อมูลไมโครอาร์เรย์ทั้งหมดที่ได้ไปใช้ในการวิเคราะห์ภายใต้วิธีการเดียวกันในขั้นตอนต่อไป เพื่อค้นหากลุ่มยีนที่มีความสำคัญและเกี่ยวข้องกับสภาวะที่ศึกษาทั้งภายในชุดข้อมูล และระหว่างชุดข้อมูลไมโครอาร์เรย์

ตารางที่ 3-1 แสดงรายละเอียดของชุดข้อมูลไมโครอาร์เรย์

GSE ID	Title	References	Platforms	Samples	Sources of LPS
GSE4712	Edema Toxin-treated Macrophage Study.	Comer <i>et al.</i> , 2006	Affymetrix Mouse Genome 430 2.0 Array (GPL1261)	9 samples untreated=3 3hr. LPS=3 6hr. LPS=3	LPS (0.01ng/ml; Sigma Aldrich)
GSE21841	Expression data from LPS-stimulated RAW 264.7 mouse macrophages treated with Hypericum perforatum fraction and bioactive constituents	Hammer <i>et al.</i> , 2010	Affymetrix Mouse Genome 430 2.0 Array (GPL1261)	6 samples untreated=3 8hr. LPS=3	LPS (1ug/ml; Sigma Aldrich)
GSE2002	CstF-64 Influenced Gene Expression	Shell <i>et al.</i> , 2005	Affymetrix Mouse Expression 430A Array (GPL339)	6 samples Untreated = 3 18 hr_LPS = 3	LPS (100 ng/ml; Sigma Aldrich)

3.3.2 การค้นหาชุดยีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบ

นำข้อมูลที่ได้มาทำการปรับมาตรฐานค่าการแสดงออกของยีนในแต่ละชุดข้อมูลให้อยู่ในรูปแบบเดียวกันนั้นโดยใช้ค่ามาตรฐาน z score ซึ่งสามารถคำนวณค่า z score ได้จากสมการที่ 1 เพื่อนำข้อมูลการแสดงออกของยีนจากชุดข้อมูลไมโครอาร์เรย์ที่คัดเลือกไปวิเคราะห์ในขั้นตอนต่อไป

$$z \text{ score} = ((x - \text{mean}) / \text{SD}) \quad (1)$$

โดยที่ x คือ ค่าการแสดงออกของของโพรบแต่ละตัวอย่างในชุดข้อมูล

Mean คือ ค่าเฉลี่ยของค่าการแสดงออกของโพรบแต่ละโพรบ

SD คือ ค่าเบี่ยงเบนมาตรฐานค่าการแสดงออกของโพรบแต่ละโพรบ

ภายหลังการปรับเทียบมาตรฐานข้อมูลไมโครอาร์เรย์ของทุกชุดข้อมูลเรียบร้อยแล้ว ขั้นตอนต่อมาคือการนำชุดข้อมูลไมโครอาร์เรย์ที่ได้มาตรวจสอบและวิเคราะห์ด้วยโปรแกรมเฉพาะทางที่เกี่ยวข้อง เพื่อทำการคัดเลือกชุดยีนที่มีความสำคัญและเกี่ยวข้องกับการตอบสนองต่อการอักเสบ หรือเป็นการตอบสนองต่อ LPS ในเซลล์แมคโครฟาจของหนูตามช่วงเวลา โดยวิธีการคัดเลือกทางสถิติที่บรรจุอยู่ภายในโปรแกรมการทำเหมืองข้อมูลออนไลน์ ArrayMining (Glaab *et al.*, 2009) ซึ่งในงานวิจัยนี้ได้เลือกใช้ 3 วิธี ได้แก่ Significance analysis of microarrays (SAM) (Tusher, Tibshirani and Chu, 2001), Empirical Bayes moderated

t-test (eBayes) (Smyth, 2004), และ PLS-CV-Partial Least Squares Cross-Validation (PLS-CV) (Hall, 2000) โดยคัดเลือกยีนที่มีการแสดงออกอย่างเด่นชัดใน 100 โพรบแรก จากนั้นทำการคัดเลือกเฉพาะโพรบที่ถูกคัดเลือกโดย 2 วิธี จากทั้งหมด 3 วิธี หลังจากนั้นทำการตรวจสอบชุดยีนที่คัดเลือกได้โดยพิจารณาค่าความถูกต้องของการจำแนก (classification accuracy) ด้วยโปรแกรม WEKA (Hall et al., 2009) โดยทำการตรวจสอบทั้งภายในชุดข้อมูล และระหว่างชุดข้อมูลของไมโครอาร์เรย์ที่นำมาศึกษาเพื่อให้ได้ชุดยีนที่มีความน่าเชื่อถือและสัมพันธ์กับลักษณะการตอบสนองต่อการอักเสบมากที่สุด (Rotter, 2008)

3.3.3 การศึกษาและออกแบบแพลตฟอร์มจากข้อมูลไมโครอาร์เรย์

ในการศึกษาและออกแบบแพลตฟอร์มจากข้อมูลไมโครอาร์เรย์ที่ผ่านการเตรียมข้อมูลมาแล้ว ในขั้นแรกของการศึกษาจะเป็นการทดลองการสร้างเครือข่ายการแสดงออกร่วมและเครือข่ายปฏิสัมพันธ์ระหว่างยีนของข้อมูลไมโครอาร์เรย์โดยเลือกใช้โปรแกรมหลากหลายชนิดเพื่อค้นหาโปรแกรมที่เหมาะสมกับการใช้งานมากที่สุด โปรแกรมที่ได้ทำการทดลองใช้เพื่อสร้างเครือข่ายประกอบไปด้วย โปรแกรม Cladist (Fukuchi et al., 2009) โปรแกรม Network Analysis ที่อยู่ในโปรแกรมออนไลน์ Arraymining (Glaab et al., 2009) ที่ จะทำการสร้างเครือข่ายการแสดงออกร่วมระหว่างยีน โปรแกรม CyNetworkBMA (Yeung et al., 2011) ซึ่งเป็นปลั๊กอินของโปรแกรม Cytoscape (Cline et al., 2007) ที่จะสร้างเครือข่ายปฏิสัมพันธ์จากการคำนวณความสัมพันธ์ของการแสดงออกของยีน และโปรแกรม DREM 2.0 (Dynamic Regulatory Events Miner version 2.0; Schulz et al., 2012) ซึ่งสร้างเครือข่ายการควบคุมการแสดงออกของยีน โดยอาศัยการวิเคราะห์และแสดงผลภายในโปรแกรม Cytoscape (Cline et al., 2007) โดยใช้โปรแกรม MEV (Multi Experiment Viewer, version 4.8.1; Eleanor et al., 2011) สำหรับการวิเคราะห์ข้อมูลไมโครอาร์เรย์เพื่อการแสดงภาพการจัดกลุ่มยีนที่สำคัญ

นอกจากนั้นยังทำการคัดเลือกกลุ่มยีนที่เกี่ยวข้องกับการอักเสบ (inflammatory genes) จากฐานข้อมูล PolySearch (Cheng, 2008) ซึ่งเป็นฐานข้อมูลที่มีการค้นหาข้อมูลจากฐานข้อมูลเอกสาร (text mining) ในการค้นหายีนที่เกี่ยวข้องกับการอักเสบโดยใช้คำสำคัญในการค้นหาว่า “inflammation” จากนั้นนำรายชื่อยีนที่สืบค้นได้มาวิเคราะห์ร่วมกับยีนที่ถูกคัดเลือกโดยวิธีการทำเหมืองข้อมูลที่ได้ทำคัดเลือกก่อนหน้านี้ และนำยีนที่ถูกคัดเลือกทั้งสองวิธีมาสร้างเครือข่ายความสัมพันธ์ระหว่างการแสดงออกร่วมของยีน (co-expression network) โดยใช้ปลั๊กอิน Expression Correlation Network ที่ถูกบรรจุภายใต้โปรแกรม Cytoscape (Cline et al., 2007) จากนั้นทำการคัดเลือกเครือข่ายย่อยที่น่าสนใจโดยใช้ยีนที่ถูกคัดเลือกได้จากฐานข้อมูล PolySearch เป็นหลัก และนำเครือข่ายย่อยที่คัดเลือกได้ไปวิเคราะห์ในขั้นตอนต่อไป เพื่อให้ได้แพลตฟอร์มของรูปแบบการส่งผ่านสัญญาณที่เกี่ยวข้องกับรูปแบบการแสดงออกระดับอาร์เอ็นเอของยีนสำคัญแต่ละชนิด ซึ่งจะเป็นข้อมูลเชิงโครงสร้างหรือแพลตฟอร์มที่สำคัญสำหรับนำไปใช้ในขั้นตอนต่อไป

3.3.4 การบูรณาการข้อมูลแพลตฟอร์มที่ได้จากข้อมูลไมโครอาร์เรย์

ในการบูรณาการข้อมูลซึ่งเป็นข้อมูลการแสดงออกคนละระดับกัน เราจะต้องใช้ความรู้พื้นฐานที่มีอยู่มาใช้เป็นข้อมูลอ้างอิงเพื่อช่วยในการบูรณาการ เช่น การใช้ข้อมูลของวิถีเมตาบอลิซึม (metabolic pathway) ที่สามารถหาได้จากฐานข้อมูลสาธารณะ เช่น ข้อมูลเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีน (protein-protein interaction network) ในฐานข้อมูลสาธารณะฐานข้อมูล STRING (Szklarczyk et al., 2015) และข้อมูลปฏิสัมพันธ์ของการควบคุมระหว่างทรานสคริปชันแฟกเตอร์กับยีนเป้าหมาย (TF-TG interaction) ซึ่งในที่นี้จะใช้ข้อมูลของหนูเมาส์ (Mouse; *Mus musculus*) (Ernst et al., 2010) เป็นต้น โดยอาศัยการบูรณาการ

ข้อมูลภายในโปรแกรม Cytoscape (Cline et al., 2007) และการใช้ปลั๊กอินภายในโปรแกรม Cytoscape ได้แก่ JactiveModules (Ideker et al., 2002) เพื่อช่วยในการคัดเลือกโมดูลย่อยของเครือข่ายที่แสดงออกถึงความสัมพันธ์อย่างเด่นชัด โปรแกรม BiNGO (Maere, Heymans and Kuiper, 2005) ที่ช่วยในการแปลผลทางชีวภาพโดยอาศัยข้อมูลยีนออนโทโลยี (GO; Gene Ontology) ของกลุ่มยีนที่ทำการศึกษา เพื่อนำมาช่วยในการตัดสินใจคัดเลือกกลุ่มยีนที่น่าจะมีบทบาทและความสำคัญในการตอบสนองต่อการอักเสบ โดยอาศัยการตรวจสอบการแปลผลทางชีวภาพเพิ่มเติมจากโปรแกรมออนไลน์ DAVID (Huang, Sherman and Lempicki, 2009) ที่มีฐานข้อมูลของการแปลผลทางชีวภาพหลากหลายชนิด โดยใช้ข้อมูลการแปลผลจากฐานข้อมูลของยีนออนโทโลยี (Gene Ontology; GO) (Ashburner et al., 2000) และ KEGG (Kyoto Encyclopedia of Genes and Genomes) (Kanehisa, 2010)

3.3.5 การตรวจสอบและยืนยันผลยีนที่ถูกคัดเลือกที่มีการตอบสนองต่อการอักเสบ เมื่อมีการกระตุ้นด้วย LPS ตามช่วงเวลา ในระดับอาร์เอ็นเอ

จากการคัดเลือกยีนเป้าหมายที่มีความสำคัญต่อการตอบสนองการอักเสบ เมื่อมีการกระตุ้นด้วย LPS ตามช่วงเวลาที่ได้จากการวิเคราะห์ข้อมูลไมโครอาร์เรย์ ร่วมกับการวิเคราะห์เครือข่ายการตอบสนองต่อการอักเสบ การบูรณาการข้อมูลแพลตฟอร์ม และการยืนยันผลการคัดเลือกกลุ่มยีนกับเอกสารงานวิจัยที่เกี่ยวข้อง โดยจะทำการตรวจสอบการแสดงออกของยีนที่ถูกคัดเลือกมากับผลที่ได้จากการวิเคราะห์ข้อมูลไมโครอาร์เรย์ โดยเริ่มจากการเพาะเลี้ยงเซลล์แมคโครฟาจสายพันธุ์ RAW 264.7 ภายใต้สภาวะที่ไม่มีการกระตุ้นด้วย LPS และมีการกระตุ้นด้วย LPS ที่เวลา 3, 6, 8, 18 และ 24 ชั่วโมง เพื่อตรวจวัดการตอบสนองที่เกิดขึ้นโดยการตรวจสอบการแสดงออกของยีนในระดับอาร์เอ็นเอโดยใช้วิธี real time PCR ของยีนที่ถูกคัดเลือกมาจำนวน 1 ยีน คือ SOCS3 โดยตรวจวัดร่วมกับยีน iNOS ซึ่งเป็นยีนที่ใช้อ้างอิงของการตอบสนองการอักเสบ และยีนที่มีการแสดงออกคงที่ EF-2 ซึ่งเป็นยีนอ้างอิง (housekeeping gene) เพื่อเป็นการยืนยันผลภายใต้การทดลองในห้องปฏิบัติการ โดยมีรายละเอียดของวิธีการ ดังนี้

3.3.5.1 การเพาะเลี้ยงเซลล์ (Cell culture) (กล่าวขวัญ ศรีสุข, 2555)

ทำการเพาะเลี้ยงเซลล์แมคโครฟาจของหนูสายพันธุ์ RAW 264.7 ในอาหารเลี้ยงเซลล์ชนิด DMEM ที่มีสารละลายยาเพนิซิลลิน(10,000 ยูนิต/มิลลิลิตร) สเตรปโตไมซิน (10,000 ไมโครกรัม/มิลลิลิตร) และ fetal bovine serum (FBS) (10%) บ่มในตู้เพาะเลี้ยงที่มีคาร์บอนไดออกไซด์ในอากาศ 5 % อุณหภูมิ 37 องศาเซลเซียส จนเซลล์มีการเจริญเติบโตเกือบเต็มพื้นผิวภาชนะเพาะเลี้ยง แล้วทำการขูดเก็บเซลล์

3.3.5.2 การ sub culture

ดูดอาหารเลี้ยงเซลล์ในจานเพาะเซลล์ทิ้ง เติมสารละลายบัฟเฟอร์ HBSS ที่ปราศจากแคลเซียมและแมกนีเซียมที่เย็นจำนวน 10 มิลลิลิตร แล้วนำไปบ่มในตู้เลี้ยงเซลล์ที่อุณหภูมิ 37 องศาเซลเซียส เป็นเวลา 10 นาที จากนั้นนำจานเพาะเซลล์ออกจากตู้และนำไปเข้าตู้ปลอดเชื้อเพื่อขูดเซลล์ให้หลุดออกจากจานเพาะเซลล์ โดยขูดเซลล์เบาๆใช้ปิเปตพลาสติกขนาด 10 มิลลิลิตร ดูดสารละลายบัฟเฟอร์ขึ้นลงหลายๆครั้ง แล้วใช้ปิเปตดูดสารแขวนลอยเซลล์ใส่หลอดพลาสติกขนาด 50 มิลลิลิตร และนำไปปั่นเหวี่ยงที่ 3000 g เป็นเวลา 3 นาที ดูดสารละลายบัฟเฟอร์ HBSS ทิ้ง แล้วเติมอาหารเลี้ยงเซลล์ที่เสริมด้วย 10% FBS ลงในหลอดพลาสติกที่ได้ปั่นเหวี่ยงแล้วทำการกระจายเซลล์โดยใช้ปิเปตขึ้นลงประมาณ 20 ครั้ง แล้วทำการนับจำนวนเซลล์ที่มีชีวิตและกระจายเซลล์ลงในภาชนะเลี้ยงเซลล์อันใหม่ตามอัตราส่วนที่ต้องการ

3.3.5.3 การสกัดอาร์เอ็นเอ

นำเซลล์แมคโครฟาจ RAW264.7 มาเพาะเลี้ยงในงานเพาะเลี้ยงเซลล์ขนาด 60 มิลลิเมตรในอาหารเลี้ยงเซลล์ DMEM ที่เสริมด้วย 10% FBS โดยให้มีปริมาณเซลล์เท่ากับ 1×10^6 เซลล์ต่อจานเป็นระยะเวลา 18-24 ชั่วโมงในตู้เพาะเลี้ยงที่มีคาร์บอนไดออกไซด์ในอากาศ 5 % อุณหภูมิ 37 องศาเซลเซียส งดอาหารเลี้ยงเซลล์ออกจนหมดและเติมอาหารเลี้ยงเซลล์ในกลุ่มควบคุมและเติมอาหารที่มี LPS ความเข้มข้น $1 \mu\text{g/ml}$ แล้วนำไปบ่มในตู้เพาะเลี้ยงเป็นเวลา 3, 6, 8, 18 และ 24 ชั่วโมง จากนั้นทำการเก็บเซลล์และสกัดอาร์เอ็นเอโดยใช้ชุดสกัดสำเร็จรูป illustra RNAspin Mini RNA Isolation Kit โดยทำตามคู่มือการสกัดจากบริษัทโดยมีขั้นตอนดังนี้

1. งดอาหารเลี้ยงเซลล์ออกให้หมด เติม lysis buffer RA1 350 μl และ β -mercaptoethanol 3.5 μl ตูดสารละลายขึ้นเพื่อเก็บตะกอนเซลล์ลงในหลอดขนาด 1.5 ml แล้วราดในโตรเจนเหลวเพื่อยับยั้งการทำงานของ RNase และ DNase โดยสามารถเก็บตัวอย่างไว้ที่อุณหภูมิ -80 องศาเซลเซียส
2. นำสารละลายตะกอนเซลล์ที่เก็บไว้มาละลาย และ vortex แรงๆ แล้วนำสารละลายไปไหลลงในคอลัมน์ที่มีวงแหวนสีม่วง นำไปปั่นเหวี่ยงด้วยความเร็ว $11000 \times g$ ที่อุณหภูมิ 4 องศาเซลเซียส เป็นเวลา 1 นาที
3. เติม 70% ethanol จำนวน 350 μl โดยดูดขึ้นลงประมาณ 5 ครั้ง เพื่อให้สารละลายเป็นเนื้อเดียวกัน
4. เตรียมคอลัมน์ที่มีวงแหวนสีฟ้าใส่ในหลอดขนาด 2 ml แล้วนำสารละลายในข้อ 3 มาไหลลงในคอลัมน์จำนวน 750 μl นำไปปั่นเหวี่ยงด้วยความเร็ว $8000 \times g$ ที่อุณหภูมิ 4 องศาเซลเซียส เป็นเวลา 30 วินาที และนำคอลัมน์ที่มีวงแหวนสีฟ้าใส่ในหลอดขนาด 2 ml (หลอดใหม่)
5. เติม MDB (Membrane Desalting Buffer) 350 μl ลงในคอลัมน์ที่มีวงแหวนสีฟ้าเพื่อล้างเกลือออก นำไปปั่นเหวี่ยงด้วยความเร็ว $10000 \times g$ ที่อุณหภูมิ 4 องศาเซลเซียส เป็นเวลา 1 นาที นำคอลัมน์ที่มีวงแหวนสีฟ้าใส่หลอดขนาด 2 ml (หลอดใหม่)
6. เติม DNase reaction mixture 95 μl (reconstituted DNase I 10 μl + DNase reaction buffer 90 μl) ลงบริเวณตรงกลางของคอลัมน์ แล้วตั้งทิ้งไว้ที่อุณหภูมิห้องเป็นเวลา 15 นาที
7. เติม buffer RA2 จำนวน 200 μl นำไปปั่นเหวี่ยงด้วยความเร็ว $11000 \times g$ ที่อุณหภูมิ 4 องศาเซลเซียส เป็นเวลา 1 นาที
8. เติม buffer RA3 จำนวน 600 μl นำไปปั่นเหวี่ยงด้วยความเร็ว $11000 \times g$ ที่อุณหภูมิ 4 องศาเซลเซียส เป็นเวลา 1 นาที
9. เติม buffer RA3 จำนวน 250 μl นำไปปั่นเหวี่ยงด้วยความเร็ว $11000 \times g$ ที่อุณหภูมิ 4 องศาเซลเซียส เป็นเวลา 2 นาที แล้วนำคอลัมน์ใส่ลงในหลอดขนาด 1.5 ml (หลอดใหม่)
10. เติม RNase-free H₂O จำนวน 50 μl (ครึ่งละ 25 μl) นำไปปั่นเหวี่ยงด้วยความเร็ว $11000 \times g$ ที่อุณหภูมิ 4 องศาเซลเซียส เป็นเวลา 1 นาที และเก็บสารละลาย RNA

หลังจากที่ทำการสกัดอาร์เอ็นเอเสร็จแล้วจะนำไปวิเคราะห์ความเข้มข้นและความบริสุทธิ์ของอาร์เอ็นเอ โดยนำ RNA จำนวน 5 μl เจือจางด้วยน้ำกลั่นที่ผสม 0.1% (v/v) DEPC จำนวน 495 μl และนำไปวัดค่าการดูดกลืนแสงที่ 260 nm และ 280 nm โดยใช้เครื่อง Spectrophotometer และคำนวณความบริสุทธิ์ของอาร์เอ็นเอโดยค่าควรอยู่ในช่วง 1.8-2.0 จะแสดงให้เห็นว่าอาร์เอ็นเอที่สกัดได้มีความบริสุทธิ์ เพื่อที่จะนำไปสังเคราะห์ cDNA ต่อไป

3.3.5.4 การสังเคราะห์ Complementary DNA (DNA)

นำอาร์เอ็นเอที่สกัดได้ไปสังเคราะห์ cDNA โดยใช้ 5X iScript™ Reverse Transcription Supermix จำนวน 4 µl RNA 2 µg และเติมน้ำที่ปราศจาก nuclease ให้ปริมาตรสุดท้ายเท่ากับ 20 µl ลงในหลอด PCR สภาวะที่ใช้ในการสังเคราะห์ cDNA ดังนี้

Priming	25 องศาเซลเซียส นาน 5 นาที
Reverse transcription	42 องศาเซลเซียส นาน 30 นาที
RT inactivation	85 องศาเซลเซียส นาน 5 นาที

จากนั้นนำ cDNA ที่ได้มาตรวจสอบคุณภาพและปริมาณโดยใช้ agarose gel electrophoresis และนำ cDNA ที่สังเคราะห์ได้ไปวิเคราะห์โดยใช้เทคนิค Real-time RT-PCR ต่อไป

3.3.5.5 การออกแบบไพรเมอร์ที่จำเพาะต่อยีน

ค้นหาลำดับเบสของยีน iNOS (NM_010927.3), EF-2 (NM_007907.2), SOCS3 (NM_007707.3) และ IL-11 (NM_008350.4) จากฐานข้อมูล NCBI (<http://www.ncbi.nlm.nih.gov/>) จากนั้นนำลำดับเบสที่ได้ไปออกแบบไพรเมอร์และตรวจสอบความจำเพาะต่อยีน โดยใช้โปรแกรม Primer blast (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) ในฐานข้อมูล NCBI

ตารางที่ 3-2 ลำดับนิวคลีโอไทด์ของไพรเมอร์ที่ใช้ในการทดลอง

Gene	Sequence	Product Length (bp)
iNOS	5' -GCACAGCACAGGAAATGTTTCAGCAC-3' (F)	156
	5' -AGCCAGCGTACCGGATGAGC- 3' (R)	
EF-2	5' -CTGAAGCGGCTGGCTAAGTCTGA- 3' (F)	155
	5' -GGGTCAGATTTCTTGATGGGGATG- 3' (R)	
SOCS3	5' -GCACAGCCTTTCAGTGCAGAGTA- 3' (F)	169
	5' -GAGACAGCGGTCGTAAGAGCAG- 3' (R)	

3.3.5.6 การวิเคราะห์การแสดงออกของยีนในระดับ mRNA โดยใช้ real time reverse transcription-polymerase chain reaction (Real-time RT-PCR)

นำ cDNA ที่สังเคราะห์เป็นต้นแบบ ทำปฏิกิริยาด้วย 2X iTag™ Universal SYBR® Green supermix และไพรเมอร์ที่จำเพาะของแต่ละยีนตามปริมาณที่ใช้ดังตารางที่ 4-3 และเมื่อผสมกันเสร็จจะนำไปทำปฏิกิริยาในเครื่อง real time PCR โดยสภาวะที่ใช้ในการทำปฏิกิริยาของยีน iNOS, EF-2, SOCS3 และ IL-11 ดังตารางที่ 4-4 จากนั้นทำการวิเคราะห์ melting curve ที่อุณหภูมิ 95 องศาเซลเซียส เป็นเวลา 10 นาที

แล้วเพิ่มอุณหภูมิจาก 65 องศาเซลเซียส ถึง 95 องศาเซลเซียสโดยเพิ่มทีละ 0.5 องศาเซลเซียส เป็นเวลา 5 วินาที และสามารถคำนวณหาปริมาณดีเอ็นเอได้ตามสูตร

$$\Delta\Delta Cq = \Delta Cq (\text{sample}) - \Delta Cq (\text{calibrator})$$

และ $\Delta Cq = Cq \text{ of target gene} - Cq \text{ of housekeeping gene}$

จำนวน target gene expression ที่ได้ normalized ด้วย reference gene และเทียบกับ calibrator จะเท่ากับ $2^{-\Delta\Delta Cq}$

ตารางที่ 3-3 สารและปริมาตรที่ใช้ในการทำปฏิกิริยา real-time PCR

สาร	ปริมาตรที่ใช้ (µl)
iTaq™ Universal SYBR® Green supermix	10
Forward primer	0.5
Reward primer	0.5
distilled water	7
cDNA	2
Total reaction mix volume	20

ตารางที่ 3-4 สภาวะที่ใช้ในปฏิกิริยา real-time PCR

ขั้นตอน	อุณหภูมิ (องศาเซลเซียส)	เวลา (นาที)
Step 1 Pre denature (1 รอบ)	95.0	3
Step 1 Denature	95.0	0.10
Step 2. annealing/extension		
	63.0	0.20
Step 1 (1 รอบ)	95.0	0.10
Melting curve	65.0-95.0	0.05

(อุณหภูมิเพิ่มขึ้นทีละ 0.5 องศาเซลเซียส)

3.3.5.7 การวิเคราะห์ข้อมูลทางสถิติ

ข้อมูลจากผลการทดลองทุกตัวอย่างในกลุ่มควบคุมและกลุ่มทดสอบจะทำซ้ำทั้งหมด 3 ซ้ำ และจะแสดงข้อมูลในรูปของค่าเฉลี่ย \pm ค่าเบี่ยงเบนมาตรฐาน (mean \pm SD) โดยใช้การวิเคราะห์ความแปรปรวน (analysis of variance) เป็นการทดสอบค่าเฉลี่ยจากข้อมูลที่ได้จากกลุ่มตัวอย่างหลายกลุ่ม และการเปรียบเทียบค่าความแตกต่างอย่างมีนัยสำคัญทางสถิติระหว่างกลุ่มควบคุมและกลุ่มทดสอบอย่างมีนัยสำคัญทางสถิติ โดยใช้ Student's t-test ค่า $p < 0.05$ ถือว่ามีนัยสำคัญทางสถิติ

บทที่ 4

ผลการทดลอง

4.1 การค้นหากระบวนการทางชีวสารสนเทศที่เหมาะสม

การศึกษาชีววิทยาเชิงระบบสำหรับสืบค้นกลไกการตอบสนองต่อการอักเสบในเซลล์แมคโครฟาจ RAW 264.7 ของหนู เป็นการอาศัยกระบวนการทางชีวสารสนเทศในการค้นหาสัญญาณและความสัมพันธ์ของกลุ่มที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบ ซึ่งสามารถนำเอารูปแบบของความสัมพันธ์ที่ได้มาใช้ในการอธิบายกลไกการตอบสนองต่อการอักเสบที่เกิดขึ้น เพื่อนำมาใช้เป็นข้อมูลที่ช่วยในการตัดสินใจสำหรับการคัดเลือกกลุ่มยีนที่มีบทบาทที่สำคัญต่อการตอบสนองต่อการอักเสบเพื่อศึกษาในระดับห้องปฏิบัติการต่อไป โดยลักษณะของแพลตฟอร์มหรือโมเดลของกลไกการอักเสบที่ได้ในที่นี่ จะอยู่ในรูปแบบของเครือข่ายปฏิสัมพันธ์ระหว่างกลุ่มยีนหรือโปรตีนที่ใช้ข้อมูลไมโครอาร์เรย์ของการแสดงของยีน ในสภาวะที่มีการตอบสนองต่อการอักเสบที่เกิดขึ้นโดยการกระตุ้นเซลล์แมคโครฟาจ RAW264.7 ด้วย LPS มาทำการวิเคราะห์ด้วยโปรแกรมทางชีวสารสนเทศซึ่งประกอบไปด้วย การคัดเลือกยีน (Gene selection) (ภาคผนวก ฉ, ช, จ) โดยโปรแกรม SVM (Support Vector Machines) ที่เป็นโปรแกรมย่อยภายในโปรแกรม WEKA (Hall et al., 2009) วิธีการเปรียบเทียบโพลด์เซนจ์ที่ระดับมากกว่า 1.4 ในระดับความเชื่อมั่นที่ $P < 0.05$ และการใช้เครื่องมือคัดเลือกยีน 3 ชนิด ได้แก่ Significance analysis of microarrays (SAM) (Tusher, Tibshirani and Chu, 2001), Empirical Bayes moderated t-test (eBayes) (Smyth, 2004), และ PLS-CV-Partial Least Squares Cross-Validation (PLS-CV) (Hall, 2000) ที่ถูกบรรจุอยู่ในโปรแกรมออนไลน์ ArrayMining (Glaab et al., 2009)

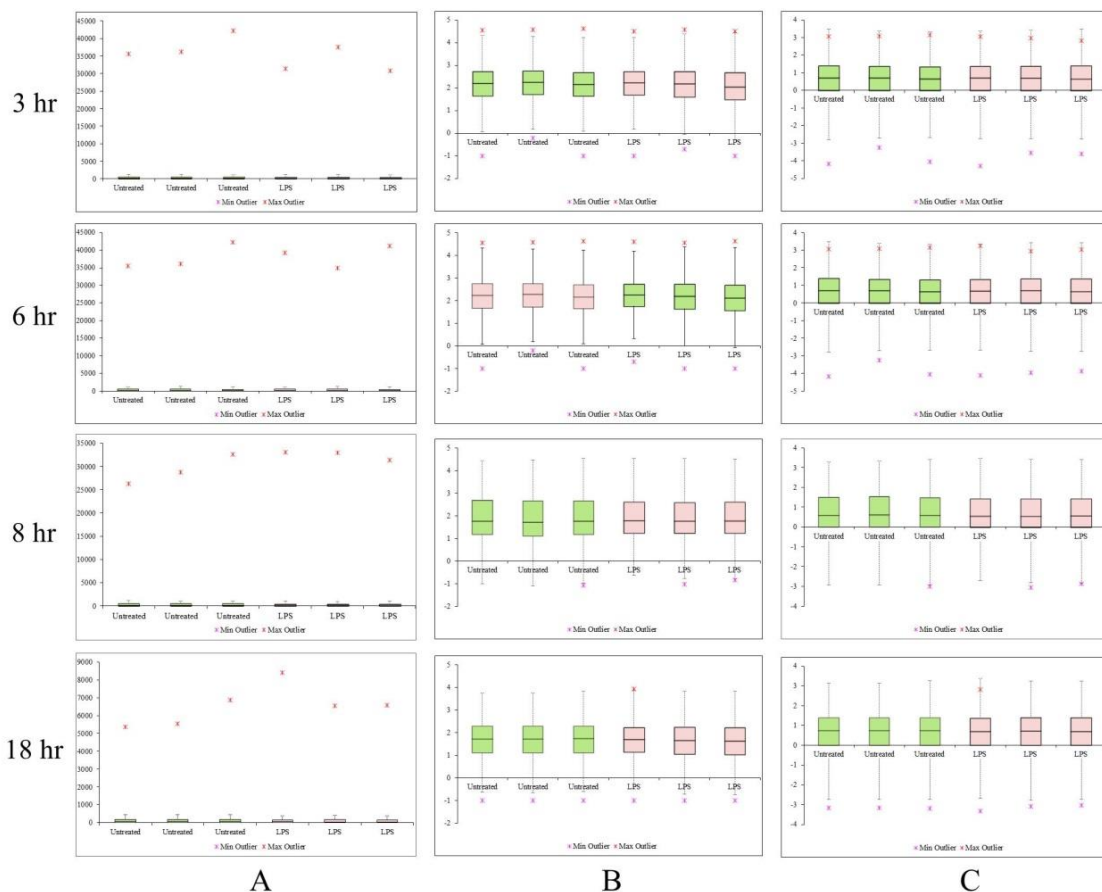
นอกจากนี้ยังอาศัยการสร้างเครือข่ายความสัมพันธ์จากระดับการแสดงออกของยีน (Co-expression network) (ภาคผนวก ช, ง, ค, ข, และ ก) โดยใช้หลักการของ Pearson Correlation Analysis ภายใต้อัลกอริทึม Cladist (Fukuchi et al., 2009) ที่ r-value 0.9 (+,-) และใช้ JActiveModules (Ideker et al., 2002) เพื่อช่วยในการคัดเลือกโมดูลย่อยของเครือข่ายที่แสดงออกถึงความสัมพันธ์อย่างเด่นชัด และการใช้โปรแกรม Network Analysis ที่อยู่ในโปรแกรมออนไลน์ Arraymining (Glaab et al., 2009) ที่จะทำการสร้างเครือข่ายการแสดงออกร่วมระหว่างยีน รวมถึงได้ใช้วิธีการสร้างเครือข่ายการควบคุมการแสดงออกของยีน (ภาคผนวก ฉ และ ก) โดยการใช้โปรแกรม DREM 2.0 (Dynamic Regulatory Events Miner version 2.0; Schulz et al., 2012) (ภาคผนวก ฉ) ที่อาศัยการบูรณาการกับข้อมูลปฏิสัมพันธ์ของการควบคุมระหว่างทรานสคริปชันแฟกเตอร์กับยีนเป้าหมาย (TF-TG interaction) ซึ่งในที่นี้จะใช้ข้อมูลของหนูเมาส์ (Mouse; *Mus musculus*) (Ernst et al., 2010) และข้อมูลที่ได้จากบทความวิจัย และการศึกษาโปรแกรม CyNetworkBMA (Yeung et al., 2011) ซึ่งเป็นปลั๊กอินของโปรแกรม Cytoscape (Cline et al., 2007) ซึ่งรันโปรแกรมร่วมกับโปรแกรม R (R Core Team, 2013) ที่จะสร้างเครือข่ายควบคุมการแสดงออกของยีนที่อาศัยหลักการของอัลกอริทึม regression-based Bayesian Model Averaging (BMA) (ภาคผนวก ก) พร้อมการคัดกรองเครือข่ายร่วมกับข้อมูลยีนการอักเสบจาก PolySearch (Cheng, 2008)

จากการศึกษาขั้นต้นทำให้ได้กระบวนการทางชีวสารสนเทศเพื่อนำมาใช้ในการคัดเลือกกลุ่มยีนและเครือข่ายปฏิสัมพันธ์ที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบ มาทำการวิเคราะห์ร่วมกับข้อมูลทางชีวภาพ

ได้แก่ ยีนออนโทโลยี (Gene Ontology; GO) (Ashburner *et al.*, 2000) และพารเวย์ KEGG (Kyoto Encyclopedia of Genes and Genomes) (Kanehisa, 2010) ก่อนที่จะนำข้อมูลที่ได้มาตัดสินใจร่วมกับข้อมูลของบทความวิจัยที่เกี่ยวข้องกับกลุ่มยีนที่คัดเลือกเหล่านั้น เพื่อนำมายืนยันผลในระดับของห้องปฏิบัติการต่อไป

4.2 การเตรียมข้อมูลไมโครอาร์เรย์ และการคัดเลือกกลุ่มยีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบ

การเตรียมข้อมูลไมโครอาร์เรย์ทั้งสามชุดข้อมูล โดยการปรับมาตรฐานข้อมูลให้อยู่มาตรฐานเดียวกัน โดยปรับข้อมูลให้อยู่ในรูปของค่ามาตรฐาน Z score จะพบว่าข้อมูลไมโครอาร์เรย์ที่ยังไม่ผ่านกระบวนการจัดเตรียมข้อมูลและกระบวนการนอร์มอไลซ์เซชันค่าตัวเลขของข้อมูลจะมีการกระจาย (ดังภาพที่ 4-1A) แต่เมื่อข้อมูลได้ผ่านกระบวนการนอร์มอไลซ์เซชันจะพบว่าค่าตัวเลขของข้อมูลจะอยู่ในมาตรฐานเดียวกัน (ดังภาพที่ 4-1B – 4-1C)

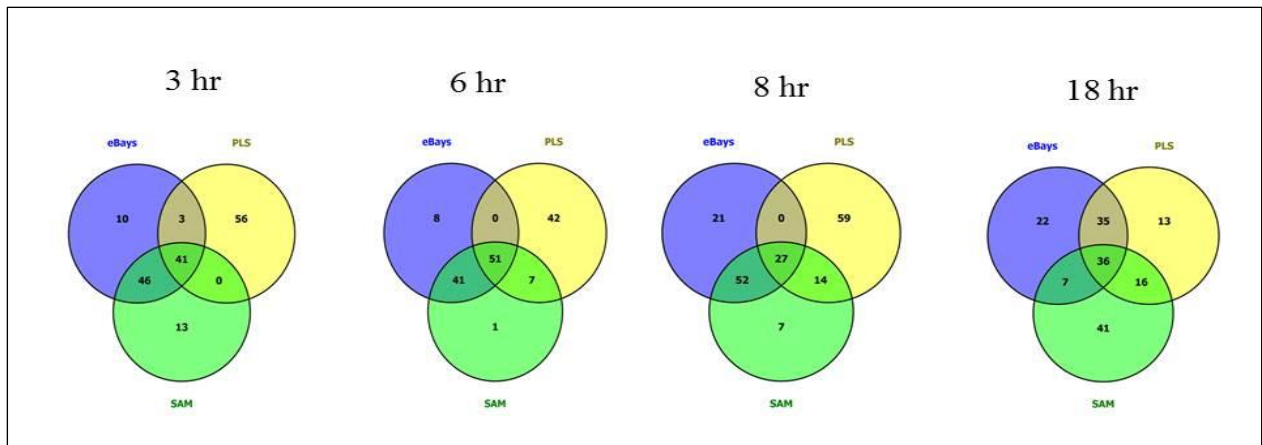


ภาพที่ 4-1 ข้อมูลไมโครอาร์เรย์ทั้ง 3 ชุดข้อมูลก่อนและหลังผ่านกระบวนการปรับมาตรฐานข้อมูล Box plot raw data (A), \log_{10} -transformed data (B) และ normalized data (C)

การคัดเลือกกลุ่มยีนที่มีการตอบสนองต่อการอักเสบ เมื่อมีการกระตุ้นต่อ LPS ที่เวลา 3, 6, 8 และ 18 ชั่วโมง โดยใช้วิธีการวิเคราะห์ทางสถิติที่ต่างกันสามวิธี คือ SAM, eBayes และ PLS-CV โดยทั้งสามวิธีนี้จะบรรจุอยู่ในโปรแกรมการทำเหมืองข้อมูลออนไลน์ ArrayMining ผลที่ได้จะคัดเลือกโพรบที่มีการแสดงออกอย่างเด่นชัด 100 โพรบแรก และเพื่อให้การคัดเลือกมีความน่าเชื่อถือเพิ่มขึ้นจึงเลือกเฉพาะโพรบที่ถูกคัดเลือกอย่างน้อยสองวิธีดังผลที่แสดงในตารางที่ 4-1 และภาพที่ 4-2 ซึ่งจะพบว่าในชุดข้อมูลที่มีการกระตุ้นต่อ LPS ที่เวลา 6 ชั่วโมงจะมีโพรบที่ถูกคัดเลือกมากที่สุด จากนั้นนำโพรบที่ถูกคัดเลือกในแต่ละช่วงเวลาไปค้นหาชื่อยีนในชุดข้อมูลไมโครอาร์เรย์จะพบว่าเมื่อการกระตุ้นด้วย LPS ที่เวลา 3 ชั่วโมง มีจำนวน 70 ยีน เมื่อกระตุ้นด้วย LPS ที่เวลา 6 ชั่วโมง จำนวน 65 ยีน เมื่อกระตุ้นด้วย LPS ที่เวลา 8 ชั่วโมง จำนวน 61 ยีน และเมื่อกระตุ้นด้วย LPS ที่เวลา 18 ชั่วโมง จำนวน 80 ยีน รายชื่อของยีนที่ถูกคัดเลือกในแต่ละช่วงเวลาแสดงในตารางที่ 4-2

ตารางที่ 4-1 จำนวนโพรบที่ถูกคัดเลือกโดยวิธี SAM, eBayes , PLS-CV และโพรบที่ถูกคัดเลือกอย่างน้อย 2 วิธี จาก 3 วิธี

ชุดข้อมูล	SAM (โพรบ)	eBayes (โพรบ)	PLS-CV (โพรบ)	จำนวนโพรบที่ถูกคัดเลือกสองในสามวิธี (โพรบ)
GSE4712 (3 hr_LPS)	100	100	100	90
GSE4712 (6 hr_LPS)	100	100	100	99
GSE21841 (8 hr_LPS)	100	100	100	93
GSE2002 (18 hr_LPS)	100	100	100	94



ภาพที่ 4-2 ไดอะแกรมแสดงจำนวนโพรบที่ถูกคัดเลือกโดยวิธี SAM, eBayes , PLS-CV และโพรบที่ถูกคัดเลือกอย่างน้อย 2 วิธี

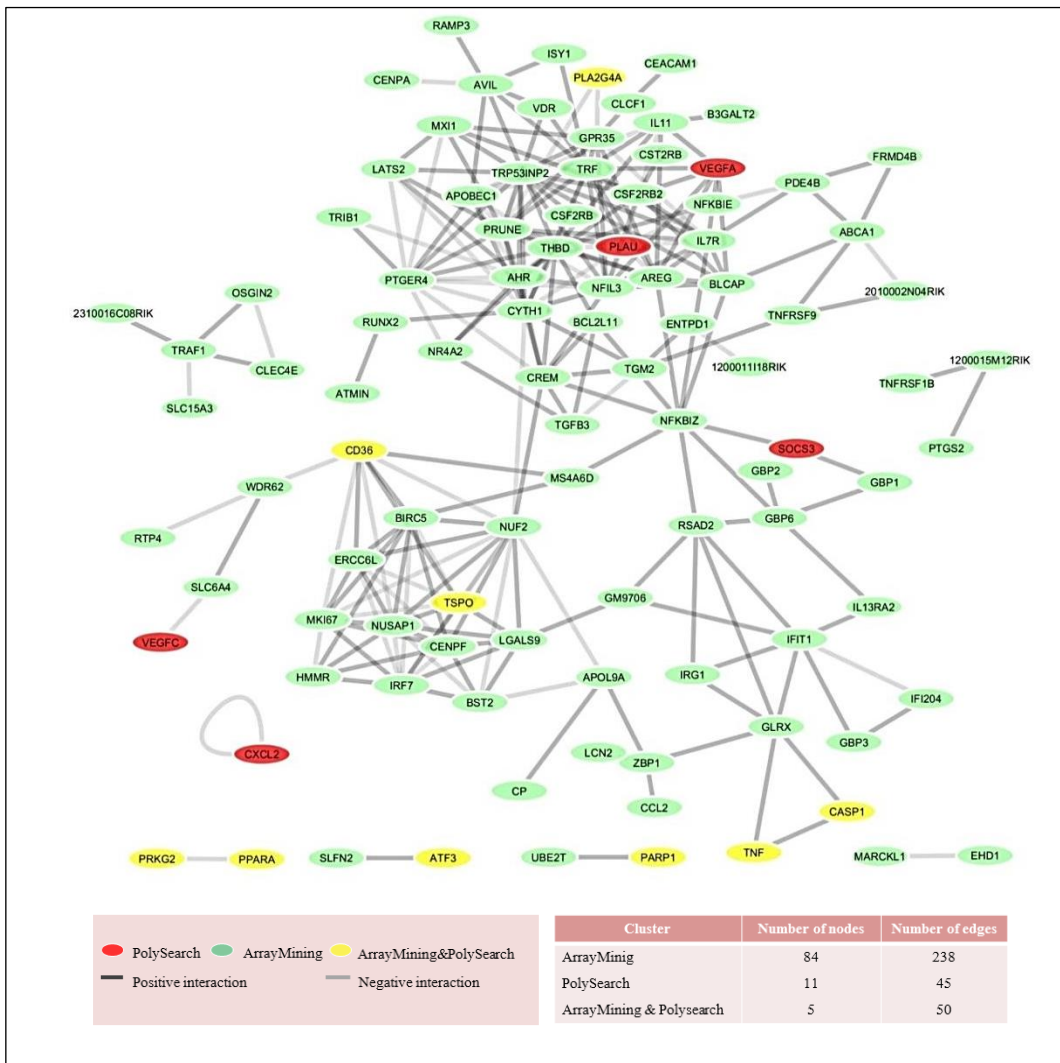
ตารางที่ 4-2 ยีนที่พบในแต่ละชุดข้อมูลไมโครอาร์เรย์จากโพรบที่ถูกคัดเลือกอย่างน้อยสองวิธี

ชุดข้อมูล	จำนวนยีน	ยีน
GSE4712 (3 hr_LPS)	70	<i>CD83, TGFB3, TGM2, 1190002H23RIK, CYTH1, CREM, NFIL3, TERF2IP, CLEC4E, SERPINA1A, BLCAP, SLC15A3, VEGFA, 2310016C08RIK, AREG, CSF2RB, THBS1, ABCA1, PLAU, PDE4B, AHR, B3GALT2, 2010002N04RIK, TRAF1, PTGER4, RUNX2, TRIB1, TNFRSF9, NUAK2, 9530003J23RIK, 4122401K19RIK, 4931439C15RIK, NFKBIE, MAST1, ATMIN, FNBP1L, CSRNP1, BCL2L11, 1110003E01RIK, CYTIP, DUSP5, FOSL2, GPR85, NR4A3, H2-EB2, D5ERTD163E, CD93, PEX19, LRIT2, E230029C05RIK, NR4A2, THBD, IL7R, HOXB2, IL11, CXCL2, MXI1, ENTPD1, APOBEC1, TRP53INP2, DZIP1, MPP7, CD302, AGPAT9, TSC22D1, 120001118RIK, SPTY2D1, FGD4, NAA11</i>
GSE4712 (6 hr_LPS)	65	<i>CD83, TGM2, VDR, CYTH1, CREM, NFIL3, AVIL, ISY1, CLEC10A, CLEC4E, RAMP3, PROCR, SLC15A3, 2310016C08RIK, AREG, CSF2RB, THBS1, ABCA1, PDE4B, AHR, SRC, 2010002N04RIK, TRAF1, PTGER4, OSGIN2, RUNX2, TRF, EDIL3, MXI1, TSC22D1, BCL2L11, EGR2, TNFRSF9, CHST11, 1110002L01RIK, NFKBIE, PRUNE, CSRNP1, 1110003E01RIK, CYTIP, EGR3, LATS2, DUSP5, FOSL2, CLCF1, FRMD4B, TMEM56, CD93, E230029C05RIK, PTCHD1, ZMYM3, NR4A2, THBD, IL7R, CSF2RB2, 120001118RIK, GPR35, CXCL2, CSF2RB , ENTPD1, ACSL4, TRP53INP2, ANKRD33B, GALNT13</i>
GSE21841 (8 hr_LPS)	61	<i>MARCKS, CD83, TNFAIP2, TNC, SOCS3, PTGS2, F8A, NFKBIZ, CCL22, CCL5, PHLDA1, ST3GAL1, CSF3, HOXB13, CCL2, ACCN5, RSAD2, LIF, IL1A, IFNB1, PDE4B, RAB3A, TPBG, TRAF1, ARHGEF3, TSC22D1, UNG, CSF2, CEACAM1, CCRL2, 1200003I10RIK , MIR155, PPFIA4, CDC42EP2, EXO1, 4921515G04RIK, 4930526H09RIK, RSPH3A, WDR37, 2610021J01RIK, PLSCR4, TNFAIP3, GBP6, 1200015M12RIK , PIM1, NEURL1B, A130040M12RIK, GBP2, E2F2, DUSP5, PRTG, CD40, RALGPS1, VEGFC, PRRG1, C130050O18RIK, PPP1R3E, KBTBD11, SDC4, CSF1, IL1B, CXCL2, IL6, DUSP2, IFT1, GADD45B, EDN1, 1200016E24RIK, 3930401B19RIK , HHAT, MAF, SHISA3, FAS</i>
GSE2002 (18 hr_LPS)	80	<i>MARCKSL1, MARCKS, EHD1, NUSAP1, SOCS3, GLRX, SLC6A4, IRF7, PTGS2, C2 /// CFB, NFKBIZ, CP, CCL22, TNFRSF1B, CCL5, GBP2, GBP3, RTP4, SLFN1, MAFF, PDPN, VEGFC, MS4A6D, IFI204, ZBP1, CLEC4N, CD274, CCL2, GBP1, SLC15A3, LGALS9, IGFBP4, IL13RA2, UBE2T, IL1RN, TRAF1, BIRC5, APOL9A , ERCC6L, WDR62, BST2, MYC, HMMR, SLC7A2, MKI67, PTPN23, CENPF, ZC3H12A, IRG1, LCN2, 1200003I10RIK , MIR155, NUF2, GM9706 , PLK3, SEMA5A, KIF11, KIF2C, ERI2, MPA2L, CDC20, C78513, MMP9, POLE, KIF20A, IL1B, DGKG, SLFN2, H2B, SAA3, CENPA, CCNB2, KIF22, CYTIP, BRCA1, HDC, 1200016E24RIK, UBE2C, OASL2, 3930401B19RIK , GM9706</i>

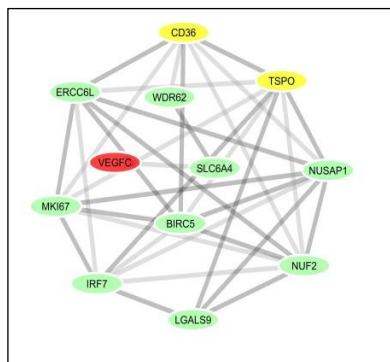
4.3 การออกแบบและบูรณาการแพลตฟอร์มจากข้อมูลไมโครอาร์เรย์

การค้นหากลุ่มยีนที่เกี่ยวข้องกับการอักเสบ (inflammatory genes) จากฐานข้อมูล PolySearch (Cheng, 2008) ในการค้นหายีนที่เกี่ยวข้องกับการอักเสบจะใช้คำสำคัญในการค้นหาว่า “inflammation” ซึ่งจะปรากฏรายชื่อยีนที่เกี่ยวข้องกับการอักเสบทั้งหมด 329 ยีน และนำรายชื่อยีนที่สืบค้นได้ไปเปรียบเทียบกับยีนที่พบในชุดข้อมูลไมโครอาร์เรย์ทั้งสามชุดข้อมูลจะมียีนทั้งหมด 187 ยีน (ดังตารางที่ 4-1) กลุ่มยีนที่เกี่ยวข้องกับการอักเสบที่ได้จากการสืบค้นจากฐานข้อมูล PolySearch และยีนที่มีการแสดงออกที่เด่นชัดที่ถูกคัดเลือกอย่างน้อยสองวิธีจากวิธีการวิเคราะห์ที่แตกต่างกันที่บรรจุภายในโปรแกรมการทำเหมืองข้อมูลออนไลน์ ArrayMining มีจำนวนทั้งหมด 312 ยีน จะถูกนำมาสร้างเครือข่ายความสัมพันธ์ระหว่างการแสดงออกร่วมของยีน โดยใช้ปลั๊กอิน Expression Correlation Network ที่ถูกบรรจุภายใต้โปรแกรม Cytoscape โดยจะเลือกความสัมพันธ์ที่มีค่า r-value ที่ ± 0.95 จะมียีนทั้งหมด 100 ยีน 333 ปฏิสัมพันธ์ ซึ่งจะเป็นยีนที่ถูกคัดเลือกด้วยวิธีการทำเหมืองข้อมูลในโปรแกรมออนไลน์ ArrayMining จำนวน 84 ยีน 238 ปฏิสัมพันธ์ ยีนที่ได้จากการสืบค้นฐานข้อมูล PolySearch จำนวน 11 ยีน 45 ปฏิสัมพันธ์ และยีนที่ถูกคัดเลือกจากการทำเหมืองข้อมูลในโปรแกรมออนไลน์ ArrayMining และจากการสืบค้นจากฐานข้อมูล PolySearch จำนวน 5 ยีน 50 ปฏิสัมพันธ์ ดังภาพที่ 4-3

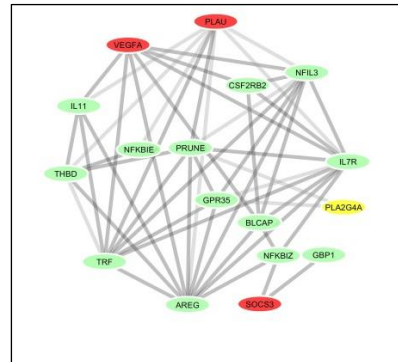
จากนั้นทำการคัดเลือกเครือข่ายย่อยจากเครือข่ายความสัมพันธ์ระหว่างการแสดงออกร่วมของยีนเพื่อให้ง่ายต่อการวิเคราะห์และการแปลผลโดยคัดเลือกจากยีนที่ได้จากการสืบค้นจากฐานข้อมูล PolySearch เป็นหลัก เนื่องจากเป็นกลุ่มยีนที่ได้มีการศึกษาแล้วว่าเกี่ยวข้องกับกระบวนการอักเสบ จะคัดเลือกเครือข่ายย่อยได้ทั้งหมด 2 เครือข่าย ดังภาพที่ 4-4 ในเครือข่ายย่อยที่ 1 จะมียีนจำนวน 12 ยีน 34 ปฏิสัมพันธ์ และเครือข่ายย่อยที่ 2 จะมียีนจำนวน 17 ยีน 50 ปฏิสัมพันธ์ โดยจะแสดงรายละเอียดของยีนทั้งสองเครือข่ายในตารางที่ 4-4



ภาพที่ 4-3 เครือข่ายการแสดงผลออกพร้อมของยีนที่ได้จากการคัดเลือกจากโปรแกรม ArrayMining และยีนที่ได้จากการสืบค้นจากฐานข้อมูล PolySearch



Module 1



Module 2

ภาพที่ 4-4 เครือข่ายย่อยที่ถูกคัดเลือกจากเครือข่ายการแสดงผลออกพร้อมของยีน

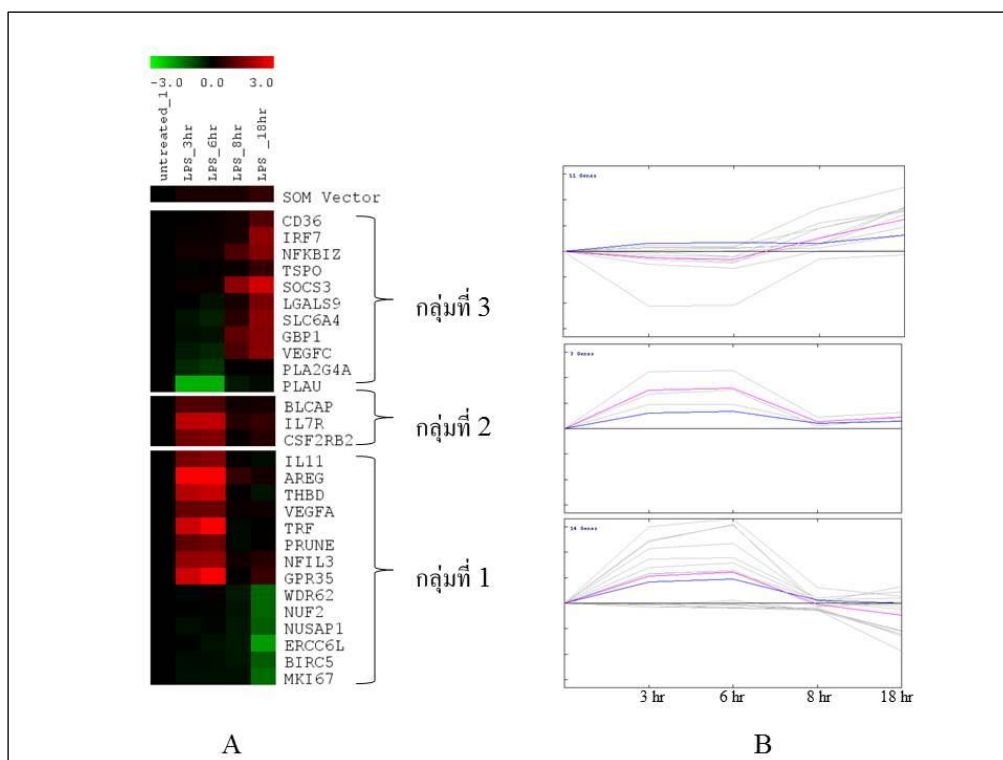
ตารางที่ 4-3 รายละเอียดของยีนในแต่ละเครือข่ายย่อยทั้งสองเครือข่าย

Gene Symbol	Gene name	RefSeq Transcript ID
Module1		
<i>CD36</i>	CD36 antigen	NM_001159555 /// NM_001159556 /// NM_001159557 /// NM_001159558 /// NM_007643
<i>WDR62</i>	WD repeat domain 62	NM_146186
<i>NUF2</i>	NUF2, NDC80 kinetochore complex component, homolog (S. cerevisiae)	NM_023284
<i>MKI67</i>	antigen identified by monoclonal antibody Ki 67	NM_001081117
<i>ERCC6L</i>	excision repair cross-complementing rodent repair deficiency complementation group 6 - like	NM_146235
<i>BIRC5</i>	baculoviral IAP repeat-containing 5	NM_001012273 /// NM_009689
<i>NUSAP1</i>	nucleolar and spindle associated protein 1	NM_001042652 /// NM_133851
<i>LGALS9</i>	lectin, galactose binding, soluble 9	NM_001159301 /// NM_010708
<i>TSPO</i>	translocator protein	NM_009775
<i>IRF7</i>	interferon regulatory factor 7	NM_016850
<i>SLC6A4</i>	solute carrier family 6 (neurotransmitter transporter, serotonin), member 4	NM_010484
<i>VEGFC</i>	vascular endothelial growth factor C	NM_009506
Module2		
<i>IL11</i>	interleukin 11	NM_008350
<i>GPR35</i>	G protein-coupled receptor 35	NM_001104529 /// NM_022320
<i>PLA2G4A</i>	phospholipase A2, group IVA (cytosolic, calcium- dependent)	NM_008869
<i>TRF</i>	transferrin	NM_133977
<i>PRUNE</i>	prune homolog (Drosophila)	NM_173347
<i>NFIL3</i>	nuclear factor, interleukin 3, regulated	NM_017373
<i>AREG</i>	amphiregulin	NM_009704
<i>BLCAP</i>	bladder cancer associated protein homolog (human)	NM_016916
<i>IL7R</i>	interleukin 7 receptor	NM_008372
<i>VEGFA</i>	vascular endothelial growth factor A	NM_001025250 /// NM_001025257 /// NM_001110266 /// NM_001110267 /// NM_001110268 /// NM_009505
<i>CSF2RB2</i>	colony stimulating factor 2 receptor, beta 2, low- affinity (granulocyte-macrophage)	NM_007781
<i>PLAU</i>	plasminogen activator, urokinase	NM_008873
<i>THBD</i>	thrombomodulin	NM_009378
<i>NFKBIZ</i>	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta	NM_001159394 /// NM_001159395 /// NM_030612
<i>SOCS3</i>	suppressor of cytokine signaling 3	NM_007707
<i>GBP1</i>	guanylate binding protein 1	NM_010259

4.4 การวิเคราะห์กลุ่มยีนที่มีแนวโน้มการแสดงออกร่วมกัน และการวิเคราะห์ร่วมกับเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีน

จากผลการวิเคราะห์เครือข่ายการแสดงออกของยีนและการคัดเลือกเครือข่ายย่อยจากนั้นนำมาจัดกลุ่มยีนที่มีตามรูปแบบการแสดงออกโดยวิธี SOM (self-organizing map) พบว่าสามารถจัดกลุ่มยีนตามรูปแบบการแสดงออกแบ่งออกได้จำนวน 3 กลุ่ม คือ กลุ่มที่ 1 จะมีรูปแบบการแสดงออกสูงในช่วงเวลา 3 และ 6 ชั่วโมงและลดการแสดงออกลงที่ช่วงเวลา 8 และ 18 ชั่วโมง กลุ่มที่ 2 จะมีรูปแบบการแสดงออกสูงในช่วงเวลา 3 และ 6 ชั่วโมงและลดการแสดงออกที่ช่วงเวลา 8 ชั่วโมงและมีการแสดงออกเพิ่มขึ้นอีกในช่วงเวลา 18 ชั่วโมง และกลุ่มที่ 3 จะมีรูปแบบการแสดงออกที่สูงในช่วงเวลา 8 และ 18 ชั่วโมง ดังภาพที่ 4-5A และ 4-5B ในกลุ่มที่ 1 จะมียีนจำนวน 14 ยีน กลุ่มที่ 2 จำนวน 3 ยีน และกลุ่มที่ 3 จำนวน 11 ยีน ดังตารางที่ 4-4 โดยยีนส่วนใหญ่จะมีรูปแบบการแสดงออกในกลุ่มที่ 1 และกลุ่มที่ 3

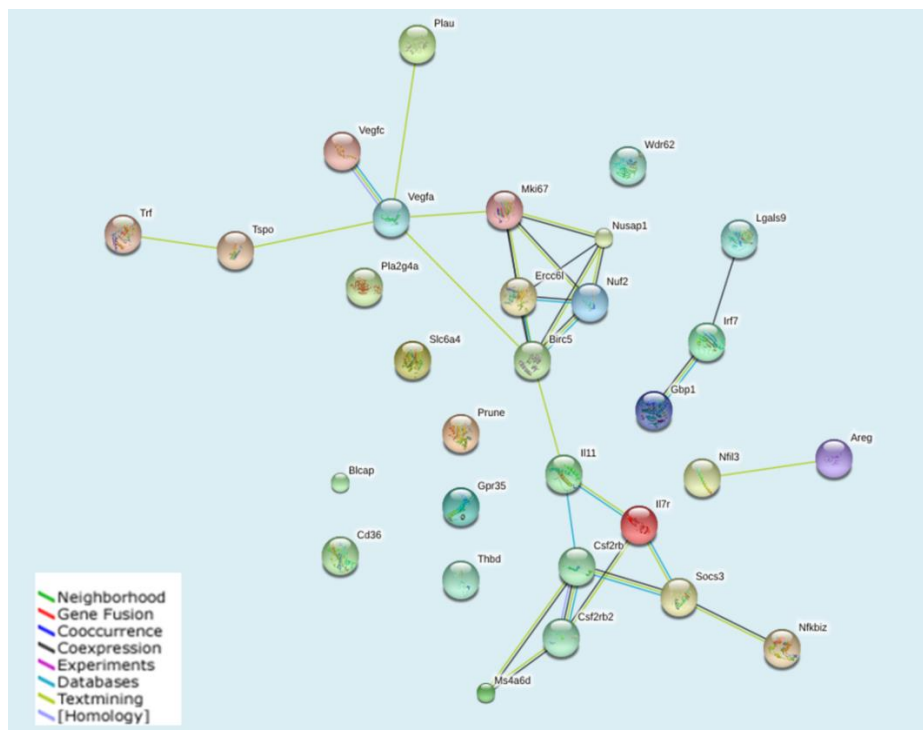
จากการคัดเลือกเครือข่ายย่อยยีนที่อยู่ภายในทั้งสองเครือข่ายย่อยจะถูกนำไปสร้างเป็นเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนในฐานข้อมูล STRING (Szklarczyk et al., 2015) ซึ่งเป็นฐานข้อมูลทางชีวภาพที่สามารถใช้ในการทำนายปฏิสัมพันธ์ระหว่างโปรตีนโดยฐานข้อมูล STRING จะรวบรวมข้อมูลจากแหล่งต่างๆไว้ประกอบด้วยข้อมูลที่ได้จากการทดลอง (experiment) ข้อมูลจากฐานข้อมูล (database) ข้อมูลจากการแสดงออกร่วม (co-expression) ข้อมูลจากการค้นหาความรู้จากข้อความ (text mining) เป็นต้น จะได้ผลดังภาพที่ 4-6



ภาพที่ 4-5 Heat map รูปแบบการแสดงออกของยีนทั้งสามกลุ่ม (A) และกราฟแสดงแนวโน้มรูปแบบการแสดงออกของยีน (B) เมื่อถูกกระตุ้นด้วย LPS ที่เวลา 3, 6, 8 และ 18 ชั่วโมง จากการจัดกลุ่มตามรูปแบบการแสดงออกของยีนโดยวิธี SOM

ตารางที่ 4-4 การจัดกลุ่มยีนที่มีการแสดงออกร่วมกันเมื่อมีการกระตุ้นด้วย LPS ที่เวลา 3 6 8 และ 18 ชั่วโมง โดยวิธี SOM (self-organizing map)

กลุ่มที่	จำนวนยีน (ยีน)	ยีน
1	14	<i>IL11, AREG, THBD, VEGFA, TRF, PRUNE, NFIL3, GPR35, WDR62, NUF2, NUSAP1, ERCC6L, BIRC5, MKI67</i>
2	3	<i>BLCAP, IL7R, CSF2RB2</i>
3	11	<i>CD36, IRF7, NFKBIZ, TSPO, SOCS3, LGALS9, SLC6A4, GBP1, VEGFC, PLA2G4A, PLAU</i>



ภาพที่ 4-6 ปฏิสัมพันธ์ระหว่างโปรตีนของยีนในเครือข่ายย่อยทั้งสองเครือข่ายในฐานข้อมูล STRING

4.5 การวิเคราะห์เครือข่ายปฏิสัมพันธ์ และการแปลผลทางชีวภาพ

จากการคัดเลือกเครือข่ายย่อยยีนที่อยู่ภายในทั้งสองเครือข่ายย่อยจะถูกนำมาวิเคราะห์หาหน้าที่และความสำคัญในกระบวนการทางชีวภาพในฐานข้อมูลสาธารณะ DAVID (Huang et al., 2009) แปลผลโดยใช้ยีนออนโทโลยี (Gene ontology) และฐานข้อมูล KEGG (Kanehisa, 2010) ซึ่งจะคัดเลือกเฉพาะกระบวนการที่มีค่า p-value < 0.05 โดยการแปลผลของยีนออนโทโลยีจะประกอบด้วย 3 กลุ่ม คือ 1) กระบวนการทางชีววิทยา (biological processes) จะพบว่ากลุ่มยีนจะเกี่ยวข้องกับกระบวนการทางชีววิทยาทั้งหมด 19 กระบวนการ โดยยีนส่วนใหญ่จะเกี่ยวข้องกับวัฏจักรของเซลล์ ดังตารางที่ 4-6 2) องค์ประกอบของเซลล์ (cellular components) จะพบว่ายีนส่วนใหญ่จะจะมีการทำงานบริเวณเยื่อหุ้มชั้นนอกและโครโมโซม ดังตารางที่ 4-7 และ 3) หน้าที่ของโมเลกุลในเซลล์ (molecular functions) จะพบว่ากลุ่มยีนจะมีหน้าที่เกี่ยวข้องกับการจับของไซโตไคน์และการกระตุ้น growth factor ดังตารางที่ 4-8 และในการค้นหาหน้าที่ของกลุ่มยีนในฐานข้อมูล KEGG pathway จะพบว่ายีนส่วนใหญ่จะเกี่ยวข้องกับวิถีการส่งสัญญาณ Jak-STAT signaling pathway และปฏิสัมพันธ์ระหว่างไซโตไคน์ (Cytokine-cytokine receptor interaction ดังตารางที่ 4-9

ตารางที่ 4-5 Biological process (BP) ของยีนในทั้งสองเครือข่ายย่อยโดยใช้ยีนออนโทโลยี

GO term ID	Go term description	P-value	Genes
GO:0000279	M phase	1.40E-03	<i>MKI67, BIRC5, ERCC6L, NUSAP1, NUF2</i>
GO:0045597	Positive regulation of cell differentiation	3.50E-03	<i>CD36, IL7R, SOCS3, VEGFC</i>
GO:0007049	Cell cycle	3.80E-03	<i>MKI67, BIRC5, ERCC6L, NUSAP1, BLCAP, NUF2</i>
GO:0007067	Mitosis	4.40E-03	<i>BIRC5, ERCC6L, NUSAP1, NUF2</i>
GO:0000280	Nuclear division	4.40E-03	<i>BIRC5, ERCC6L, NUSAP1, NUF2</i>
GO:0000087	M phase of mitotic cell cycle	4.70E-03	<i>BIRC5, ERCC6L, NUSAP1, NUF2</i>
GO:0048285	Organelle fission	4.90E-03	<i>BIRC5, ERCC6L, NUSAP1, NUF2</i>
GO:0007059	Chromosome segregation	5.60E-03	<i>BIRC5, NUSAP1, NUF2</i>
GO:0001666	Response to hypoxia	5.60E-03	<i>PLAU, TRF, VEGFA</i>
GO:0070482	Response to oxygen levels	5.80E-03	<i>PLAU, TRF, VEGFA</i>
GO:0051094	Positive regulation of developmental process	6.10E-03	<i>CD36, IL7R, SOCS3, VEGFC</i>
GO:0008283	Cell proliferation	9.10E-03	<i>AREG, MKI67, IL7R, VEGFA</i>
GO:0051301	Cell division	1.30E-02	<i>BIRC5, ERCC6L, NUSAP1, NUF2</i>
GO:0000226	Microtubule cytoskeleton organization	1.60E-02	<i>BIRC5, NUSAP1, NUF2</i>
GO:0001763	Morphogenesis of a branching structure	2.00E-02	<i>AREG, SOCS3, VEGFA</i>
GO:0048010	Vascular endothelial growth factor receptor signaling pathway	2.40E-02	<i>VEGFA, VEGFC</i>
GO:0060749	Mammary gland alveolus development	2.60E-02	<i>AREG, VEGFA</i>
GO:0000910	Cytokinesis	4.70E-02	<i>BIRC5, NUSP1</i>
GO:0006955	Immune response	4.90E-02	<i>IRF7, IL7R, GBP1, VEGFA</i>

ตารางที่ 4-6 Cellular component (CC) ของยีนในทั้งสองเครือข่ายย่อยโดยใช้ยีนออนโทโลยี

GO term ID	Go term description	P-value	Genes
GO:0000775	Chromosome centromeric region	1.20E-03	<i>NUF2, BIRC5, MKI67, ERCC6L</i>
GO:0005615	Extracellular space	2.50E-03	<i>IL11, AREG, THBD, TRF, VEGFA, VEGFC</i>
GO:0044421	Extracellular region part	1.40E-02	<i>AREG, IL11, THBD, TRF, VEGFA, VEGFC</i>
GO:0000793	Condensed chromosome	1.80E-02	<i>MKI67, ERCC6L, NUF2</i>
GO:0044427	Chromosomal part	2.20E-02	<i>MKI67, BIRC5, NUF2, ERCC6L</i>
GO:0005876	Spindle microtubule	3.20E-02	<i>BIRC5, NUSAP1</i>
GO:0005576	Extracellular region	3.40E-02	<i>AREG, IL11, LGALS9, PLAU, THBD, TRF, VEGFA, VEGFC</i>
GO:0005694	Chromosome	3.50E-02	<i>MKI67, BIRC5, NUF2, ERCC6L</i>

ตารางที่ 4-7 Molecular function ของยีนในทั้งสองเครือข่ายย่อยโดยใช้ยีนออนโทโลยี

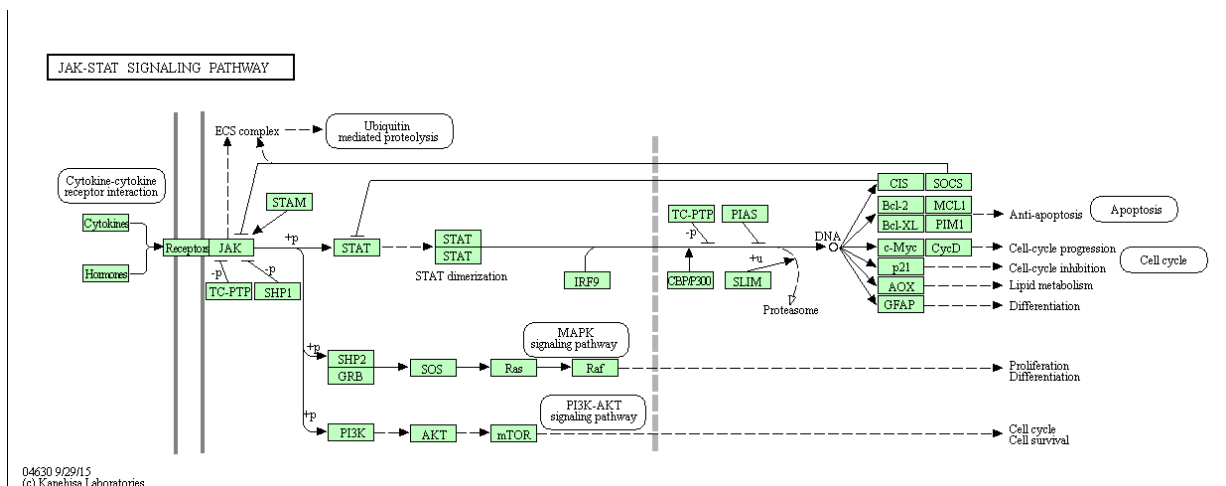
GO term ID	Go term description	P-value	Genes
GO:0008083	Growth factor activity	1.20E-03	<i>AREG, IL11, VEGFA, VEGFC</i>
GO:0004896	Cytokine receptor activity	3.00E-03	<i>CSF2RB2, IL7R</i>
GO:0019955	Cytokine binding	7.60E-03	<i>CSF2RB2, IL7R</i>

ตารางที่ 4-8 หน้าที่และกระบวนการทางชีวภาพของยีนในทั้งสองเครือข่ายย่อยฐานข้อมูล Kyoto Encyclopedia of Genomes pathway enrichment analysis (KEGG)

Term	P-value	Genes
mmu04060 Cytokine-cytokine receptor interaction	4.00E-04	<i>CSF2RB2, IL11, VEGFA, VEGFC</i>
mmu04630 Jak-STAT signaling pathway	6.70E-04	<i>CSF2RB2, IL11, IL7R, SOCS3</i>
mmu04640 Hematopoietic cell lineage	2.20E-02	<i>CD36, IL11, IL7R</i>

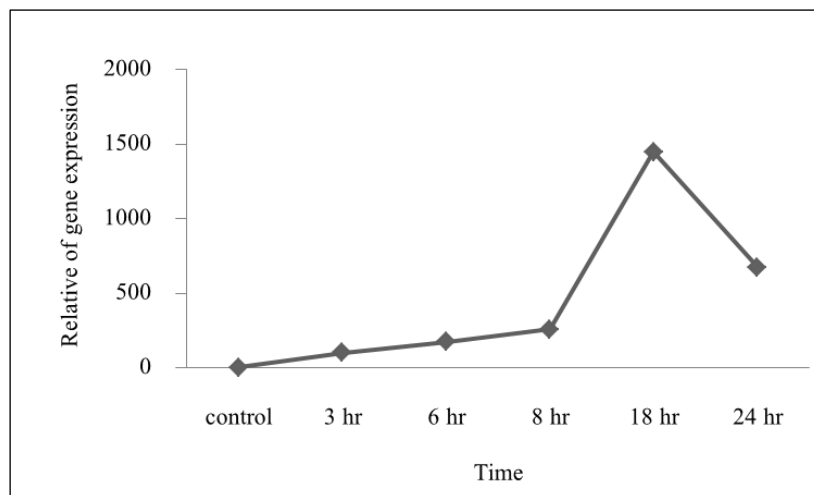
4.6 การคัดเลือกยีนและการตรวจสอบการแสดงออกของยีนระดับ mRNA ในห้องปฏิบัติการ โดยใช้เทคนิค real time PCR

จากการวิเคราะห์หาหน้าที่และความสำคัญในกระบวนการทางชีวภาพของยีนในทั้งสองเครือข่ายย่อย จากฐานข้อมูล KEGG pathway จะพบว่ายีนส่วนใหญ่จะเกี่ยวข้องกับวิถีการส่งสัญญาณ Jak-STAT signaling pathway ดังภาพที่ 4-7 (Jatiani *et al.*, 2010) โดยยีนที่ได้จากการคัดเลือกที่เกี่ยวข้องกับวิถีนี้ประกอบด้วย 4 ยีน คือ *CSF2RB2*, *IL11*, *IL7R* และ *SOCS3* ซึ่งยีน *SOCS* เป็นยีนที่มีความน่าสนใจที่นำไปตรวจสอบระดับการแสดงออกของยีนในห้องปฏิบัติการต่อไป เนื่องจากยีน *SOCS3* เป็นยีนที่ถูกคัดเลือกได้จากทั้งวิธีการทำเหมืองข้อมูลและการสืบค้นจากฐานข้อมูล PolySearch และมีรูปแบบการแสดงออกสูงในช่วงเวลา 8 และ 18 ชั่วโมง (ช่วงท้าย) และทำการตรวจสอบระดับการแสดงออกของยีนร่วมกับยีน *iNOS* ซึ่งเป็นยีนเครื่องหมายที่สำคัญของการตอบสนองต่อการอักเสบที่มีผู้ศึกษากันมาก เพื่อให้มั่นใจว่ามีการอักเสบเกิดขึ้นแน่นอน

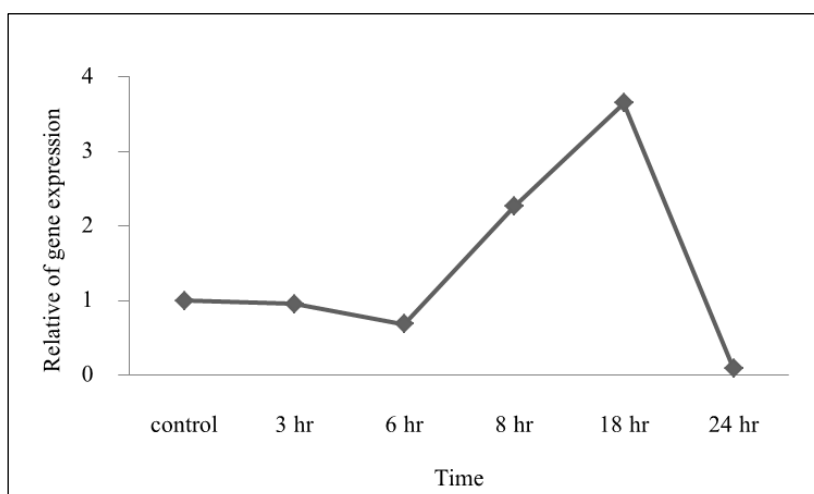


ภาพที่ 4-7 วิถีการส่งสัญญาณ Jak-STAT signaling pathway (ที่มา: http://www.genome.jp/kegg-bin/show_pathway?hsa04630 สืบค้นเมื่อวันที่ 23 สิงหาคม 2559)

ผลการวิเคราะห์หัตถปริมาณ mRNA ของยีน *iNOS* และ *SOCS3* โดยมียีน *EF2* เป็นยีนอ้างอิง (Housekeeping gene) โดยใช้เทคนิค real time PCR พบว่ายีน *iNOS* จะมีระดับการแสดงออกในสภาวะที่มีการกระตุ้นด้วย LPS ที่ช่วงเวลา 3 6 8 และ 18 ชั่วโมงตามลำดับ และมีการแสดงออกลดลงเมื่อมีการกระตุ้นด้วย LPS ที่ช่วงเวลา 24 ชั่วโมง โดยในช่วงที่มีการกระตุ้นด้วย LPS ที่เวลา 24 ชั่วโมง จะมีระดับการแสดงออกที่สูงที่สุด (ภาพที่ 4-8) และยีน *SOCS3* จะมีการแสดงออกที่ลดลงในสภาวะที่มีการกระตุ้นด้วย LPS ที่ช่วงเวลา 3 และ 6 ชั่วโมงตามลำดับ จากนั้นเมื่อมีการกระตุ้นด้วย LPS ที่ช่วงเวลา 8 และ 18 ชั่วโมง จะมีระดับการแสดงออกเพิ่มขึ้นโดยช่วงเวลาที่มีการกระตุ้นด้วย LPS ที่ 18 ชั่วโมงจะมีระดับการแสดงออกสูงสุด และเมื่อมีการกระตุ้นด้วย LPS ที่ช่วงเวลา 24 ชั่วโมงจะมีระดับการแสดงออกลดลงต่ำสุด (ภาพที่ 4-9)



ภาพที่ 4-8 ระดับการแสดงออกของ mRNA ของยีน *iNOS* ในเซลล์แมคโครฟาจ RAW264.7 ภายใต้สภาวะที่มีการกระตุ้นด้วย LPS ตามช่วงเวลา 3, 6, 8, 18, และ 24 ชั่วโมง



ภาพที่ 4-9 ระดับการแสดงออกของ mRNA ของยีน *SOCS3* ในเซลล์แมคโครฟาจ RAW264.7 ภายใต้สภาวะที่มีการกระตุ้นด้วย LPS ตามช่วงเวลา 3, 6, 8, 18, และ 24 ชั่วโมง

บทที่ 5

อภิปราย สรุปผลการทดลอง และข้อเสนอแนะ

5.1 อภิปรายผลการทดลอง

จากการศึกษาเบื้องต้นโดยทีมงานของผู้วิจัยที่เกี่ยวข้องกับการคัดเลือกยีนที่เกี่ยวข้องกับการอักเสบจากชุดข้อมูลไมโครอาร์เรย์ให้มีความน่าเชื่อถือมากยิ่งขึ้น โดยอาศัยกระบวนการของการจัดจำแนกกลุ่มตัวอย่าง (Classification) มาใช้ในการสร้างโมเดลของการจัดจำแนก ก่อนทำการทดสอบโมเดลที่ได้ทำให้พบว่าถึงแม้ค่าความถูกต้องของการจัดจำแนกจะมีค่าสูง แต่เนื่องจากจำนวนตัวอย่างที่ใช้ในการตรวจสอบมีเป็นจำนวนน้อยคือมีจำนวนตัวอย่างในแต่ละกลุ่มเพียง 3 ตัวอย่าง ทำให้สามารถเกิดข้อผิดพลาดในการยืนยันผลของกลุ่มยีนที่ถูกคัดเลือกมาได้เช่นเดียวกัน (ปรัศนี, มินตรา และพิทักษ์, 2556) และเมื่อทำการคัดเลือกกลุ่มยีนที่มีการแสดงออกแตกต่างกันชัดเจนมาวิเคราะห์ร่วมกับข้อมูลพาธเวย์จำนวน 6 พาธเวย์ ที่เกี่ยวข้องกับการตอบสนองการอักเสบ ซึ่งคัดเลือกมาจากฐานข้อมูล KEGG (<http://www.genome.jp/kegg/>; Kanehisa, 2010) พบว่ายังคงพบข้อจำกัดในแง่ของการอธิบายกลไกที่เกิดขึ้นของกระบวนการตอบสนองการอักเสบที่เกิดขึ้น เนื่องจากการคัดเลือกเฉพาะกลุ่มยีนซึ่งมีจำนวนน้อยอาจจะทำให้ไม่สามารถอธิบายภาพรวมของการตอบสนองที่ควรจะเป็นได้ (เหมือนฝัน, กอบเกตุ และพิทักษ์, 2556) จึงเป็นที่มาของการศึกษาการตอบสนองการอักเสบเพื่อทำการบูรณาการการวิเคราะห์ข้อมูลไมโครอาร์เรย์ โดยอาศัยกระบวนการทางชีวสารสนเทศ ทั้งการคัดเลือกยีน (Gene selection) (ภาคผนวก ฉ, ช, จ) การสร้างเครือข่ายความสัมพันธ์จากระดับการแสดงออกของยีน (Co-expression network) (ภาคผนวก ช, ง, ค, ข, และ ก) การสร้างเครือข่ายการควบคุมการแสดงออกของยีน (ภาคผนวก ฉ และ ก) โดยนำข้อมูลของกลุ่มยีนหรือเครือข่ายที่ถูกคัดเลือกมาทำการวิเคราะห์ร่วมกับข้อมูลทางชีวภาพ ได้แก่ ยีนออนโทโลยี (Gene Ontology; GO) (Ashburner *et al.*, 2000) และพาธเวย์ KEGG (Kyoto Encyclopedia of Genes and Genomes) (Kanehisa, 2010) ก่อนที่จะนำข้อมูลที่ได้มาตัดสินใจร่วมกับข้อมูลของบทความวิจัยที่เกี่ยวข้องกับกลุ่มยีนที่คัดเลือกเหล่านั้น เพื่อนำมายืนยันผลในระดับของห้องปฏิบัติการต่อไป

จากการรวบรวมข้อมูลไมโครอาร์เรย์ของการแสดงออกของยีนที่บรรจุอยู่ในฐานข้อมูลสาธารณะ GEO (Edgar 2002, <http://www.ncbi.nlm.nih.gov/geo/>) โดยชุดข้อมูลไมโครอาร์เรย์ที่ได้นำมาวิเคราะห์นั้น จะคัดเลือกมา 2 กลุ่มการทดลองคือ กลุ่มที่ถูกกระตุ้นด้วย LPS และกลุ่มที่ไม่ถูกกระตุ้นด้วย LPS ตามช่วงเวลาในเซลล์แมคโครฟาจเซลล์ไลน์ RAW264.7 ทำให้ได้ชุดข้อมูลไมโครอาร์เรย์ที่ครอบคลุม 4 ช่วงเวลาของการกระตุ้นเซลล์ด้วย LPS จำนวน 3 ชุดข้อมูล ได้แก่ GSE4712 ในช่วงเวลา 3 และ 6 ชั่วโมง (Comer *et al.*, 2006) GSE21841 ในช่วงเวลา 8 ชั่วโมง (Hammer *et al.*, 2010) และ GSE2002 ในช่วงเวลา 18 ชั่วโมง (Shell *et al.*, 2005) จากข้อมูลไมโครอาร์เรย์ที่ได้จึงรองรับการศึกษาการอักเสบใน 4 ช่วงเวลา คือ 3, 6, 8 และ 18 ชั่วโมง ในการคัดเลือกยีน (Gene selection) จากข้อมูลที่ได้ของการศึกษาใน ภาคผนวก ช จะพบว่าเมื่อนำกระบวนการคัดเลือกยีนโดยใช้วิธี SVM (Support Vector Machines) ที่เป็นโปรแกรมย่อยภายในโปรแกรม WEKA (Hall *et al.*, 2009) ทำการคัดเลือกยีนสำคัญ 100 อันดับแรก จากยีนทั้งหมด 21,678 ยีน ร่วมกับใช้วิธีการเปรียบเทียบโพลด์เซนจ์ที่ระดับมากกว่า 1.4 ในระดับความเชื่อมั่นที่ $P < 0.05$ กลุ่มยีนที่ถูกคัดเลือกออกมายังไม่สามารถแปลผลให้สอดคล้องกับการตอบสนองการอักเสบที่เกิดขึ้นได้อย่างชัดเจน แสดงให้เห็นว่าการคัดเลือกยีนโดยวิธีการข้างต้นอาจจะยังไม่มีเหมาะสมเพียงพอ ดังนั้นเมื่อลอง

ปรับเปลี่ยนวิธีการคัดเลือกยีนมาใช้เครื่องมือที่มีการจัดทำไว้โดยเฉพาะสำหรับการคัดเลือกข้อมูลไมโครอาร์เรย์ที่มีจำนวนตัวอย่างน้อยในขณะที่จำนวนยีนมีเป็นจำนวนมาก (ภาคผนวก จ) ซึ่งเครื่องมือการคัดเลือกเหล่านี้ถูกรวบรวมไว้ในโปรแกรมออนไลน์ ArrayMining (Glaab et al., 2009) โดยใช้เครื่องมือคัดเลือกยีน 3 ชนิด ได้แก่ Significance analysis of microarrays (SAM) (Tusher, Tibshirani and Chu, 2001), Empirical Bayes moderated t-test (eBayes) (Smyth, 2004), และ PLS-CV-Partial Least Squares Cross-Validation (PLS-CV) (Hall, 2000) โดยคัดเลือกยีนที่ถูกคัดเลือกอย่างน้อย 2 จาก 3 วิธี เป็นยีนที่ตอบสนองต่อการอักเสบในแต่ละช่วงเวลา เมื่อนำมาแปลผลทางชีวภาพโดยใช้ DAVID (Huang, Sherman and Lempicki, 2009) ที่มีฐานข้อมูลของการแปลผลทางชีวภาพหลากหลายชนิด โดยใช้ข้อมูลการแปลผลจากฐานข้อมูลของยีนออนโทโลยี (Gene Ontology; GO) (Ashburner et al., 2000) ที่ประกอบไปด้วยการแปลผล 3 ลักษณะคือ Biological Process (BP), Cell Component (CC), และ Molecular Function (MF) พบว่า สามารถแปลผลออกมาสอดคล้องกับกลไกการตอบสนองการอักเสบที่เกิดขึ้นในช่วงเวลา 3, 6, 8, และ 18 ชั่วโมง แต่เนื่องจากไม่สามารถอธิบายความเชื่อมโยงของการแสดงออกของกลุ่มยีนแต่ละชนิดระหว่างกันได้ ทางผู้วิจัยจึงมีความคิดที่จะนำเอารูปแบบของการสร้างเครือข่ายปฏิสัมพันธ์ และเครือข่ายการควบคุมการแสดงออกของยีน รวมถึงการบูรณาการกับข้อมูลทางชีวภาพอื่น เพื่อมาช่วยอธิบายกลไกการตอบสนองการอักเสบที่เกิดขึ้นให้มีความละเอียดและชัดเจนยิ่งขึ้น

การสร้างเครือข่ายความสัมพันธ์ของระดับการแสดงออกของยีน (Co-expression network) จากภาคผนวก ค โดยใช้หลักการของ Pearson Correlation Analysis ภายใต้โปรแกรม Cladist (Fukuchi et al., 2009) ของกลุ่มยีนที่คัดเลือกมาจากพารเวย์ที่เกี่ยวข้องกับการตอบสนองการอักเสบจำนวนทั้งสิ้น 659 ยีน ทำให้ได้จำนวนปฏิสัมพันธ์ที่เกิดขึ้นทั้งสิ้น 434,281 ปฏิสัมพันธ์ เมื่อคัดเลือกเฉพาะยีนที่มีค่าโพลด์เซนจ์มากกว่า 1.4 ก็จะเหลือจำนวนยีน 48 ยีน 2,304 ปฏิสัมพันธ์ เมื่อเลือกเฉพาะที่ r-value 0.9 (+,-) ก็จะได้ 352 ปฏิสัมพันธ์ และเมื่อใช้ JActiveModules (Ideker et al., 2002) เพื่อช่วยในการคัดเลือกโมดูลย่อยของเครือข่ายที่แสดงออกถึงความสัมพันธ์อย่างเด่นชัด ก็จะทำให้ได้ 34 ยีน 155 ปฏิสัมพันธ์ เมื่อนำเครือข่ายที่ได้มาทำการจัดกลุ่มและใส่ข้อมูลระดับการแสดงออกในแต่ละช่วงเวลาเข้าไป ก็จะสามารถนำมาใช้ในการอธิบายกลไกของการอักเสบที่เกิดขึ้นในแต่ละช่วงเวลาได้ (ภาคผนวก ค) แต่เนื่องจากการสร้างเครือข่ายความสัมพันธ์ที่เกิดขึ้นนี้มีการใช้ยีนเพียงแค่ 659 ยีน จึงทำให้การตรวจสอบกลไกที่ได้เกิดความจำเพาะกับวิถีพารเวย์ที่เลือกมาเท่านั้น ไม่สามารถจะค้นพบวิถีพารเวย์ใหม่ได้ ทีมผู้วิจัยจึงได้ทดลองศึกษาการสร้างเครือข่ายความสัมพันธ์ของระดับการแสดงออกของยีนโดยใช้ข้อมูลจากโพรบทั้งหมด 45,037 โพรบ ที่ครอบคลุมยีนทั้งหมด 21,678 ยีน (ภาคผนวก ก) ภายใต้โปรแกรม Network Analysis ที่อยู่ในโปรแกรมออนไลน์ Arraymining (Glaab et al., 2009) ที่จะทำการสร้างเครือข่ายการแสดงออกร่วมระหว่างยีน ซึ่งการคัดเลือกนี้จะเหลือเครือข่ายที่มีขนาด 2,000 โพรบ และ 61,096 ปฏิสัมพันธ์ ก่อนนำเครือข่ายที่ได้มาวิเคราะห์ร่วมกับกลุ่มยีนที่ถูกคัดเลือกมาเช่นเดียวกับการศึกษาก่อนหน้านี้ที่มีการคัดเลือกยีนที่ถูกคัดเลือกอย่างน้อย 2 จาก 3 วิธี ได้แก่ SAM, eBays, และ PLS-CV (ภาคผนวก จ) แต่ผลที่ได้ไม่เป็นไปตามที่คาดหวังไว้เนื่องจากกลุ่มยีนที่ถูกคัดเลือกมาซึ่งมีความสัมพันธ์กันในเครือข่ายส่วนใหญ่ไม่มีความเกี่ยวข้องกับการตอบสนองการอักเสบ ทำให้ไม่สามารถอธิบายกลไกการตอบสนองการอักเสบได้อย่างชัดเจน

นอกจากการสร้างเครือข่ายความสัมพันธ์ของการแสดงออกร่วมของยีนข้างต้นแล้ว ทางผู้วิจัยยังได้ทดลองสร้างเครือข่ายควบคุมการแสดงออกของยีน โดยการใช้โปรแกรม DREM 2.0 (Dynamic Regulatory Events Miner version 2.0; Schulz et al., 2012) (ภาคผนวก ฉ) ซึ่งทำการสร้างเครือข่ายการควบคุมการ

แสดงออกของยีนจำนวนทั้งสิ้น 21,678 ยีน โดยอาศัยการบูรณาการกับข้อมูลปฏิสัมพันธ์ของการควบคุมระหว่างทรานสคริปชันแฟกเตอร์กับยีนเป้าหมาย (TF-TG interaction) ซึ่งในที่นี้จะใช้ข้อมูลของหนูเมาส์ (Mouse; *Mus musculus*) (Ernst et al., 2010) ทำให้ได้เครือข่ายที่ประกอบไปด้วยทรานสคริปชันแฟกเตอร์ 335 ยีน และยีนเป้าหมาย 1,636 ยีน ที่มีปฏิสัมพันธ์ทั้งสิ้น 44,000 ปฏิสัมพันธ์ และเมื่อคัดกรองข้อมูลกับงานวิจัยที่เกี่ยวข้องกับการอักเสบทำให้เหลือเครือข่ายที่ประกอบไปด้วย ทรานสคริปชันแฟกเตอร์ 5 ยีน ยีนเป้าหมาย 621 ยีน และปฏิสัมพันธ์ 1,022 ปฏิสัมพันธ์ และสุดท้ายทำการคัดกรองด้วยยีนที่มีค่าโพลด์เซนจ์มากกว่า 1.4 และการแบ่งกลุ่มฟังก์ชัน ทำให้เหลือเครือข่ายที่ประกอบไปด้วย ทรานสคริปชันแฟกเตอร์ 5 ยีน ยีนเป้าหมาย 6 ยีน และปฏิสัมพันธ์ 10 ปฏิสัมพันธ์ ที่ง่ายและสะดวกต่อการตรวจสอบ แต่เนื่องจากขนาดของเครือข่ายที่มีขนาดเล็กที่อาจไม่ครอบคลุมการอธิบายกลไกโดยรวมของการตอบสนองการอักเสบได้ดี และความยุ่งยากของการใช้งานโปรแกรม DREM ที่ยังต้องการการปรับเปลี่ยนข้อมูลเพื่อให้สามารถนำมาใช้งานได้กับการวิเคราะห์และแสดงผลภายในโปรแกรม Cytoscape (Cline et al., 2007) จึงทำให้ทีมผู้วิจัยได้ทดลองใช้โปรแกรม CyNetworkBMA (Yeung et al., 2011) ซึ่งเป็นปลั๊กอินของโปรแกรม Cytoscape (Cline et al., 2007) ซึ่งรันโปรแกรมร่วมกับโปรแกรม R (R Core Team, 2013) ที่จะสร้างเครือข่ายควบคุมการแสดงผลของยีนที่อาศัยหลักการของอัลกอริทึม regression-based Bayesian Model Averaging (BMA) (ภาคผนวก ก) แต่การทดลองนี้ได้ใช้ชุดข้อมูลการแสดงออกของยีน GSE52320 (Schott et al., 2014) ซึ่งเป็นข้อมูลการแสดงออกของยีนที่ได้จากเทคโนโลยี RNA-seq ที่ประกอบไปด้วยช่วงเวลาของการบ่ม LPS ที่ 0, 15, 30, 45, 60, 75, 90, และ 120 นาที โดยอาศัยการคัดกรองเครือข่ายร่วมกับข้อมูลยีนการอักเสบจาก PolySearch (Cheng, 2008) การตัดยีนที่มีการแสดงออกน้อย การวิเคราะห์โครงสร้างเครือข่าย พบทรานสคริปชันแฟกเตอร์จำนวน 8 ยีน และยีนเป้าหมายจำนวน 85 ยีน และเมื่อทำการคัดเลือกเครือข่ายย่อยโดยอาศัยการจัดกลุ่มยีนตามรูปแบบการแสดงออก (Expression profile) ออกเป็น 5 กลุ่ม จะพบว่ายีนแต่ละกลุ่มมีการตอบสนองต่อการกระตุ้นของ LPS ที่ช่วงเวลาแตกต่างกัน แต่เนื่องด้วยรูปแบบของข้อมูลที่ใช้มีความแตกต่างจากข้อมูลไมโครอาร์เรย์ในเรื่องของจำนวนซ้ำของตัวอย่างข้อมูล ทำให้ทีมผู้วิจัยยังไม่สามารถนำวิธีการ CynetworkBMA มาใช้ร่วมกับชุดข้อมูลไมโครอาร์เรย์เดิมที่มีอยู่ได้

จากผลการศึกษาข้างต้นทางทีมผู้วิจัยจึงได้รวบรวมแนวทางการเลือกใช้เครื่องมือทางชีวสารสนเทศ วิธีการวิเคราะห์และแปลผลข้อมูลทางชีวภาพของกลุ่มยีนที่ได้จากการคัดเลือกโดยการใช้เครือข่ายปฏิสัมพันธ์การแสดงออกของยีนและการคัดเลือกยีน (ภาคผนวก ข) เพื่อทำการค้นหากลุ่มยีนที่มีบทบาทสำคัญต่อการตอบสนองการอักเสบที่เกิดขึ้นโดยการใช้ชุดข้อมูลไมโครอาร์เรย์ทั้งสามชุดข้อมูล ได้แก่ GSE4712 (Comer et al., 2006), GSE21841 (Hammer et al., 2010), และ GSE2002 (Shell et al., 2005) ที่รองรับการศึกษากลไกการอักเสบใน 4 ช่วงเวลา คือ 3, 6, 8 และ 18 ชั่วโมง ก่อนที่จะทำการคัดเลือกยีนโดยการเลือกกลุ่มยีนที่ถูกคัดเลือกโดย 2 วิธี จาก 3 วิธี ได้แก่ SAM (Tusher, Tibshirani and Chu, 2001), eBayes (Smyth, 2004), และ PLS-CV (Hall, 2000) ก่อนที่จะนำมาคัดกรองกลุ่มยีนที่ถูกคัดเลือกได้อีกครั้งด้วยข้อมูลยีนอักเสบที่สืบค้นได้จาก PolySearch (Cheng, 2008) แล้วนำไปสร้างเครือข่ายการแสดงออกพร้อมด้วยโปรแกรมปลั๊กอิน ExpressionCorrelationViewer ของโปรแกรม Cytoscape (Cline et al., 2007) หลังจากนั้นก็ทำการคัดเลือกเครือข่ายย่อยซึ่งมีปฏิสัมพันธ์ของกลุ่มยีนที่น่าสนใจ ซึ่งคาดหวังว่าจะได้พบกลุ่มยีนที่มีการแสดงออกไม่เด่นชัด แต่อยู่ท่ามกลางปฏิสัมพันธ์ของยีนที่มีการแสดงออกอย่างเด่นชัดหรือเป็นยีนที่มีความสำคัญต่อการตอบสนองการอักเสบ ที่อาจจะเป้าหมายของการศึกษาในอนาคตต่อไป

จากผลการทดลองเพื่อวิเคราะห์หาหน้าที่และความสำคัญในกระบวนการทางชีวภาพของยีนในเครือข่ายย่อย (ภาพที่ 4-4) โดยอาศัยฐานข้อมูลของยีนออนโทโลยี (Gene Ontology; GO) (Ashburner *et al.*, 2000) และฐานข้อมูล KEGG pathway จะพบว่ายีนส่วนใหญ่จะเกี่ยวข้องกับวิถีการส่งสัญญาณ Jak-STAT signaling pathway เป็นวิถีที่ส่งสัญญาณจากสารเคมีที่อยู่นอกเซลล์ (extracellular chemical signals) เข้าไปยังภายในนิวเคลียสทำให้เกิดการถอดรหัสและเกิดการแสดงออกของยีนที่เกี่ยวข้องกับระบบภูมิคุ้มกัน การเพิ่มจำนวนเซลล์ (proliferation) และการตายของเซลล์ (apoptosis) ในการกระตุ้น cytokine receptor นั้นส่วนใหญ่จะผ่านวิถี Jak-STAT pathway โดย JAKs เป็น tyrosine kinases ที่จับอยู่กับ cytokine receptors และ STAT เป็น transcription factor ซึ่งการจับของไซโตไคน์กับ cytokine receptor จะส่งผลให้เกิด dimerization ของ cytokine receptors และเกิดการกระตุ้น JAKs ให้มีการเติมฟอสเฟตให้กับตัว JAKs และ receptor ทำให้ STATs สามารถมาจับกับ receptor และถูกเติมฟอสเฟตโดย JAKs ทำให้ STAT ที่ถูกเติมฟอสเฟตเกิด dimerization กลายเป็น STAT ที่อยู่ในรูปที่ทำงานได้ (active form) ของ transcription factor ที่ไปกระตุ้นการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองต่อไซโตไคน์นั้นๆ ดังภาพที่ 4-7 (Jatiani *et al.*, 2010) ดังนั้นจะเห็นได้ว่าวิถีการส่งสัญญาณ Jak-STAT นั้นมีความสำคัญ โดยยีนที่ได้จากการคัดเลือกที่เกี่ยวข้องกับวิถีนี้ประกอบด้วย 4 ยีน คือ *CSF2RB2*, *IL11*, *IL7R* และ *SOCS3* ซึ่งยีน *SOCS3* เป็นยีนที่มีความน่าสนใจที่นำไปตรวจสอบระดับการแสดงออกของยีนในห้องปฏิบัติการต่อไป เนื่องจากยีน *SOCS3* เป็นยีนที่ถูกคัดเลือกได้จากทั้งวิธีการทำเหมืองข้อมูลและการสืบค้นจากฐานข้อมูล PolySearch และมีรูปแบบการแสดงออกสูงในช่วงเวลา 8 และ 18 ชั่วโมง (ช่วงท้าย) และทำการตรวจสอบร่วมกับยีน *iNOS* เพื่อให้มั่นใจว่ามีการอักเสบเกิดขึ้นแน่นอน เนื่องจาก *iNOS* เป็นยีนที่ได้มีการศึกษาอย่างมากว่าเกี่ยวข้องกับการอักเสบ โดยใช้ยีน *EF2* เป็นยีนอ้างอิงหรือ housekeeping gene

จากผลการตรวจวัดปริมาณ mRNA ของยีน *iNOS* และ *SOCS3* โดยมียีน *EF2* เป็นยีนอ้างอิง (Housekeeping gene) โดยใช้เทคนิค real time PCR จะพบว่ายีน *iNOS* จะมีระดับการแสดงออกในสถานะที่มีการกระตุ้นด้วย LPS ที่ช่วงเวลา 3, 6, 8, และ 18 ชั่วโมงตามลำดับ และมีการแสดงออกลดลงเมื่อมีการกระตุ้นด้วย LPS ที่ช่วงเวลา 24 ชั่วโมง โดยในช่วงที่มีการกระตุ้นด้วย LPS ที่เวลา 24 ชั่วโมง จะมีระดับการแสดงออกที่สูงสุด (ภาพที่ 4-11) ซึ่งสอดคล้องกับงานวิจัยของทีมผู้วิจัยที่ใช้ยีน *iNOS* เป็นยีนที่แสดงออกถึงสถานะการตอบสนองต่อการอักเสบที่ช่วงเวลา 18 ชั่วโมง ของเซลล์แมคโครฟาจ RAW264.7 ที่ถูกกระตุ้นด้วย LPS ซึ่งจะใช้เป็นโมเดลในการศึกษาฤทธิ์ต้านการอักเสบของสารออกฤทธิ์ ตัวอย่างเช่น การศึกษาโดย Srisook และคณะ (2011) ที่ได้ทำการศึกษาศาสตร์สกัดเอทิลอะซิเตท (ethyl acetate extract) จากต้นเพชรสังฆาตที่แสดงให้เห็นถึงฤทธิ์การต่อต้านการอักเสบ (anti-inflammatory activity) ที่ผลการทดลองแสดงให้เห็นว่าสารสกัดจากพืชสมุนไพรนี้สามารถยับยั้งการผลิตไนตริกออกไซด์ในเซลล์แมคโครฟาจ ซึ่งสารสกัดจากพืชสมุนไพรนี้จะส่งผลทำให้การทำงานของเอนไซม์ *iNOS* ลดลง โดยลดการแสดงออกของยีน *iNOS* ในส่วนของยีน *SOCS3* ที่ถูกคัดเลือกมาจากเครือข่ายย่อยจะมีการแสดงออกที่ลดลงในสถานะที่มีการกระตุ้นด้วย LPS ที่ช่วงเวลา 3 และ 6 ชั่วโมงตามลำดับ จากนั้นเมื่อมีการกระตุ้นด้วย LPS ที่ช่วงเวลา 8 และ 18 ชั่วโมง จะมีระดับการแสดงออกเพิ่มขึ้น โดยช่วงเวลาที่มีการกระตุ้นด้วย LPS ที่ 18 ชั่วโมงจะมีระดับการแสดงออกสูงสุด และเมื่อมีการกระตุ้นด้วย LPS ที่ช่วงเวลา 24 ชั่วโมงจะมีระดับการแสดงออกที่ต่ำสุด (ภาพที่ 4-12) ซึ่งมีระดับการแสดงออกของยีน *SOCS3* ที่สอดคล้องกับงานวิจัยของ Qin และคณะ (2007) ซึ่งยีน *SOCS3* นี้เป็นยีนที่จะถูกแปลรหัสต่อไปได้เป็นโปรตีนในกลุ่ม Suppressor of cytokine signaling (SOCS) proteins ซึ่งมีบทบาทในการเป็นตัวควบคุมเชิงลบ (negative regulators) ของการตอบสนองในระบบภูมิคุ้มกันทั้งแบบที่ติดตัวมาแต่

กำเนิดหรืออินเนต (innate immunity) และแบบที่เกิดขึ้นมาภายหลังหรืออะแดปทีฟ (adaptive immunity) จึงมีความเกี่ยวข้องกับกลไกการตอบสนองต่อการอักเสบที่มักเกี่ยวข้องกับระบบภูมิคุ้มกัน

5.2 สรุปผลการทดลอง

1. จากการศึกษาและทดลองใช้กระบวนการทางชีวสารสนเทศหลากหลายชนิดเพื่อมาช่วยในการวิเคราะห์ข้อมูลไมโครอาร์เรย์การแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองต่อ LPS ในเซลล์แมคโครฟาจของหนู เพื่อให้ได้การศึกษาชีววิทยาเชิงระบบสำหรับสืบค้นกลไกการตอบสนองต่อการอักเสบ ทำให้ได้กระบวนการทางชีวสารสนเทศที่ประกอบไปด้วย การจัดเตรียมข้อมูลการแสดงออกของยีน (GEO: Edgar 2002) การคัดเลือกกลุ่มยีนที่มีบทบาทสำคัญ (ArrayMining: Glaab et al., 2009) การสร้างเครือข่ายปฏิสัมพันธ์ (Cladist: Fukuchi et al., 2009 และ ExpressionCorrelation Viewer: Cline et al., 2007) และเครือข่ายการควบคุมการแสดงออกของยีน (DREM: Schulz et al., 2012 และ CyNetworkBMA: Yeung et al., 2011) และการบูรณาการเครือข่ายกับข้อมูลทางชีวภาพ (DREM: Schulz et al., 2012; PolySearch: Cheng, 2008; STRING: Szklarczyk et al., 2015; Gene Ontology: Ashburner et al., 2000; และ KEGG: Kanehisa, 2010) ที่ทำให้ได้เครือข่ายย่อยซึ่งมีกลุ่มยีนที่มีบทบาทและความสำคัญที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบที่สามารถตรวจสอบผลในระดับห้องปฏิบัติการได้ จากตัวอย่างการตรวจสอบของยีน *SOCS3*
2. การวิจัยนี้เป็นต้นแบบของการศึกษาเพื่อพัฒนาแพลตฟอร์มของกลไกการตอบสนองต่อ LPS ในเซลล์แมคโครฟาจของหนู สำหรับนำมาเป็นโมเดลเพื่อใช้ในการศึกษาผลการออกฤทธิ์ของสารสกัดจากพืชสมุนไพร หรือสารสังเคราะห์ในการยับยั้งการอักเสบต่อไป

5.3 ข้อเสนอแนะ

1. เป็นโครงการวิจัยที่ต้องใช้เวลากับการคัดเลือกเครื่องมือและกระบวนการทางชีวสารสนเทศที่เหมาะสมเป็นเวลา และด้วยข้อจำกัดของพื้นฐานความรู้ของนิสิตที่เข้าร่วมการทำวิจัยที่ส่วนใหญ่เป็นนิสิตระดับปริญญาตรี ทำให้มีปัญหาในการทำความเข้าใจและการใช้งานเครื่องมือทางชีวสารสนเทศ ดังนั้นโครงการวิจัยแบบนี้จึงเหมาะสมกับนิสิตหรือนักศึกษาในระดับปริญญาโทและปริญญาเอกมากกว่า เพราะต้องการทักษะและความเชี่ยวชาญที่สูงกว่าระดับปริญญาตรี
2. ต้นทุนในการศึกษาวิจัยในห้องปฏิบัติการมีสูงมากจึงทำให้โครงการวิจัยนี้จำเป็นต้องลดเนื้องานลงให้เหลือเพียงการตรวจสอบในห้องปฏิบัติการในระดับ mRNA เท่านั้น จึงควรมีการทดลองในระดับโปรตีนต่อไป
3. โครงการวิจัยยังคงมีปัญหาในเรื่องของการนำเสนอรูปแบบของกลไกการอักเสบที่ได้ เนื่องจากยังคงนำเสนอในรูปแบบของเครือข่ายปฏิสัมพันธ์ ทำให้ยากต่อการแปลผลและการทำความเข้าใจ ดังนั้นถ้าหากสามารถนำเสนอกลไกการตอบสนองตามช่วงเวลาได้ในรูปแบบของภาพกราฟฟิก ก็จะทำให้ทำความเข้าใจและแปลผลกลไกการตอบสนองได้ง่ายยิ่งขึ้น

เอกสารอ้างอิง

- ปรีศนี นาคคำ มินตรา บัณฑิต และพิทักษ์ สุตรอนันต์. 2556. การคัดเลือกยีนที่เกี่ยวข้องกับการเกิดการอักเสบจากชุดข้อมูลไมโครอาร์เรย์ให้มีความน่าเชื่อถือมากยิ่งขึ้น. รายงานสืบเนื่องจากการประชุมทางวิชาการระดับชาติ วิทยาศาสตร์วิจัย ครั้งที่ 5, ระหว่างวันที่ 4–5 มีนาคม (2556), มหาวิทยาลัยพะเยา จังหวัดพะเยา, หน้า BIO-87–BIO-94.
- ปรีศนา โภคาพานิช. 2551. การต้านการอักเสบของพืชบางชนิดในวงศ์ขิง. วิทยานิพนธ์วิทยาศาสตรบัณฑิต ภาควิชาชีวเคมี คณะวิทยาศาสตร์.
- มัลลิกา ปาละโชติ. 2550. ฤทธิ์ในการต้านการอักเสบของส่วนสกัดเอทิลอะซิเตทและเมทานอลของต้นเพชรสังฆาตในเซลล์แมคโครฟาจ. วิทยานิพนธ์วิทยาศาสตรบัณฑิต ภาควิชาชีวเคมี คณะวิทยาศาสตร์.
- เหมือนฝัน โวหารกล้า กอบเกตุ วงษา และพิทักษ์ สุตรอนันต์. 2556. การใช้ข้อมูลไมโครอาร์เรย์และพาธเวย์เพื่อสำรวจกลไกการอักเสบในเซลล์แมคโครฟาจ. รายงานสืบเนื่องจากการประชุมทางวิชาการระดับชาติ วิทยาศาสตร์วิจัย ครั้งที่ 5, ระหว่างวันที่ 4–5 มีนาคม (2556), มหาวิทยาลัยพะเยา จังหวัดพะเยา, หน้า BIO-248–BIO-255.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Tarver LI, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. (2000). Gene Ontology: tool for the unification of biology. Nature Genetics. 25(1): 25–29.
- Barrett, T., and Edgar, R. 2006. Gene Expression Omnibus (GEO): Microarray data storage, submission, retrieval, and analysis. Methods in Enzymology, 411: 352–369, doi:10.1016/S0076-6879(06)11019-8.
- Basil, C.F., Zhao, Y., Zavaglia, K., Jin, P., Panelli, M.C., Voiculescu, S., Mandruzzato, S., Lee, H.M., Seliger, B., Freedman, R.S., Taylor, P.R., Hu, N., Zanovello, P., Marincola, F.M., and Wang, E., 2006. Common Cancer Biomarkers. Cancer Research, 66 (6): 2953–2961.
- Brown, P.O. and Botstein, D. 1999. Exploring the new world of the genome with DNA microarrays. Nature Genetics, 21: 33–37.
- Chantimal, J. 2012. Proteome analysis of machphage cells threated with ethyl 2-acetyl-3-(4-hydroxy-3-methoxyphenyl) acrylate. Master's thesis in Biological Science Program, Graduate School, Burapha University.
- Cheng, D., Knox, C., Young, N., Stothard, P., Damaraju, S. and Wishart D.S. 2008. PolySearch: a web-based text mining system for extracting relationships between human diseases, genes, mutations, drugs and metabolites. Nucleic Acids Research. 2008, Jul 1;36(Web Server issue):W399–405. doi: 10.1093/nar/gkn296. Epub 2008 May 16.

- Cline, M.S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., Christmas, R., Campilo, I.A., Creech, M., Gross, B., Hanspers, K., Isserlin, R., Kelley, R., Killcoyne, S., Lotia, S., Maere, S., Morris, J., Ono, K., Pavlovic, V., Pico, A.R., Vailaya, A., Wang, P.L., Adler, A., Conklin, B.R., Hood, L., Kuiper, M., Sander, C., Schmulevich, I., Schwikowski, B., Warner, G.J., Ideker, T., and Bader, G.D. 2007. Integration of biological networks and gene expression data using Cytoscape. Nature Protocols. 2(10): 2366-2382.
- Comer JE, Galindo CL, Zhang F, Wenglikowski AM, Bush KL, Garner HR, Peterson JW, Chopra AK. 2006. Murine macrophage transcriptional and functional responses to *Bacillus anthracis* edema toxin. Microbial Pathogenesis. 41(2-3) : 96–110.
- Edgar, R., Domrachev, M., Lash, A. E. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Research. 30: 207–210.
- Eleanor, A.H., Raktim, S., Daniel, S. and John, Q. (2011). RNA-Seq analysis in MeV. Bioinformatics, 27(22), 3209–3210.
- Ernst, J., Plasterer, H.L., Simon, I., and Joseph, Z.B. 2010. Integrating multiple evidence sources to predict transcription factor binding in the human genome. Genome Research. 20: 526–536.
- Fukuchi S, Homma K, Minezaki Y, Gojobori T, Nishikawa K. (2009). Development of an accurate classification system of proteins into structured and unstructured regions that uncovers novel structural domains its application to human transcription factors. BMC Structural Biology. 9 : 26–35.
- Ge, H., Walhout, A. J. M., and Vidal, M. 2003. Integrating ‘omic’ information: a bridge between genomics and systems biology. Trends in Genetics, 19 (10): 551–560.
- Gold, E., Ramsey, S., Sartain, M., Selinummi, J., Podolsky, I., Rodriguez, D. Moritz, R., Aderem, A. 2012. ATF3 protects against atherosclerosis by suppressing 25-hydroxycholesterol induced lipid body formation. Journal of Experimental Medicine, 209: 807–817.
- Gollub, J., Ball, C.A., Binkley, G., Demeter, J., Finkelstein, D.B., Hebert, J.M., Hernandez-Boussard, T., Jin, H.; Kaloper, M., Matese, J.C., Schroeder, M., Brown, P.O., Botstein, D., and Sherlock, G. 2003. The Stanford Microarray Database: data access and quality assessment tools. Nucleic Acids Research, 31 (1): 94–96.
- Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gassenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S., 1999. Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. Science, 286 (5439): 531–537.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Ian, H.W. (2009). The WEKA Data Mining Software : An Update. SIGKDD Explorations, 11(1), 10–18.

- Hall, M.A. 2000. Correlation-based feature selection for discrete and numeric class machine learning. Proceedings of the Seventeenth International Conference on Machine Learning, Stanford University, Stanford, United States of America.
- Hammer, K.D., Yum, M.Y., Dixon, P.M. and Birt, D.F. 2010. Identification of JAK-STAT pathways as important for the anti-inflammatory activity of a *Hypericum perforatum* fraction and bioactive constituents in RAW 264.7 mouse macrophages. Phytochemistry. 71(7) : 716–725.
- Hammer, M., Mages, J., Dietrich, H., Servatius, A., et al. 2006. Dual specificity phosphatase 1 (DUSP1) regulates a subset of LPS-induced genes and protects mice from lethal endotoxin shock. Journal of Experimental Medicine, 203 (1): 15–20.
- Hoffmann, J. A., and Reichhart, J. M. 2002. Drosophila innate immunity: an evolutionary perspective. Nature Immunology, 3: 121–126.
- Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW, Lane HC, and Lempicki RA. (2007). DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Research. 35: W169–W175.
- Ideker, T., Ozier, O., Schwikowski, B., and Siegel AF (2002). Discovering regulatory and signaling circuits in molecular interaction networks. Bioinformatics, 18 Suppl. 1 : S233–S240.
- Joyce, A., and Palsson, B. 2006. The model organism as a system: integrating ‘omics’ data sets. Molecular Cell Biology, 7: 198–210.
- Kitano, H., 2002. Systems Biology: A Brief Overview. Science, 296 (5560): 1662–1664.
- Kollipara, R.K., Perumal, N.B. 2010. Motif prediction to distinguish LPS-stimulated pro-inflammatory vs. antibacterial macrophage genes. Immunome Research, 6 (5), doi:10.1186/1745-7580-6-5.
- Lander, E. S., Linton, L.M., Nusbaum, C., et al. 2001. Initial sequencing and analysis of the human genome. Nature, 409: 860–921.
- Lipshutz, R. J., Morris, D., Chee, M., Hubbell, E., Kozal, M. J., Shah, N. Shen, N., Yang, R., Fodor, S. P. 1995. Using oligonucleotide probe arrays to access Genetic diversity. Biotechniques, 19: 442–447.
- Maere S, Heymans K, Kuiper M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics. 21(16): 3448–3449.

- Murray, R.Z., Wylie, F.G., Khromykh, T., Hume, D.A., et al. 2005. Syntaxin 6 and Vti1b form a novel SNARE complex, which is up-regulated in activated macrophages to facilitate exocytosis of tumor necrosis Factor- α . The Journal of Biological Chemistry, 280 (11): 10478–10483.
- Palachot, M. 2012. Investigation of anti-inflammatory activity and molecular mechanism of the action of selected zingiberaceae plant extracts. Master's thesis in Biological Science Program, Graduate School, Burapha University.
- Poltorak, A., He, X., Smirnova, I., Liu, M. Y., Van Huffel, C., Du, X., Birdwell, D., Alejos, E., Silva, M., Galanos, C., Freudenberg, M., Ricciardi-Castagnoli, P., Layton, B., Beutler, B. 1998. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. Science, 282: 2085–2088.
- Pusztai, L., and Hess, K.R. 2004. Clinical trial design for microarray predictive marker discovery and assessment. Annals of Oncology, 15: 1731–1737.
- Qin, H., Roberts, K.L., Niyongere, S.A., Cong, Y., Elson, C.O., and Benveniste, E.N. 2007. Molecular mechanism of lipopolysaccharide-induced SOCS-3 gene expression in macrophages and microglia. The Journal of Immunology. 179: 5966–5976.
- R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Rotter, A, Hren, M., Baebler, Š., Blejec, A., Gruden, K. 2008. Finding differentially expressed genes in two-channel DNA microarray datasets: How to increase reliability of data preprocessing. A Journal of Integrative Biology, 12 (3): 171–182.
- Sandvik, A.K., Alsberg, B.K., Nørsett, K.G., Yadetie, F., Waldum, H.L., and Lægreid, A., 2006. Gene expression analysis and clinical diagnosis. Clinica Chimica Acta, 363: 157–164.
- Schena, M., Shalon, D., Davis, R.W., Brown, P. O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science, 270: 467–470.
- Schott, J., Reitter, S., Philipp, J., Haneke, K., Schafer, H., and Stoecklin, G. 2014. Translational Regulation of Specific mRNAs Controls Feedback Inhibition and Survival during Macrophage Activation. PLoS Genetics. 10(6): e1004368
- Schulz MH, Devanny WE, Gitter A, Zhong S, Ernst J, Joseph ZB. (2012). DREM 2.0: Improved reconstruction of dynamic regulatory networks from time-series expression data. BMC Systems Biology. 6: 104.
- Shell, S.A., Hesse, C., Morris, S.M. Jr, Milcarek, C. 2005. Elevated levels of the 64-kDa cleavage stimulatory factor (CstF-64) in lipopolysaccharide-stimulated macrophages influence gene expression and induce alternative poly(A) site selection. The Journal of Biological Chemistry, 280 (48): 39950–39961.

- Smyth, G.K. 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Statistical Application in Genetics and Molecular Biology. (3) : 1027–1030.
- Srisook, K., Palachot, M., Mongkol, N., Srisook, E., and Sarapusit, S. 2001. Anti-inflammatory effect of ethyl acetate extract from *Cissus quadrangularis* Linn may be involved with induction of heme oxygenase-1 and suppression of NF- κ B activation. Journal of Ethnopharmacology, 133 (3): 1008–1014.
- Tan, S. L., Ganji, G., Paeper, B., and Katze, M. G. 2007. Systems biology and the host response to viral infection. Nature Biotechnology, 25: 1383–1389.
- Tusher, V.G., Tibshirani, R., Chu, G. 2001. Significance analysis of microarrays applied to the ionizing radiation response. Proceeding of the National Academy of Sciences of the United States of America. 98(9) : 5116–5121.
- Yap, Y. L., Zhang, X. W., Smith, D., Soong, R., and Hill, J., 2007. Molecular gene expression signature patterns for gastric cancer diagnosis. Computational Biology and Chemistry, 31: 275–287.
- Yeung KY, Dombek KM, Lo K, Mittler JE, Zhu J, Schadt EE, Bumgarner RE, Raftery AE. 2011. Construction of regulatory networks using expression time-series data of a genotyped population. Proceedings of the National Academy of Sciences. 48: 108.
- Zhang, X., Wei, D., Yap, Y., and Chen, F. 2007. Moving cancer diagnostics from bench to bedside. TREND in Biotechnology, 25 (4): 166–173, doi:10.1016/j.tibtech.2007.02.006.

ภาคผนวก

ภาคผนวก ก

นำเสนอรายงานสืบเนื่องจากงานประชุมทางวิชาการระดับชาติ

การประชุมวิชาการระดับชาติ “วิทยาศาสตร์วิจัย” ครั้งที่ 8

วันที่ 30 – 31 พฤษภาคม 2559 มหาวิทยาลัยพะเยา จังหวัดพะเยา ประเทศไทย

อรอนงค์ จันทุมมี ปราณิ พรหมคล้าย จิตภา สอนศิริ เหมือนฝัน ไหวหารกล้า และพิทักษ์ สุตรอนันต์. (2559). การคัดเลือกกลุ่มยีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบด้วยการจัดกลุ่มยีนและการวิเคราะห์เครือข่ายการแสดงออกของยีน. รายงานสืบเนื่องจากการประชุมวิชาการระดับชาติ “วิทยาศาสตร์วิจัย” ครั้งที่ 8, วันที่ 30 – 31 พฤษภาคม 2559, มหาวิทยาลัยพะเยา จังหวัดพะเยา, หน้า 241 – 246.

กิตติพันธ์ ธาระเขตร์ ฐิติรัตน์ นุชศิลา เหมือนฝัน ไหวหารกล้า และพิทักษ์ สุตรอนันต์. (2559). การค้นหายีนที่เกี่ยวข้องกับการตอบสนองการอักเสบจากข้อมูลไมโครอาร์เรย์โดยการคัดเลือกและการวิเคราะห์เครือข่ายการแสดงออกร่วม. รายงานสืบเนื่องจากการประชุมวิชาการระดับชาติ “วิทยาศาสตร์วิจัย” ครั้งที่ 8, วันที่ 30 – 31 พฤษภาคม 2559, มหาวิทยาลัยพะเยา จังหวัดพะเยา, หน้า 235 – 240.

**การคัดเลือกกลุ่มยีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบด้วยการจัดกลุ่มยีน
และการวิเคราะห์เครือข่ายการแสดงออกของยีน**
**Selection of inflammatory responsive genes by gene clustering
and gene expression network analysis**

อรอนงค์ จันทุมมี¹, ปราณี พรหมคล้าย¹, จิดาภา สอนศิริ², เหมือนฝัน ไหวหารกล้า², และ พิทักษ์ สุตรอนันต์^{1*}
Onanong Jantummee¹, Pranee Promklai¹, Jidapa Sornsiri², Muanfan Wohankla² and Pitak Sootanan^{1*}

¹ ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

² โปรแกรมวิทยาศาสตร์ชีวภาพ บัณฑิตศึกษา คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อตรวจสอบกลไกการตอบสนองต่อการอักเสบในเซลล์แมคโครฟาจ RAW 264.7 ที่ถูกกระตุ้นด้วย LPS จากข้อมูลการแสดงออกของยีนตามช่วงเวลา โดยอาศัยการคัดเลือกยีนที่เกี่ยวข้องกับการอักเสบจากการจัดกลุ่มยีน และการสร้างเครือข่ายการแสดงออกของยีน ผลที่ได้จากการจัดกลุ่มยีนโดยเทคนิค k-Means/Median clustering ในโปรแกรม MeV จะทำให้ได้กลุ่มยีนที่มีการแสดงออกแตกต่างกัน 5 แบบ ที่แสดงให้เห็นถึงกลไกการตอบสนองต่อการอักเสบอย่างเป็นลำดับขั้นภายในช่วงเวลาสองชั่วโมงแรก เมื่อนำข้อมูลมาคัดกรองร่วมกับเครือข่ายการแสดงออกของยีนที่สร้างโดย CyNetworkBMA ภายใต้โปรแกรม Cytoscape และการวิเคราะห์ร่วมกับข้อมูลทรานสคริปชันแฟกเตอร์ จะทำให้ได้กลุ่มยีนภายในเครือข่ายย่อยที่มีปฏิสัมพันธ์ของทรานสคริปชันแฟกเตอร์กับยีนเป้าหมายที่มีหน้าที่เกี่ยวกับกิจกรรมของโปรตีนตัวรับ และการเข้าจับของโปรตีน ซึ่งยีนกลุ่มนี้มีบทบาทสำคัญในกระบวนการควบคุมการตอบสนองต่อการอักเสบของระบบภูมิคุ้มกันในช่วงเวลาตั้งแต่หนึ่งชั่วโมงเป็นต้นไป

คำสำคัญ : การคัดเลือกยีน / เครือข่ายการแสดงออกของยีน / การจัดกลุ่มยีน / การอักเสบ

Abstract

The objective of this research aimed to investigate the mechanisms of the inflammatory response in macrophages RAW264.7 cells stimulated with LPS by gene selection which is integration of gene clustering and gene expression network. The results from gene clustering with k-Means/Median Clustering in MeV program demonstrated five different profiles of gene expression. They showed the mechanism steps of inflammation within two hours. After analyzing the previous results with gene expression network reconstructed by CyNetworkBMA in Cytoscape and transcription factor data, interested group of genes were found. These results showed sub-network of transcription factor and gene targets which had the functions of protein receptor activity and protein binding. These genes were important in the role of regulation of inflammatory response process of immune system since one hour.

Keywords : Gene selection / Gene expression network / Gene clustering / Inflammation

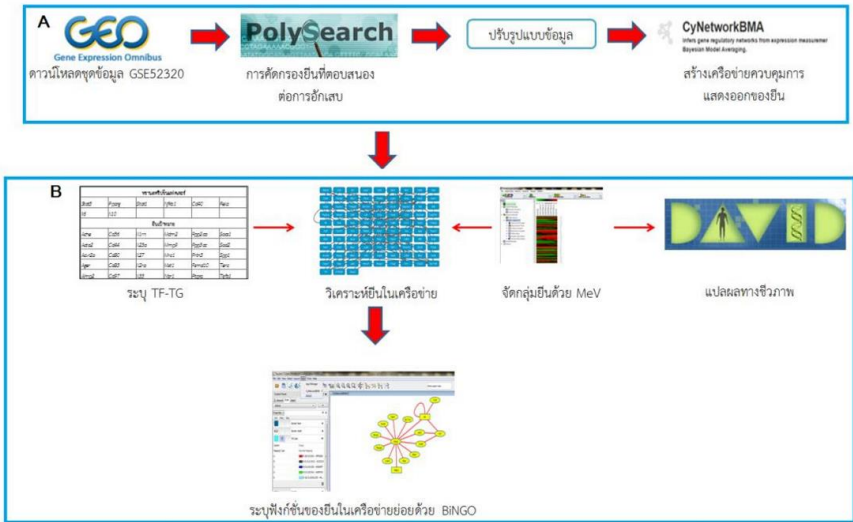
*Corresponding author. E-mail : pitak@buu.ac.th

1. บทนำ

การอักเสบ (Inflammation) เป็นกลไกป้องกันสิ่งแปลกปลอมหรือสิ่งกระตุ้นที่ทำให้เซลล์หรือเนื้อเยื่อถูกทำลาย ทำให้เกิดการกระตุ้นเซลล์ในระบบภูมิคุ้มกัน ได้แก่ เซลล์แมคโครฟาจ ให้มีการตอบสนองต่อการอักเสบโดยการหลั่งสารสื่อกลางการอักเสบต่างๆ (pro-inflammatory cytokine) เช่น ไนตริกออกไซด์ (nitric oxide, NO) และพรอสตาแกลนดิน (prostaglandins, PGs) (Kwon et al, 2013) ในการศึกษาการตอบสนองต่อการอักเสบต้องอาศัยการศึกษาชีววิทยาเชิงระบบ โดยการเรียนรู้ข้อมูลการแสดงออกของยีนทั้งหมดที่มีในสิ่งมีชีวิตเป้าหมาย จากการศึกษาวิเคราะห์การแสดงออกของยีนที่ถูกรวบรวมไว้ในฐานข้อมูลสาธารณะ GEO (Edgar, Domrachev and Lash, 2002) ซึ่งเป็นฐานข้อมูลสาธารณะที่สามารถนำข้อมูลมาใช้เพื่อการศึกษาและวิจัย โดยมีผู้ศึกษาการตอบสนองต่อการอักเสบโดยใช้ชุดข้อมูลการแสดงออกของยีนในเซลล์แมคโครฟาจ RAW 264.7 ที่ถูกกระตุ้นด้วยไลโปพอลิแซคคาไรด์ (LPS) มาใช้วิเคราะห์ปฏิสัมพันธ์ของทรานสคริปชันแฟกเตอร์กับยีนเป้าหมาย (TF-TG interaction) และสร้างเครือข่ายการแสดงออกของยีนด้วยโปรแกรม Cytoscape (Cline et al., 2007) เพื่อใช้ในการอธิบายกลไกการแสดงออกของยีนที่ตอบสนองต่อการอักเสบ (พศิกา, วิไลพร และ พิทักษ์, 2557) และยังมีโปรแกรมอื่นอีกมากมายที่ใช้ในการสร้างเครือข่ายการแสดงออกของยีน เช่น โปรแกรม CyNetworkBMA (Young, Raftery and Yeung, 2014) ซึ่งเป็นโปรแกรมที่มีอยู่ภายใต้ Cytoscape (Cline et al., 2007) ที่มีการทำงานร่วมกับโปรแกรม R ซึ่งโปรแกรมจะคัดเลือกยีนที่มีความสัมพันธ์ของการแสดงออกมาสร้างเป็นเครือข่ายการแสดงออกของยีน ในการศึกษาที่ใช้ข้อมูลการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบของเซลล์แมคโครฟาจ Raw 264.7 ที่ถูกกระตุ้นด้วย LPS ในชุดข้อมูล GSE52320 (Schott et al., 2014) มาทำการคัดกรองรายชื่อยีนที่เกี่ยวข้องกับการอักเสบ และสร้างเครือข่ายการแสดงออกของยีนด้วยโปรแกรม CyNetworkBMA (Young, Raftery and Yeung, 2014) ซึ่งจะได้เครือข่ายเพื่อใช้ในการอธิบายกลไกการตอบสนองการอักเสบที่เกิดขึ้นตามช่วงเวลาการศึกษา ก่อนระบุชนิดของยีนว่าเป็นทรานสคริปชันแฟกเตอร์หรือยีนเป้าหมาย จัดกลุ่มยีนตามรูปแบบการแสดงออกและแปลผลทางชีวภาพในแต่ละกลุ่ม เลือกเครือข่ายย่อยจากเครือข่ายการแสดงออกของยีนโดยอาศัยข้อมูลการจัดกลุ่มยีนที่น่าสนใจมาแปลผลทางชีวภาพ เพื่ออธิบายกลไกการตอบสนองการอักเสบที่เกิดขึ้นและหน้าที่การทำงานของยีนในกลุ่มที่สนใจ ซึ่งทำให้สามารถเรียนรู้และทำความเข้าใจกลไกการอักเสบในกระบวนการทางชีวภาพได้ดียิ่งขึ้น

2. วิธีการ

การสร้างเครือข่ายการแสดงออกของยีนเพื่ออธิบายกลไกการตอบสนองการอักเสบจากข้อมูลการแสดงออกของยีน ในเซลล์แมคโครฟาจ RAW 264.7 ที่ถูกเหนี่ยวนำด้วย LPS ประกอบด้วยขั้นตอนดังแสดงในภาพที่ 1



ภาพที่ 1 ภาพรวมการทดลอง

โดยวิธีการจะเริ่มต้นจากการสืบค้นข้อมูลการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบฐานข้อมูลสาธารณะ GEO (Gene Expression Omnibus, <http://www.ncbi.nlm.nih.gov/geo/>, Edgar, Domrachev and Lash, 2002) ได้แก่ ชุดข้อมูล GSE52320 (Schott et al., 2014) ซึ่งเป็นข้อมูลการแสดงออกของยีนที่เวลา 0, 15, 30, 45, 60, 75, 90 และ 120 นาที ในเซลล์แมคโครฟาจ RAW 264.7 ของหนูที่ถูกกระตุ้นด้วย LPS นำมาคัดกรองยีนจากการสืบค้นด้วยรายชื่อยีนที่เกี่ยวข้องกับการอักเสบจากฐานข้อมูลสาธารณะ PolySearch (Cheng et al, 2008) เพื่อให้การวิเคราะห์มีความครอบคลุมเฉพาะกลุ่มยีนที่มีการตอบสนองต่อการอักเสบ ก่อนนำมาสร้างเครือข่ายการแสดงออกของยีนในโปรแกรม CyNetworkBMA (Young, Raftery and Yeung, 2014) ที่บรรจุในโปรแกรม Cytoscape (Cline et al., 2007) ซึ่ง CyNetworkBMA จะอาศัยการทำงานผ่านแพ็คเกจในโปรแกรม R (Young, Raftery and Yeung, 2014) โดยโปรแกรมจะคัดเลือกเฉพาะยีนที่มีการแสดงออกสัมพันธ์กันมาสร้างเป็นเครือข่ายการแสดงออกของยีน (ภาพที่ 1A)

จากเครือข่ายการแสดงออกของยีนนำมาระบุชนิดของยีนว่าเป็นทรานสคริปชันแฟคเตอร์หรือยีนเป้าหมายโดยเทียบกับข้อมูลทรานสคริปชันแฟคเตอร์จากโปรแกรม DREM 2.0 (Schulz et al., 2012) และจัดกลุ่มตามรูปแบบการแสดงออกของยีนโดยใช้ k-Means/Medians Clustering ในโปรแกรม MeV 4.9.0 (Saeed et al., 2003) และนำกลุ่มยีนที่ถูกจัดกลุ่มมาแปลผลทางชีวภาพด้วยโปรแกรม DAVID (<http://david.abcc.ncifcrf.gov/>; Huang, Sherman and Lempicki, 2008) ข้อมูลทรานสคริปชันและยีนเป้าหมาย และข้อมูลการจัดกลุ่มยีนตามรูปแบบการแสดงออกของยีนถูกนำมาใช้ในการวิเคราะห์ร่วมกับเครือข่ายในโปรแกรม Cytoscape (Cline et al., 2007) เพื่อพิจารณาความสัมพันธ์ของยีนและคัดเลือกเครือข่ายย่อยที่มีรูปแบบการแสดงออกที่น่าสนใจโดยดูจากรูปแบบการแสดงออกของทรานสคริปชันแฟคเตอร์และยีนที่มีปฏิสัมพันธ์ ก่อนนำมาแปลผลทางชีวภาพ และค้นหาหน้าที่ทางชีวภาพของยีนในเครือข่ายย่อยโดยใช้โปรแกรม BiNGO (Maere, Heymans and Kuiper, 2005) ที่บรรจุอยู่ในโปรแกรม Cytoscape (Cline et al., 2007) (ภาพที่ 1B)

3. ผลและอภิปราย

ผลจากการสร้างเครือข่ายการแสดงออกของยีนเพื่อศึกษากลไกการตอบสนองต่อการอักเสบของยีน จากข้อมูลการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบในเซลล์แมคโครฟาจ RAW 264.7 เมื่อถูกกระตุ้นด้วย LPS ประกอบด้วยยีนจำนวน 23,143 ยีน และถูกคัดกรองโดยรายชื่อยีนที่เกี่ยวข้องกับการอักเสบทำให้เหลือ 314 ยีน และตัดยีนที่มีการแสดงออกน้อยออกทำให้เหลือ 179 ยีน สร้างเครือข่ายการแสดงออกของยีนและระบุชนิดของยีนจะพบว่า มีทรานสคริปชันแฟคเตอร์ 8 ยีน ยีนเป้าหมาย 85 ยีน ซึ่งมี 112 ปฏิสัมพันธ์ ซึ่งจำนวนยีนที่ผ่านการคัดกรองและการวิเคราะห์ในขั้นตอนต่างๆ แสดงในตารางที่ 1 เมื่อวิเคราะห์โครงสร้างเครือข่ายพบว่า มีจำนวน 3 ยีนที่ถูกตัดออกเนื่องจากไม่มีปฏิสัมพันธ์กับเครือข่ายใหญ่ ทำให้ได้เครือข่ายที่มี 81 ยีนเป้าหมาย ที่ถูกควบคุมโดยทรานสคริปชันแฟคเตอร์ 8 ยีน ได้แก่ *Stat3, Pparg, Stat1, Nfkb1, Cd40, Rela, Il6* และ *Il10* แสดงดังภาพที่ 2 เมื่อได้เครือข่ายการแสดงออกของยีน นำยีนทั้งหมดในเครือข่ายมาจัดกลุ่มตามรูปแบบการแสดงออกของยีนตามช่วงเวลาด้วย k-Means/Medians Clustering และแปลผลทางชีวภาพ พบว่าสามารถจัดกลุ่มยีนตามรูปแบบการแสดงออกได้ 5 กลุ่ม (ตารางที่ 2) ที่แสดงให้เห็นถึงกลไกการตอบสนองการอักเสบอย่างเป็นลำดับขั้นภายในช่วงเวลาสองชั่วโมงแรก ในการแปลผลทางชีวภาพพบว่ายีนส่วนใหญ่จะเกี่ยวข้องกับกับปฏิสัมพันธ์ของตัวรับไซโตไคน์-ไซโตไคน์ (Cytokine-cytokine receptor interaction) เช่น ยีน *Il10, Il6, Il1b, Tnf, Ccl2, Ccl5* เป็นต้น และเกี่ยวข้องกับวิถี JAK-STAT pathway เช่น ยีน *Jak2, Ilfng1, Il10, Il6, Stat1, Stat3, Socs1* เป็นต้น เนื่องจากการกระตุ้น cytokine receptors จะผ่านวิถีการส่งสัญญาณ JAK-STAT pathway โดย JAKs จะจับอยู่กับ cytokine receptors และ STAT ซึ่งเป็น transcription factor ที่จะไปกระตุ้นการแสดงออกของยีนที่เกี่ยวข้องกับไซโตไคน์นั้น เช่น ยีน *Il6* และ *Il10* จะกระตุ้น *Jak1* และ *Stat3* (Niemand et al., 2003)

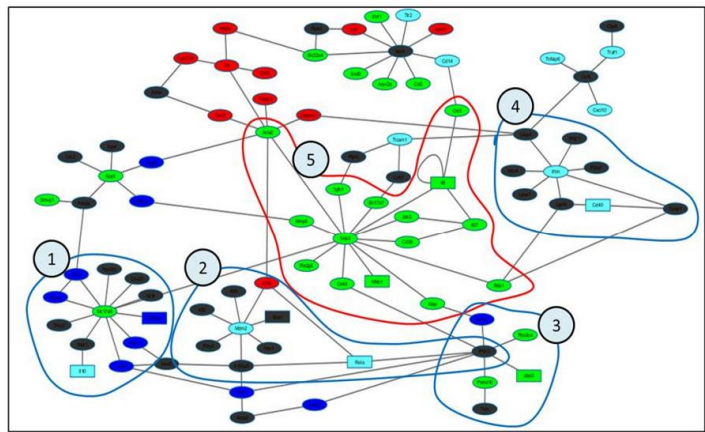
จากข้อมูลการระบุทรานสคริปชันแฟคเตอร์และยีนเป้าหมาย และการจัดกลุ่มยีน ที่นำข้อมูลมาวิเคราะห์ร่วมกับเครือข่ายได้เป็นเครือข่ายปฏิสัมพันธ์ของทรานสคริปชันแฟคเตอร์แต่ละตัวกับยีนเป้าหมายทั้งหมด และนำข้อมูลการจัดกลุ่มมาใช้ในการแปลผลทางชีวภาพซึ่งแบ่งสี่ตามกลุ่มการจัดรูปแบบการแสดงออกโดยสี่ค่าคือกลุ่ม A ระดับการแสดงออกของยีนไม่มีการเปลี่ยนแปลง สี่ห้าคือกลุ่ม B การแสดงออกของยีนเปลี่ยนแปลงขึ้นลงเล็กน้อย สี่แดงคือกลุ่ม C มีระดับการแสดงออกของยีนเพิ่มขึ้นตั้งแต่ช่วง 30 นาทีแรก สี่ฟ้าคือกลุ่ม D มีการแสดงออกเพิ่มขึ้นตั้งแต่ 45 นาทีขึ้นไป และสี่เขียวคือ E มีการแสดงออกเพิ่มขึ้นที่ช่วงหลังจากการกระตุ้นผ่าน 1 ชั่วโมงไปแล้ว จะเห็นได้ว่ากลุ่ม A และกลุ่ม B ยังไม่ถูกกระตุ้นในช่วง 2 ชั่วโมงแรก ยีนกลุ่มแรกที่ถูกกระตุ้นให้มีการแสดงออกเพิ่มขึ้นคือ กลุ่ม C ต่อมาคือ กลุ่ม D และ E ตามลำดับ ซึ่งยีนแต่ละกลุ่มมีปฏิสัมพันธ์กันดังภาพที่ 2

ตารางที่ 1 จำนวนทรานสคริปชันแฟกเตอร์ ยีนเป้าหมาย และจำนวนปฏิสัมพันธ์ที่ผ่านการคัดกรองและการวิเคราะห์ในขั้นตอนต่างๆ

การคัดกรองและการวิเคราะห์	จำนวน TF (ยีน)	จำนวน TG (ยีน)	จำนวน Interaction (Edges)
ชุดข้อมูล GSE 52320	23,143		267,799,225
การคัดกรองยีนด้วย PolySearch	314		49,298
การตัดยีนที่มีการแสดงออกน้อย	179		16,021
วิเคราะห์ด้วย CyNetworkBMA	8	84	112
วิเคราะห์โครงสร้างเครือข่าย	8	81	110
การเลือกเครือข่ายย่อย	2	13	17

คัดเลือกยีนและทรานสคริปชันแฟกเตอร์ที่มีรูปแบบการแสดงออกที่น่าสนใจ เพื่อคัดเลือกเครือข่ายย่อยมาใช้อธิบายกลไกการแสดงออกของยีนที่มีการกระตุ้นการอักเสบ โดยกลุ่มยีนที่สนใจ ได้แก่ ยีนกลุ่มที่ 5 ที่ควบคุมด้วยทรานสคริปชันแฟกเตอร์สีเขียวที่มีรูปแบบการแสดงออกของยีนที่มีปฏิสัมพันธ์โดยตรงและยีนที่มีปฏิสัมพันธ์ร่วมในรูปแบบเดียวกันคือมีการแสดงออกของยีนเพิ่มขึ้นในช่วงเวลา 90 นาทีแรกของการกระตุ้นการอักเสบ โดยทรานสคริปชันแฟกเตอร์ที่เลือกศึกษาได้แก่ *I/6* และ *Nfkb1* ซึ่งควบคุมยีนตัวเดียวกันคือ *Nos2* และยีนที่มีปฏิสัมพันธ์กับ *Nos2* มีรูปแบบการแสดงออกในรูปแบบเดียวกัน ซึ่งก็คือกลุ่ม E และ *I/6* ยังควบคุมการแสดงออกของยีนอีก 2 ยีน คือ *Ccl5* และ *I/27* ได้เป็นเครือข่ายย่อยที่มี 2 ทรานสคริปชันแฟกเตอร์กับ 13 ยีนเป้าหมาย โดยแสดงเครือข่ายย่อยที่คัดเลือกจากรูปแบบการแสดงออกที่สนใจดังภาพที่ 2

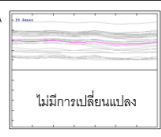
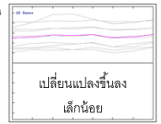
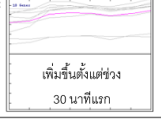
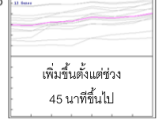
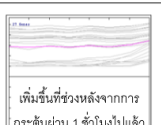
โปรตีน *Nfkb1* เป็นทรานสคริปชันแฟกเตอร์ที่ควบคุมยีน *Nos2* เมื่อมีการกระตุ้นจากสิ่งเร้าต่างๆทั้งจากภายในและภายนอก *Nfkb* จะเข้าไปในนิวเคลียสและกระตุ้นการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบ และ *I/6* เป็นทรานสคริปชันแฟกเตอร์ที่ควบคุมการแสดงออกของยีน *Nos2*, *Ccl5* และ *I/27* ซึ่งยีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบ โดย *I/6* ทำหน้าที่เป็นไซโตไคน์ซึ่งจะหลั่งจาก T cell และแมคโครฟาจที่กระตุ้นให้เกิดตอบสนองต่อการอักเสบ นอกจากนี้ *Nos2* จะมีปฏิสัมพันธ์กับทรานสคริปชันแฟกเตอร์ทั้ง 2 ชนิดแล้วยังเป็นยีนที่มีปฏิสัมพันธ์กับยีนอื่นๆ อีกหลายยีน ซึ่งการแสดงออกของยีนเหล่านั้นมีรูปแบบการแสดงออกเหมือนกับ *Nos2* เนื่องจากมีปฏิสัมพันธ์ที่ทอ้มกับทรานสคริปชันแฟกเตอร์ทั้งสองโดยมี *Nos2* เป็นศูนย์กลาง (hub)



ภาพ 2 เครือข่ายแสดงปฏิสัมพันธ์ของทรานสคริปชันแฟกเตอร์กับยีนเป้าหมายโดยแบ่งสีตามกลุ่มการจัดรูปแบบการแสดงออก โดยสีค่าน้ำเงิน แดง ฟ้ำ และเขียว แทนกลุ่ม Custer A, B, C, D และ E ตามลำดับ

จากการแปลผลทางชีวภาพของยีนในเครือข่ายย่อยที่ทำการศึกษา พบว่ายีนส่วนใหญ่มีหน้าที่เกี่ยวกับกิจกรรมของโปรตีนตัวรับ และการเข้าจับของโปรตีน โดยมีการทำงานทั้งภายนอกและภายในเซลล์ ยีนกลุ่มนี้มีบทบาทสำคัญในกระบวนการควบคุมการตอบสนองต่อการอักเสบของระบบภูมิคุ้มกันในช่วงเวลาตั้งแต่หนึ่งชั่วโมงเป็นต้นไป ดังนั้นการสร้างเครือข่ายการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบจากข้อมูลการแสดงออกของยีน และการใช้ข้อมูลที่เกี่ยวข้อง ได้แก่ ข้อมูลยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบ ข้อมูลการจัดกลุ่มยีน และข้อมูลทรานสคริปชันแฟคเตอร์ สามารถนำมาใช้ในการคัดเลือกเครือข่ายย่อยที่สามารถนำมาช่วยในการอธิบายและค้นหากลุ่มยีนที่มีความสำคัญต่อการแสดงออกในระดับ mRNA ของยีนที่มีการตอบสนองต่อการอักเสบ

ตารางที่ 2 ผลทางชีวภาพของกลุ่มการแสดงออก A, B, C, D และ E ตามลำดับ

รูปแบบการแสดงออกของยีน	คำอธิบาย	ยีน
A  ไม่มีการเปลี่ยนแปลง	โรคหลอดเลือดหัวใจหรือกล้ามเนื้ออ่อนแอของหลอดเลือด วิธีการส่งสัญญาณของแคลเซียม วิธีการส่งสัญญาณของคีโมไคน์ วิธีการส่งสัญญาณ Jak-STAT ปฏิสัมพันธ์ของตัวรับไซโตไคน์-ไซโตไคน์	Stat1, Lace1, Sirt4, Aimp2, Cbx8, Cd97, Ccr5, Nbr1, Tnip2, Casp1, Ccr2, Cox20, Ilngf1, Klf2, Lgmn, Mst1, Nos1, P2rx7, Il6ra, Klf9, Ltbp4, Nlrc4, Nos3, P2ry6, Pde6a, Ptn3, Ripk3, Terc, Ager, Ache, Ppp3cc, Ptpro, S100a8
B  เปลี่ยนแปลงจึ้นลงเล็กน้อย	โรคอัลไซเมอร์ วิถี Jak-STAT โปรตีนไกลแคนในโรคมะเร็ง ไมโครอาร์เอ็นเอในโรคมะเร็ง ปฏิสัมพันธ์ของตัวรับไซโตไคน์-ไซโตไคน์ การควบคุมที่ผิดปกติของการถอดรหัสในโรคมะเร็ง วิธีการส่งสัญญาณ HIF-1 การเสกติดแอมเฟตามีน วิธีการส่งสัญญาณ FoxO	Socs1, Vegfa, Zfp36, Pparg, Apoe, Hmox1, Il2ra, Ly96, Mrc1, Sirt1
C  เพิ่มขึ้นตั้งแต่ช่วง 30 นาทีแรก	วิธีการส่งสัญญาณ NOD-like receptor ปฏิสัมพันธ์ของตัวรับไซโตไคน์-ไซโตไคน์ วิธีการส่งสัญญาณอะดิโพไซโตไคน์ วิธีการส่งสัญญาณ RIG-I-like receptor	Cd83, Cxcl2, Cyp27b1, Tnf, Icam1, Il23a, Tnfaip3, Nfkb1a, Serpine1
D  เพิ่มขึ้นตั้งแต่ช่วง 45 นาทีขึ้นไป	วิธีการส่งสัญญาณ Toll-like receptor ร่างกายปฏิเสธร้อยวะที่นำมาปลูกถ่ายให้ โรคไตเรื้อรังอักเสบแบบภูมิคุ้มกันตนเอง เครือข่ายของระบบภูมิคุ้มกันในลำไส้สำหรับการผลิต IgA วิธีการตรวจจับ Cytosolic DNA ปฏิสัมพันธ์ของตัวรับไซโตไคน์-ไซโตไคน์ โรคหลอดเลือดหัวใจ โรคหัวใจ ภาวะสเต็มเซลล์ในลำไส้ร่างกายผู้ป่วย เมาหวานชนิด 1 วิธีการส่งสัญญาณ MAPK	Cd40, Il1b, Cd14, Cxcl10, Ticam1, Tnfaip6, Il10, Tlr2, Cd80, Il1rn, Mdm2, Traf1, Rela
E  เพิ่มขึ้นที่ช่วงหลังจากกระตุ้นผ่าน 1 ชั่วโมงไปแล้ว	วิธีการส่งสัญญาณ NOD-like receptor วิถีในโรคมะเร็ง วิธีการตรวจจับ Cytosolic DNA วิธีการส่งสัญญาณอะดิโพไซโตไคน์ วิธีการส่งสัญญาณของคีโมไคน์ วิธีการส่งสัญญาณ Toll-like receptor ปฏิสัมพันธ์ของตัวรับไซโตไคน์-ไซโตไคน์ มะเร็งตับอ่อน ปฏิสัมพันธ์ของตัวรับ ECM มะเร็งปอดชนิดเซลล์เล็ก การเจริญเติบโตของเซลล์เม็ดเลือด วิธีการส่งสัญญาณ TGF-beta	Il6, Cd44, Il27, Pla2g6, Tgfb1, Acta2, Ccl2, Mmp9, Slc17a7, Spp1, Acvr2a, Il33, Il1r1, Nos2, Smug1, Xiap, Bdp1, Ppp2ca, Jak2, Slc17a6, Sod2, Nfkb1, Cd36, Psmc10, Stat3, Ccl5, Slc22a4

4. บทสรุป

จากการสร้างเครือข่ายควบคุมการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบทำให้ได้เครือข่ายที่มี 93 ยีน 112 ปฏิสัมพันธ์ ที่ประกอบไปด้วยทรานสคริปชันแฟคเตอร์ 8 ยีน คือ *Stat3*, *Pparg*, *Stat1*, *Nfkb1*, *Cd40*, *Rela*, *Il6* และ *Il10* และยีนเป้าหมาย 85 ยีน จากการจัดกลุ่มของยีนตามรูปแบบการแสดงออกเป็น 5 กลุ่ม จะเห็นว่ายีนแต่ละกลุ่มมีการตอบสนองต่อการกระตุ้นที่เวลาต่างกัน และในเครือข่ายย่อยที่เลือกศึกษาจะพบทรานสคริปชันแฟคเตอร์ ได้แก่ *Il6* และ *Nfkb1* ซึ่งควบคุมยีนตัวเดียวกันคือ *Nos2* และยีนที่มีปฏิสัมพันธ์กับ *Nos2* และ *Il6* ยังควบคุมการแสดงออกของยีนอีก 2 ยีน คือ *Ccl5* และ *Il27* กับ 13 ยีนเป้าหมาย โดยมีรูปแบบการแสดงออกในรูปแบบเดียวกัน ซึ่งก็คือกลุ่ม E ในการระบุฟังก์ชันพบว่ายีนกลุ่มนี้มีบทบาทสำคัญในกระบวนการควบคุมการตอบสนองต่อการอักเสบของระบบภูมิคุ้มกันในช่วงเวลาตั้งแต่ 1 ชั่วโมงเป็นต้นไป นอกจากนี้การสร้างเครือข่ายยังเป็นตัวช่วยในการคัดเลือกเครือข่ายย่อยสำหรับนำมาศึกษาการตอบสนองของยีนต่อการกระตุ้นการอักเสบในระดับต่อไป

5. กิตติกรรมประกาศ

ขอขอบพระคุณคณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา ที่ให้ทุนสนับสนุนการนำเสนอผลงานในงานวิจัย และงานวิจัยที่ได้รับทุนอุดหนุนการวิจัย มหาวิทยาลัยบูรพา ประจำปีงบประมาณ พ.ศ.2557 สัญญาเลขที่ 87/2557

6. เอกสารอ้างอิง

- พดิกา ไบยา, วิไลพร ศรีตะบุตร และ พิทักษ์ สุตรอนันต์. (2557). เครื่องช่วยควบคุมการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบจากข้อมูลไมโครอาร์เรย์. *วารสารวิทยาศาสตร์บูรพา ฉบับพิเศษ การประชุมวิชาการระดับชาติ วิทยาศาสตร์วิจัย ครั้งที่ 6*, 173 – 184.
- Cheng, D., Knox, C., Young, N., Stothard, P., Damaraju, S. and Wishart, D.S. (2008). PolySearch: a web-based text mining system for extracting relationships between human diseases, genes, mutation, drugs and metabolites. *Nucleic Acids Research*, 36, W399 – 405.
- Cline, M.S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman C, et al. (2007). Integration of biological networks and gene expression data using Cytoscape. *Nature Protocols*, 2(10), 2366 – 2382.
- Edgar, R., Domrachev, M. and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, 30(1), 207 – 210.
- Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4, 44 – 57.
- Kwon, D.J., Ju, S.M, Youn, G.S., Choi, S.Y. and Park, j. (2013). Suppression of iNOS and COX-2 expression by flavokawain A via blockade of NF- κ B and AP-1 activation in RAW 264.7 macrophages. *Food and Chemical Toxicology*, 58, 479 – 486.
- Maere, S., Heymans, K. and Kuiper, M. (2015). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, 21(16), 3448 – 3449.
- Niemand, C., Nimmegern, A., Haan, S., Fischer, P., Schaper, F., Rossaint, R., et al. (2003). Activation of STAT3 by IL-6 and IL-10 in Primary Human Macrophages Is Differentially Modulated by Suppressor of Cytokine Signaling 3. *The Journal of Immunology*, 170(6), 3263 – 3272.
- Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., et al. (2003). TM4: a free, open-source system for microarray data management and analysis. *Biotechniques*, 34(2), 374 – 378.
- Schott, J., Reitter, S., Philipp, J., Haneke, K., Schafer H. and Stoecklin, G. (2014). Translational Regulation of Specific mRNAs Controls Feedback Inhibition and Survival during Macrophage Activation. *PLoS Genetics*, 10(6), e1004368.
- Schulz, M.H., Devanny, W.E, Gitter, A., Zhong, S., Ernst, J. and Joseph, Z.B. (2012). DREM 2.0: Improved reconstruction of dynamic regulatory networks from time-series expression data. *BMC Systems Biology*, 6, 104.
- Young, W.C., Raftery, A.E. and Yeung, K.Y. (2014). Fast Bayesian inference for gene regulatory networks using ScanBMA. *BMC Systems Biology*, 8, 47.

**การค้นหายีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบจากข้อมูลไมโครอาร์เรย์
โดยการคัดเลือกยีนและการวิเคราะห์เครือข่ายการแสดงออกร่วม**
**Inflammatory responsive gene finding from microarray data by gene selection
and co-expression network analysis**

กิตติพันธ์ ธาระเขตร์¹, ฐิติรัตน์ นุชศิลา², เหมือนฝัน โวหารกล้า² และ พิทักษ์ สุตอรันันต์¹

Kittipan Tharakhet¹, Thitirat Nuchsilā², Muanfan Wohankla² and Pitak Sootanan¹

¹ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

²โปรแกรมวิทยาศาสตร์สุขภาพ บัณฑิตศึกษา คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

บทคัดย่อ

การค้นหายีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบ มีส่วนสำคัญต่อการนำมาใช้ในการอธิบายกลไกระดับโมเลกุลของการอักเสบที่เกี่ยวข้อง โดยสามารถค้นหายีนผ่านการวิเคราะห์ข้อมูลไมโครอาร์เรย์ ในงานวิจัยนี้เลือกใช้ข้อมูลไมโครอาร์เรย์จำนวน 3 ชุดข้อมูล ของเซลล์แมคโครฟาจ RAW 264.7 ที่มีช่วงเวลากการบ่มด้วย LPS แตกต่างกัน 4 ช่วงเวลา ทำการสร้างเครือข่ายการแสดงออกของแต่ละชุดข้อมูลด้วย Network analysis ของ โปรแกรมออนไลน์ ArrayMining และแสดงเครือข่ายผ่านโปรแกรม Cytoscape ก่อนทำการคัดเลือกเครือข่ายย่อยด้วยยีนที่คัดเลือกได้จากโปรแกรม Gene selection ของ ArrayMining แล้วทำการตรวจสอบหน้าที่ของกลุ่มยีนที่ถูกคัดเลือก พบว่ายีนทั้งหมดไม่เคยตรวจพบมาก่อนว่าเกี่ยวข้องกับการอักเสบเมื่อเปรียบเทียบกับยีนที่ได้ฐานข้อมูล PolySearch จึงสามารถสรุปเบื้องต้นได้ว่าอาจเป็นกลุ่มยีนใหม่ที่เกี่ยวข้องกับการตอบสนองการอักเสบ แต่จำเป็นต้องทำการพิสูจน์ในการศึกษาคั้งต่อไป

คำสำคัญ : การอักเสบ / ข้อมูลไมโครอาร์เรย์ / การคัดเลือกยีน / การวิเคราะห์เครือข่ายการแสดงออกร่วม

Abstract

Inflammatory response gene discovery are important to explain the mechanism of inflammation in molecular level. Microarray data analysis has been applied. Three microarray datasets of lipopolysaccharide (LPS)-stimulated macrophage RAW 264.7 with four different incubation times were used in this study. Selected microarray data were used to reconstruct co-expression network with online tool, Network analysis of ArrayMining. These networks were shown through Cytoscape. Sub-networks were chosen with a set of genes which were selected from Gene selection tools of ArrayMining. These results showed that the selected were not reported to be inflammatory responsive genes when compare with genes from PolySearch database. In summary, these genes may be novel inflammatory responsive genes which must be proved in the next study.

Keywords: Inflammation / Microarray data / Gene selection / Co-expression network analysis

*Corresponding author. E-mail : pitak@buu.ac.th

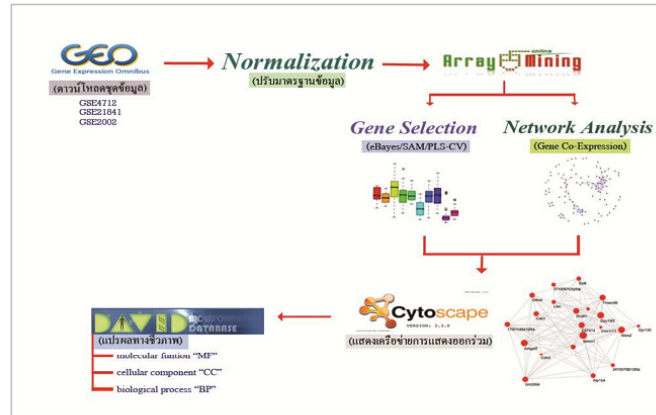
1. บทนำ

การอักเสบ (Inflammation) เป็นกระบวนการที่ร่างกายตอบสนองต่อสิ่งที่ทำให้เนื้อเยื่อของร่างกายได้รับบาดเจ็บ โดยจะกระตุ้นให้เซลล์ในระบบภูมิคุ้มกันมีการตอบสนองต่อสิ่งกระตุ้นหรือสิ่งแปลกปลอมที่ถูกกำจัดออกไป หากมีการอักเสบที่มากเกินไปจะก่อให้เกิดโรคได้ ดังนั้นหากเข้าใจกลไกการตอบสนองของเซลล์เมื่ออยู่ในสภาวะที่เกิดการอักเสบ อาจนำมาประยุกต์ใช้ในทางการแพทย์เพื่อค้นหาวิธีลดหรือต้านการอักเสบได้ (Nilsson et al., 2006) การศึกษากลไกการอักเสบจะอาศัยข้อมูลไมโครอาร์เรย์ซึ่งเป็นข้อมูลระดับการแสดงออกของยีนจำนวนมากในเวลาเดียวกัน โดยจะอาศัยวิธีการวิเคราะห์ข้อมูลไมโครอาร์เรย์ในการคัดเลือกเฉพาะกลุ่มยีนที่มีบทบาทสำคัญในการตอบสนองต่อการอักเสบ เช่น การคัดเลือกยีนโดยวิธี Gene selection จากงานวิจัยของ วันเจลิม จินอู และพิทักษ์ สุตรอนันต์ (2557) ในการคัดเลือกกลุ่มยีนที่มีการแสดงออกที่แตกต่างกันอย่างเด่นชัด จากข้อมูลไมโครอาร์เรย์ภายใต้สภาวะการอักเสบของเซลล์แมคโครฟาจ RAW 264.7 ที่มีการกระตุ้นด้วย LPS โดยใช้วิธีการคัดเลือกที่อยู่ภายในโปรแกรม ArrayMining (<http://www.arraymining.net/>) (Glaab et al., 2009) ทำให้ได้กลุ่มยีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบ การสร้างเครือข่ายการแสดงออกร่วมของยีน (gene co-expression network; GCN) (Mabbott et al., 2013) จะช่วยให้เข้าใจภาพรวมของยีนที่มีความสัมพันธ์กัน ซึ่งช่วยทำให้อธิบายกลไกที่เกี่ยวข้องได้อย่างเป็นระบบ ดังเช่นตัวอย่างจากงานวิจัยของเทพพนม อารัญญาน และคณะ (2558) ดังนั้นในงานวิจัยนี้จึงได้รวมสองเทคนิคข้างต้นไว้ด้วยกัน เพื่อให้การค้นหายีนที่มีความสำคัญต่อการตอบสนองต่อการอักเสบมีความครอบคลุมและช่วยให้เข้าใจกลไกการอักเสบได้ดียิ่งขึ้น

งานวิจัยนี้จะทำการตรวจสอบข้อมูลการแสดงออกของยีนในการตอบสนองต่อการอักเสบของเซลล์แมคโครฟาจ RAW 264.7 ที่ถูกกระตุ้นด้วย LPS จากฐานข้อมูลสาธารณะ GEO (Gene Expression Omnibus) (Edgar and Barrett, 2006) มาคัดเลือกกลุ่มยีนที่มีการแสดงออกอย่างเด่นชัด การวิเคราะห์เครือข่ายการแสดงออกร่วมของยีน (gene co-expression network analysis) โดยอาศัยโปรแกรม ArrayMining (<http://www.arraymining.net/>) (Glaab et al., 2009) การคัดเลือกเครือข่ายย่อยโดยใช้โปรแกรม Cytoscape (Comer et al., 2006; Hammer et al., 2010) และการแปลผลทางชีวภาพโดยใช้ยีนออนโทโลยี (Gene ontology; GO) ภายใต้โปรแกรมออนไลน์ DAVID (Huang et al., 2009) เพื่อนำมาใช้ในการอธิบายกลไกระดับโมเลกุลที่เกิดขึ้นของการตอบสนองการอักเสบในช่วงเวลาที่แตกต่างกัน

2. วิธีการ

รวบรวมข้อมูลไมโครอาร์เรย์ที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบจากฐานข้อมูลสาธารณะ GEO (Gene Expression Omnibus) (Edgar and Barrett, 2006) ได้แก่ GSE2002, GSE4712 และ GSE21841 โดยทั้ง 3 ชุดข้อมูลจะมีช่วงเวลากระตุ้นเซลล์แมคโครฟาจ RAW 264.7 ด้วย LPS ที่แตกต่างกัน คือ 3, 6, 8, และ 18 ชั่วโมง จากนั้นทำการปรับมาตรฐานและรูปแบบของข้อมูลตามแต่ละโปรแกรม กำหนด นำชุดข้อมูลไปทำการวิเคราะห์ในโปรแกรม ArrayMining (<http://www.arraymining.net/>) (Glaab et al., 2009) โดยในงานวิจัยนี้เลือกใช้การคัดเลือกโพรบ (probe) โดย Gene selection ที่คัดเลือกโพรบจาก 3 วิธี คือ SAM (Tusher, Tibshirani and Chu, 2001), eBayes (Smyth, 2004) และ PLS-CV (Hall, 2000) ซึ่งโพรบที่ถูกคัดเลือกอย่างน้อย 2 วิธี จะถูกนำมาเลือกเครือข่ายย่อยจากเครือข่ายการแสดงออกร่วมของโพรบภายในข้อมูลไมโครอาร์เรย์ที่ถูกสร้างขึ้นโดยเครื่องมือ Network Analysis ในโปรแกรม ArrayMining (<http://www.arraymining.net/>) (Glaab et al., 2009) ก่อน แปลผลโพรบเป็นยีน แล้วนำผลที่ได้ไปแสดงผลและคัดเลือกเครือข่ายย่อยของยีนด้วยโปรแกรม Cytoscape (Comer et al., 2006; Hammer et al., 2010) และอาศัยการแปลผลทางชีวภาพจากฐานข้อมูล DAVID (Huang et al., 2009) ของยีนที่พบในเครือข่ายย่อย โดยการแปลผลนี้จะทำให้ทราบถึงตำแหน่งการทำงานของยีน และฟังก์ชันการทำงานรวมไปถึงหน้าที่การทำงานของยีนนั้นว่าเกี่ยวข้องอย่างไรในกระบวนการอักเสบ โดยรายละเอียดของขั้นตอนการศึกษาแสดงดังภาพที่ 1



ภาพที่ 1 ภาพรวมวิธีการวิจัยของการค้นหายีนจากข้อมูลไมโครอาร์เรย์

3. ผลและอภิปราย

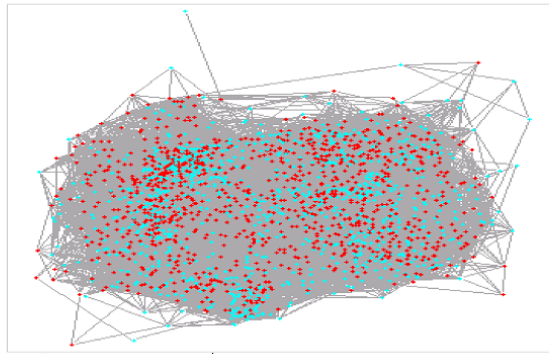
จากการคัดเลือกโพรบที่มีการตอบสนองต่อการอักเสบของเซลล์แมคโครฟาจ RAW 264.7 ที่ถูกกระตุ้นด้วย LPS ตามช่วงเวลา 3, 6, 8, และ 18 ชั่วโมง จากชุดข้อมูลไมโครอาร์เรย์ โดยเลือกใช้วิธีการคัดเลือกโพรบ 3 วิธี คือ SAM (Tusher, Tibshirani and Chu, 2001), eBayes (Smyth, 2004) และ PLS-CV (Hall, 2000) ในโปรแกรม ArrayMining (<http://www.arraymining.net/>) (Glaab et al., 2009) โดยคัดเลือกโพรบที่มีการแสดงออกที่เด่นชัดจำนวน 100 โพรบแรก (top 100) เพื่อยืนยันว่าโพรบที่ถูกคัดเลือกมานั้นเป็นโพรบที่มีความเกี่ยวข้องกับการอักเสบจริง จึงได้ทำการคัดเลือกโพรบที่ถูกคัดเลือกด้วยวิธีการข้างต้นอย่างน้อย 2 วิธี เนื่องจากแต่ละวิธีมีกระบวนการคำนวณที่แตกต่างกัน (วันเฉลิม จินอู และพิทักษ์ สุตรอนันต์, 2557) ได้ผลดังตารางที่ 1

ตารางที่ 1 ผลของโพรบที่ถูกคัดเลือกอย่างน้อย 2 วิธีจาก 3 วิธี ได้แก่ SAM, eBayes และ PLS-CV

ชุดข้อมูล	จำนวนของโพรบที่ถูกคัดเลือกได้ (โพรบ, probe)			
	SAM	eBayes	PLS-CV	โพรบที่พบอย่างน้อย 2 วิธี
GSE4712 ที่ 3 ชั่วโมง	5	100	100	35
GSE4712 ที่ 6 ชั่วโมง	100	100	100	65
GSE21841 ที่ 8 ชั่วโมง	20	100	100	50
GSE2002 ที่ 18 ชั่วโมง	100	100	100	96

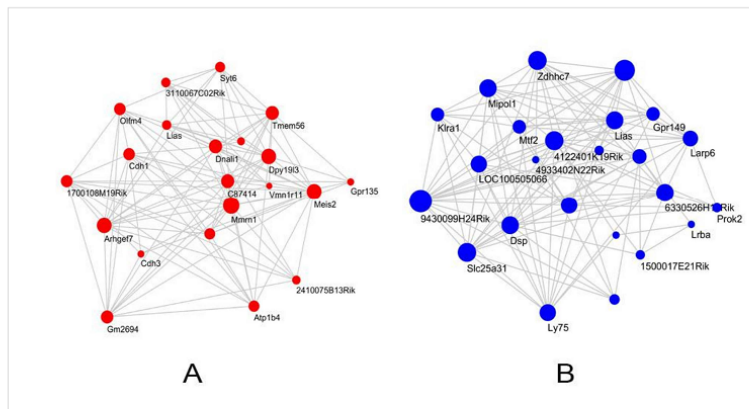
เมื่อนำชุดข้อมูลไมโครอาร์เรย์ที่ทำการปรับมาตรฐานและจัดรูปแบบแล้ว มาทำการสร้างเครือข่ายการแสดงออกร่วมของโพรบ ด้วยโปรแกรมย่อย Network analysis ภายในโปรแกรมออนไลน์ ArrayMining (<http://www.arraymining.net/>) (Glaab et al., 2009) โดยข้อมูลเครือข่ายการแสดงออกร่วมของโพรบจะถูกนำมาแสดงผลด้วยโปรแกรม Cytoscape (Comer et al., 2006; Hammer et al., 2010) ซึ่งผลที่ได้พบว่าชุดข้อมูล GSE4712 ที่ 3 ชั่วโมง ได้เครือข่ายที่มีจำนวนโพรบ 2,000 โพรบ (probes หรือ nodes) และ 61,096 ปฏิสัมพันธ์ (interactions หรือ edges) ที่ 6 ชั่วโมง ได้เครือข่ายขนาด 2,000 nodes และ 60,627 edges แต่ในชุดข้อมูล GSE21841 ที่ 8 ชั่วโมง และชุดข้อมูล GSE2002 ที่ 18 ชั่วโมง ไม่สามารถสร้างเครือข่ายการแสดงออกร่วมของโพรบได้ เนื่องจากการแสดงออกของโพรบในชุดข้อมูลมีความสัมพันธ์ต่ำมาก อาจมีสาเหตุมาจากขนาดของข้อมูลที่มีจำนวนตัวอย่างน้อย และข้อมูลการแสดงออกไม่เด่นชัด โปรแกรมจึงไม่สามารถให้ค่านำหน้า

ความสัมพันธ์ระหว่างโหนดได้ และเมื่ออาศัยข้อมูลการคัดเลือกโหนดที่มีการแสดงออกอย่างเด่นชัดและถูกคัดเลือกจาก 2 วิธี จาก 3 วิธีข้างต้นของแต่ละชุดข้อมูล มาแปลงข้อมูลเป็นยีนและใช้ในการคัดเลือกเครือข่ายย่อย ทำให้ได้เครือข่ายย่อยในชุดข้อมูล GSE4712 ที่ 3 ชั่วโมง ที่มีจำนวนยีนเท่ากับ 21 ยีน (genes หรือ nodes) และ 111 ปฏิสัมพันธ์ (interactions หรือ edges) และ ชุดข้อมูล GSE4712 ที่ 6 ชั่วโมง จำนวน 24 nodes และ 218 edges แสดงดังภาพที่ 3A และ 3B ตามลำดับ



ภาพที่ 2 เครือข่ายปฏิสัมพันธ์ที่ได้จากการวิเคราะห์ Network analysis ของชุดข้อมูล GSE4712 ที่ 3 ชั่วโมง ซึ่งแสดงผลในโปรแกรม Cytoscape

จากเครือข่ายย่อยที่แสดงดังภาพที่ 3 เมื่อทำการปรับขนาดของโหนด (nodes) ให้แปรตามจำนวนปฏิสัมพันธ์หรือเอดจ์ (edges) ซึ่งโหนดขนาดใหญ่จะมีปฏิสัมพันธ์สูง โดยชุดข้อมูล GSE4712 ที่ 3 ชั่วโมงพบว่ายีน *MMRN1* มีจำนวนยีนที่มีปฏิสัมพันธ์ร่วมด้วย 17 ยีน เช่น ยีน *Arhgef7*, *Dnal1*, *Dpy19l3* และ *C87414* ดังภาพที่ 3A และในชุดข้อมูล GSE4712 ที่ 6 ชั่วโมง พบยีน *MIPOL1* มียีนที่มีปฏิสัมพันธ์ร่วมด้วยจำนวน 23 ยีน เช่น ยีน *Larp6*, *Lias*, *Ly75*, *Zdhc7* และ *Dsp* ดังภาพที่ 3B เมื่อทำการตรวจสอบหน้าที่ของกลุ่มยีนที่ถูกคัดเลือก พบว่ายีนทั้งหมดไม่เคยตรวจพบมาก่อนว่าเกี่ยวข้องกับการอักเสบเมื่อเปรียบเทียบกับยีนที่ได้ฐานข้อมูล PolySearch (Cheng et al, 2008) เนื่องจากฐานข้อมูล PolySearch เป็นฐานข้อมูลที่เกี่ยวข้องกับการตรวจสอบรายชื่อยีนที่มีความสัมพันธ์หรือเกี่ยวข้องกับโรคที่สนใจ ผ่านการค้นหาข้อมูลตัวอักษรด้วยกระบวนการเทกซ์ไมนิ่ง (text mining) (Cheng et al, 2008) ภายในบทความวิจัย ซึ่งทำให้ความได้ว่ายีนกลุ่มนี้อาจเป็นยีนใหม่ที่น่าจะมีความเกี่ยวข้องกับการอักเสบก็เป็นได้



ภาพที่ 3 ภาพแสดงเครือข่ายย่อยของการแสดงออกร่วมของยีน ในชุดข้อมูล GSE4712 ที่ 3 ชั่วโมง (A) และชุดข้อมูล GSE4712 ที่ 6 ชั่วโมง (B)

จากการแปลผลทางชีวภาพของกลุ่มยีนจากเครือข่ายย่อยดังภาพที่ 3 ด้วยฐานข้อมูล DAVID (<http://david.abcc.ncifcrf.gov/>) (Huang *et al.*, 2009) สามารถอธิบายถึงคุณสมบัติที่เป็นเอกลักษณ์ของยีน และกลุ่มยีนที่ถูกคัดเลือกมีความเกี่ยวข้องกับการตอบสนองการอักเสบของเซลล์แมคโครฟาจ RAW264.7 ที่ถูกกระตุ้นด้วย LPS ดังแสดงในตารางที่ 2 และ ภาพที่ 4 โดยยีนที่มีความโดดเด่นสูงในชุดข้อมูล GSE4712 ที่ 3 ชั่วโมง ได้แก่ยีน *Arhgef7*, *1700108M19Rik*, *Mmm1*, *Olfm4*, *Gm2694*, *Tmem56*, *Meis2*, *Dnal1*, *Dpy19l3* และ *C87414* เป็นยีนในกลุ่ม cell adhesion ที่อยู่ในชุดของข้อมูลที่เกี่ยวข้องกับกลไกที่ชักนำให้เซลล์เม็ดเลือดขาวเข้ามาในเนื้อเยื่อบริเวณที่เกิดการอักเสบ จะมีการแสดงออกมากในส่วนของชุดข้อมูล GSE4712 ที่ 6 ชั่วโมง ที่มีความโดดเด่นสูง ได้แก่ยีน *Larp6*, *9430099H24Rik*, *4122401K19Rik*, *Mipol1*, *Lias*, *Ly75*, *Dsp*, *Slc25a31*, *Zdhc7* และ *6330526H18Rik* ซึ่งกลุ่มของยีนที่พบส่วนใหญ่จะมีผลผลิตโปรตีนที่ทำงานภายใต้พื้นฐานของการถอดรหัสและแปลรหัสของชุดข้อมูล DNA และ RNA ที่เกี่ยวข้องกับการรับส่งสัญญาณในกระบวนการเกิดการอักเสบ ซึ่งยีนเหล่านี้เป็นยีนที่มีปฏิสัมพันธ์สูงในแต่ละกลุ่มของข้อมูล

ตารางที่ 2 molecular function, cellular component และ biological process ของชุดข้อมูล GSE4712 ที่ 3 และ 6 ชั่วโมง

	GOTERM_MF	GOTERM_CC	GOTERM_BP
3hr	binding	cell junction	cell adhesion
	protein binding	membrane	biological adhesion
	ion binding	cytoplasm	cellular process
6hr	binding	mitochondrion	regulation of biological process
	ion binding	nucleus	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process
	molecular transducer activity	membrane	localization



ภาพที่ 4 แสดงถึงความสอดคล้องของการแปลผลทางชีวภาพจากการคัดเลือกกลุ่มยีนในแต่ละชุดข้อมูลตามช่วงเวลาที่แตกต่างกันและของข้อมูลพาทเวีย TLR signaling pathway ที่ได้จากฐานข้อมูล BioCarta (<http://www.biocarta.com/>)

4. บทสรุป

การวิเคราะห์ข้อมูลไมโครอาร์เรย์ในเซลล์แมคโครฟาจ RAW 264.7 ที่ถูกกระตุ้นด้วย LPS โดยใช้กระบวนการการคัดเลือกยีน (Gene selection) และการสร้างเครือข่ายการแสดงออกร่วม (Network analysis) ที่อยู่ในโปรแกรม ArrayMining (<http://www.arraymining.net/>) (Glaab *et al.*, 2009) ทำให้ได้เครือข่ายย่อยของกลุ่มยีนที่มีการแสดงออกร่วมกัน เฉพาะในช่วงเวลาที่ 3 และ 6 ชั่วโมง เมื่อทำการตรวจสอบรายชื่อยีนกับฐานข้อมูล PolySearch (Cheng *et al.*, 2008) พบว่ายีนทั้งหมดไม่เคยตรวจพบมาก่อนว่าเกี่ยวข้องกับการอักเสบ ทำให้คาดว่ากลุ่มยีนเหล่านี้ อาจเป็นกลุ่มยีนใหม่ที่มีความเกี่ยวข้องกับการตอบสนองการอักเสบ ที่จะต้องทำการตรวจสอบ

เพิ่มเติมในการศึกษาคั้งต่อไป ซึ่งในอนาคตอาจเป็นยีนที่มีความสำคัญในการเป็นตัวบ่งชี้ถึงระดับการแสดงออกของยีนของการตอบสนองการอักเสบได้

5. กิตติกรรมประกาศ

ขอขอบพระคุณคณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา ที่ให้ทุนสนับสนุนการนำเสนอผลงานในงานวิจัย และงานวิจัยนี้ได้รับทุนอุดหนุนการวิจัย มหาวิทยาลัยบูรพา ประจำปีงบประมาณ พ.ศ.2557 สัญญาเลขที่ 87/2557

6. เอกสารอ้างอิง

- เทพพนม อารัญญาณ, จรัญ รุ่งเรือง, เหมือนฝัน ไหวหารกล้า และ พิทักษ์ สุตรอนันต์. (2558). เครือข่ายความสัมพันธ์การแสดงออกร่วมของกลุ่มยีนในเซลล์แมคโครฟาจ RAW 264.7 เมื่อถูกกระตุ้นด้วย LPS. *Proceedings of The 7th National Science Research Conference*, 30 – 31 มีนาคม 2558, Naresuan University.
- วันเฉลิม จินธุ์ และ พิทักษ์ สุตรอนันต์. (2557). กลไกการตอบสนองการอักเสบจากการวิเคราะห์ข้อมูลไมโครอาร์เรย์. รายงานสืบเนื่องจากการประชุมวิชาการ "การศึกษาทางพันธุศาสตร์และจีโนมิกส์ด้วยเครื่องวิเคราะห์ยุคใหม่" *Next Generation Sequencing for Genetic and Genomic Studies (NGS2014)*. 29 – 30 กรกฎาคม 2557, โรงแรมวินเซอร์ สวีทส์ สุขุมวิท กรุงเทพมหานคร.
- Comer, J.E., Galindo, C.L., Zhang, F., Wenglikowski, A.M., Bush, K.L., Garner, H.R., et al. (2006). Murine macrophage transcriptional and functional responses to *Bacillus anthracis* edema toxin. *Microbial Pathogenesis*, 41(2-3), 96 – 110.
- Cheng, D., Knox, C., Young, N., Stothard, P., Damaraju, S. and Wishart, D.S. (2008). PolySearch: a web-based text mining system for extracting relationships between human diseases, genes, mutation, drugs and metabolites. *Nucleic Acids Research*, 36, W399 – 405.
- Edgar, R., Domrachev, M. and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, 30(1), 207 – 210.
- Glaab, E., Garibaldi, J.M. and Krasnogor, N. (2009). Array Mining: a modular web-application for microarray analysis combining ensemble and consensus methods with cross-study normalization. *BMC Bioinformatics*, 10, 358 (1 – 7).
- Hall, M.A. (2000). Correlation-based feature selection for discrete and numeric class machine learning. *Proceedings of the Seventeenth International Conference on Machine Learning*. Stanford University, Stanford, United States of America, 359-366.
- Hammer, K.D., Yum, M.Y., Dixon, P.M. and Birt, D.F. (2010). Identification of JAK-STAT pathways as important for the anti-inflammatory activity of a *Hypericum perforatum* fraction and bioactive constituents in RAW 264.7 mouse macrophages. *Phytochemistry*, 71(7), 716 – 725.
- Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4, 44 – 57.
- Mabbott, N.A., Baillie, J.K., Brown, H., Freeman, T.C. and Hume, D.A. (2013). An expression atlas of human primary cells inference of gene function from coexpression networks. *BMC Genomics*, 14, 632 – 645.
- Nilsson R, Bajic VB, Suzuki H, Bernardo DD, Björkegren J, Katayama S, Reid JF, Sweet MJ, Gariboldi M, Carninci P, Hayashizaki Y, Hume DA, Tegner J, Ravasi T. (2006). Transcriptional network dynamics in macrophage activation. *Genomics*. 88(2), 133-142.
- Smyth, G.K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Application in Genetics and Molecular Biology*, 3, 1027 – 1030.
- Tusher, V.G., Tibshirani, R. and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proceeding of the National Academy of Sciences of the United States of America*, 98(9), 5116 – 5121.

ภาคผนวก ข

นำเสนอรายงานสืบเนื่องจากงานประชุมทางวิชาการระดับนานาชาติ

The 27th Annual Meeting of the Thai Society for Biotechnology and
International Conference (TSB2015).

17 – 20 November 2015, Bangkok, Thailand.

Wohankla, M., Srisook, K. and Soonanan, P. (2015). Identification of low differentially expressed responsive genes in microarray data of LPS-stimulated macrophage RAW264.7. The 27th Annual Meeting of the Thai Society for Biotechnology and International Conference (TSB2015). November 17-20, 2015 Mandarin Hotel Bangkok by Centre Point, Bangkok, Thailand. 752 – 757.

Identification of low differentially expressed responsive genes in microarray data of LPS-stimulated macrophage RAW264.7

Muanfan Wohankla¹, Klaokwan Srisook² and Pitak Soonanan^{2*}

¹Biological Science Program, Faculty of Science, Burapha University, Chonburi, Thailand

²Department of Biochemistry, Faculty of Science, Burapha University, Chonburi, Thailand

*e-mail: pitak@buu.ac.th

Abstract

Lipopolysaccharide (LPS), an endotoxin present in gram-negative bacteria, is known to activate inflammatory signals in macrophages. LPS-stimulated macrophage causes stimulation of pro-inflammatory cytokine gene expression and cytokine production. The finding of key responsive genes helped for better understanding the mechanisms of systematic inflammation. Feature selection methods in microarray data such as SAM, eBayes, and PLS-CV were used for searching of high differentially expressed genes to be the key responsive ones. However, these methods could lose some important genes that have low differential expression levels. Therefore, microarray gene expression datasets involving LPS-response in macrophage cell obtained from the open-access database were investigated in this study. The low differentially expressed responsive genes were identified by using the integration of feature selection methods, literature mining tool and gene co-expression network approach. The results demonstrated three low differentially expressed inflammatory responsive genes, i.e. *CD36*, *TSPO* and *PLA2G4A* from two co-expression modules. These genes interacted with 19 high differentially expressed genes. Three selected inflammatory responsive genes and some high differentially expressed genes will be further validated on time-series wet-lab in the near future.

Keywords: Inflammation, Feature selection, Literature mining, Gene co-expression network, Responsive genes

Introduction

Molecular mechanisms within cells that respond to stimuli can be understood from whole gene expression. Microarray technology can measure the expression level of thousands of genes within a particular mRNA (Watson et al., 1998). Microarray data are submitted from several research groups in public databases such as Gene Expression Omnibus (GEO) (Edgar et al., 2002). Some of these data are associated with studies of the lipopolysaccharide (LPS) response in RAW264.7 macrophage cell (Hammer et al., 2010). LPS activates inflammatory signaling in macrophages and stimulates pro-inflammatory cytokine gene expression (Fujihara et al., 2003). Thus, identification of key responsive genes is important to understand inflammatory response mechanisms. Feature selection step is a challenge in microarray data analysis as datasets are commonly very large. It is extremely useful to select gene sets important to the conditions under investigation (Selvaraj et al., 2011). ArrayMining includes several feature selection methods to identify high differentially expressed genes such as Significance analysis in microarrays (SAM), Empirical Bayes *t*-statistic (eBayes) and Partial-Least-Squares based filter (PLS-CV) (Glaab et al., 2009). These methods might also suffer from incompleteness or missing some important responsive genes that have low differential expression levels. The issue is being corrected by integration analysis of gene selection with the information of known inflammatory genes. It may be arriving at a comprehensive to identify low differentially expressed inflammatory responsive genes. Inflammatory responsive genes can be found with literature mining tools, i.e. PolySearch database (Cheng et al., 2008). Another integration analysis is to analyze with gene co-expression network (Stuart, Segal, Koller and Kim, 2003). This co-expression network helps to find the correlation among important inflammatory responsive genes. In this study, low differentially expressed inflammatory responsive genes were discovered by integration of gene selection methods, literature mining tool, and gene co-expression network. These inflammatory responsive genes are important to explain the mechanism of inflammation in the LPS-stimulated macrophage system. It could be applied to be a model for the other related studies, e.g. investigating the mechanism of anti-inflammatory effects of plant extracts.

Materials and methods

Identification of low differentially expressed responsive inflammatory genes consists of six steps, summarized in Fig. 1.

Step 1: Retrieval of three microarray gene expression datasets involving LPS response in RAW264.7 macrophage cell line from Gene Expression Omnibus (GEO) database (Edgar et al., 2002; <http://www.ncbi.nlm.nih.gov/geo/>). All datasets were normalized by \log_{10} followed by z-score transformations.

Step 2: Selection of differentially expressed responsive genes by three feature selection methods such as SAM, eBayes and PLS-CV on a web-application ArrayMining (Glaab et al., 2009, <http://www.arraymining.net>). A total of 100 top ranked probes were selected from each method. Only probes that selected more than one method were chosen. These probes are annotated as their genes and had high differential expression levels.

Step 3: Retrieval of inflammatory genes from the literature using PolySearch text mining tool (Cheng et al., 2008, <http://wishart.biology.ualberta.ca/polysearch/>).

Step 4: Integrative analysis of inflammatory genes and the high differentially expressed responsive genes.

Step 5: Construction of gene co-expression network with data obtained from Step 4 via Cytoscape plugin, ExpressionCorrelationViewer (Cline et al., 2007).

Step 6: Selection of sub-network for gene co-expression network from Step 5 that have interesting interactions among these genes. Low differentially expressed responsive inflammatory genes were identified and their function analyzed with Cytoscape plugin, Golorize (Garcia et al., 2007).

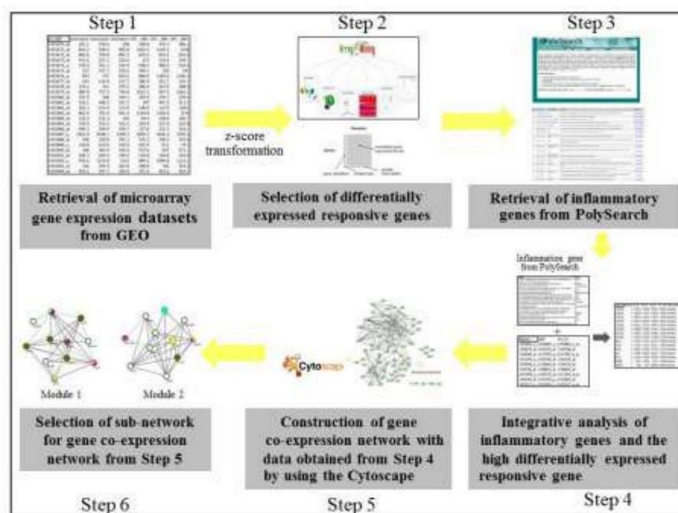


Fig 1: A brief overview of the identification of low differentially expressed responsive inflammatory genes.

Results and discussion

High differentially expressed responsive genes were selected with more than one gene selection method including SAM, eBayes, and PLS-CV (Table 1). These methods are available in ArrayMining (Glaab et al., 2009, <http://www.arraymining.net>). SAM is a method to identify genes with statistically significant changes in expression by assimilating a set of gene-specific *t*-tests (Tusher et al., 2001). eBayes is method used to rank genes in order of evidence for differential

expression. Bayes approach based on this rank statistic ranks genes by testing whether all pairwise contrasts between different outcome-classes are zero (Smyth, 2004). PLS-CV is estimated from magnitudes of the efficient obtained from training a Partial Least Squares classifier. The number of PLS-components is selected based on the cross-validation accuracies for 20 random 2/3-partitions of the data for all possible values of n (Hall, 2000). These three methods were selected dependent on the analysis of their different algorithms. This would be enhancing the credibility for selecting the high differentially expressed responsive genes. These finding could lose some important genes that have low differential expression levels. However, integration analysis of gene selection with the information of known inflammatory genes may be useful for more reliable analysis to identify targets of key responsive genes. Inflammatory genes can be retrieved from PolySearch database (Cheng et al., 2008, <http://wishart.biology.ualberta.ca/polysearch/>).

Table 1: Number of probes selected by three different methods (SAM, eBayes, and PLS-CV), number of probes selected more than one method and the number of genes found in microarray data.

GSE ID	SAM (probe)	eBayes (probe)	PLS-CV (probe)	Number of probes selected more than one methods (probe)	Number of Gene Symbol (gene)
GSE4712 (3 hr_LPS)	100	100	100	89	70
GSE4712 (6 hr_LPS)	100	100	100	98	64
GSE21841 (8 hr_LPS)	100	100	100	92	61
GSE2002 (18 hr_LPS)	100	100	100	93	80

Low differentially expressed responsive genes were found by their interaction with high differentially expressed inflammatory responsive genes in the gene co-expression network. This network was constructed by using the expression data of 312 selected genes which had high differentially expressed responsive genes and were in the list of inflammatory genes. The network consisted of 100 nodes (genes) and 333 edges (interactions). It was created with default correlation cut-off values of ± 0.95 (Fig 2A-2B). Two modules were found with their interested interaction among the nodes which represent high differentially expressed responsive genes (ArrayMining), inflammatory genes (PolySearch), and both of ArrayMining and Polysearch (Fig 2C). Three low differentially expressed inflammatory genes, i.e. *CD36*, *TSPO* and *PLA2G4A*, were found from these two co-expression modules (Fig 2C-2D). These were inflammatory genes which had interactions with 19 high differentially expressed responsive genes. Cytoscape plugin (Golorize; Garcia et al., 2007), was used to analyze and visualize these genes in the network with their biological processes. Overall, 13 GO terms associated with inflammation process were selected as represented in Fig 3. The main biological processes form GO were as following: regulation of macrophage derived foam cell differentiation, immune system process, cytokine metabolic process, activation of innate immune response, regulation of I-kappaB kinase/NF-kappaB cascade, regulation of tumor necrosis factor production, regulation of MAPKK cascade, regulation of cytokine production involved in immune response, lipid biosynthetic process, arachidonic acid metabolic process and cell cycle. The results of three genes demonstrated that these genes are associated with various processes that will lead to inflammation (Fig 3).

These three genes were associated with those that are highly differentially expressed. CD36, family members of class B scavenger receptors, is widely expressed in a variety of cell types, including vascular endothelial, smooth muscle, macrophages, adipocytes, and hepatocytes. It plays an important role in physiological processes, including regulation of macrophages derived foam cell differentiation, tumor necrosis factor production and regulation of MAPKK cascade (Febbraio, Hajjar and Silverstein, 2001). TSPO is a 18 kDa outer-mitochondrial membrane transporter which has many functions including roles in immune response, steroid synthesis and apoptosis (Daugherty et al., 2013). PLA2G4A is a member of the cytosolic phospholipase A2 group IV family. This enzyme catalyzes the release of arachidonic acid from membrane phospholipids which is further metabolized to eicosanoids that play important roles in inflammatory responses, and other intracellular pathways (Sheridan et al., 2001). The approach used in this study could be implemented to find low differentially expressed genes that are the important inflammatory responsive genes which could not

be found by several feature selection methods. A combination of other network approaches such as gene regulatory network and protein-protein interaction network can help increasing the reliability of a network result. This approach can be implemented for prediction of other candidate responsive genes that can help to explain the mechanism of the inflammatory response process and might help in the search for a cure for inflammation.

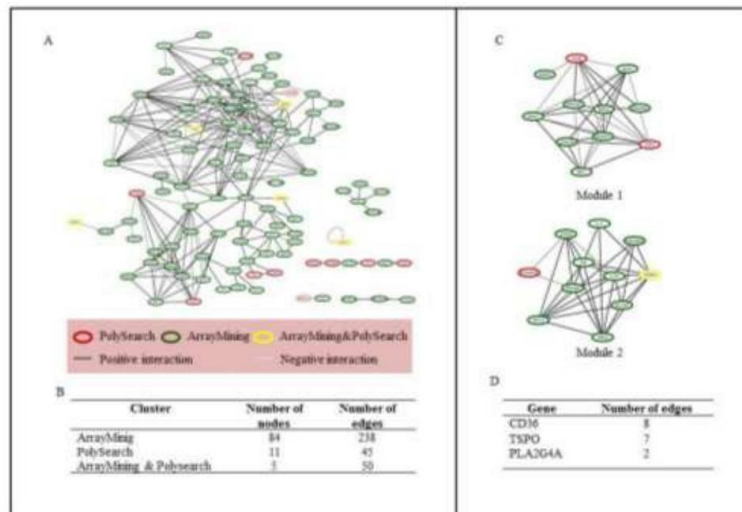


Fig 2: Co-expression network analysis of selected genes from literature mining tool (PolySearch database) and feature selection methods (ArrayMining) (A), Number of nodes and edges in co-expression network (B), sub-networks of co-expression network were selected (C) and number of edges of three low differentially expressed inflammatory responsive genes from two sub-networks (D).

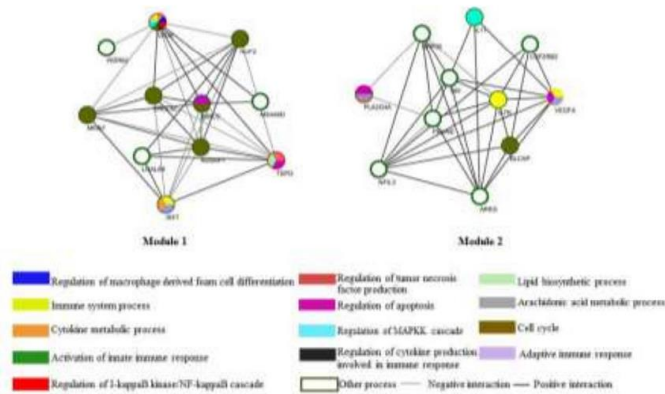


Fig 3: Biological functions of genes in two sub-networks (Fig 2C) interpreted with biological processes in Gene ontology (GO) using Cytoscape plugin, Golorize.

Conclusions

In this study, feature selection methods, literature mining tool and gene co-expression network approach were integrated to identify the low differentially expressed inflammatory responsive genes. This approach resulted in the identification of low differentially expressed genes that could be potentially important in the inflammatory processes. The results help to better

understand the mechanisms in the inflammatory response of LPS-stimulated macrophage RAW264.7 cell and represent targets of key genes for further investigation in wet-lab.

Acknowledgements

This research was supported by Burapha University through its research grant number 87/2557. We would like to thank Faculty of Science, Burapha University with its grant for graduate student in year 2558. We would like to gratefully thank to Prof. Dr. Frederick W.H. Beamish, who helped to proofread our proceeding.

References

- Cheng, D., Knox, C., Young, N., Stotharc, P., Damaraju, S., and Wishart, D.S. 2008. PolySearch: a web-based text mining system for extracting relationships between human diseases, genes, mutations, drugs and metabolites. *Nucleic Acids Research*, 36 (Web Server issue): W399–W405. doi:10.1093/nar/gkn296.
- Cline, M.S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., et al. (2007). Integration of biological networks and gene expression data using cytoscape. *Nature Protocols*, 2(10), 2366-2382.
- Daugherty, D.J., Selvaraj, V., Chechneva, O.V., Liu, X.B., Pleasure, D.E., and Deng, W. 2013. A TSPO ligand is protective in a mouse model of multiple sclerosis. *EMBO Molecular Medicine*, 5: 891-903.
- Edgar, R., Domrachev, M., and Lash A.E. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, 30: 207-210.
- Febbraio M., Hajjar D.P., and Silverstein R.L. 2001. CD36: a class B scavenger receptor involved in angiogenesis, atherosclerosis, inflammation, and lipid metabolism. *Journal of clinical investigation*. 108: 785-791.
- Fujihara, M., Muroi, M., Tanamoto, K., Suzuki, T., Azuma, H., and Ikeda, H. 2003. Molecular mechanisms of macrophage activation and deactivation by lipopolysaccharide: roles of the receptor complex. *Pharmacology and Therapeutics*, 100: 171-194.
- Garcia, O., Saveanu, C., Cline, M., Fromont-Racine, M., Jacquier, A., Schwikowski, B., and Aittokallio, T. 2007. Golorize: a Cytoscape plug-in for network visualization with Gene Ontology-based layout and coloring. *Bioinformatics*, 23: 394-396.
- Glaab, E., Garibaldi, J.M., and Krasnogor, N. 2009. Array Mining: a modular web-application for microarray analysis combining ensemble and consensus methods with cross-study normalization. *BMC Bioinformatics*, 10: 1-7.
- Hall, M.A. 2000. Correlation-based feature selection for discrete and numeric class machine learning. *Proceedings of the Seventeenth International Conference on Machine Learning*, Stanford University, Stanford, United States of America.
- Hammer, K.D.P., Yum, M.Y., Dixon, P.M., and Birt D.F. 2010. Identification of JAK-STAT pathways as important for the anti-inflammatory activity of a *Hypericum perforatum* fraction and bioactive constituents in RAW264.7 mouse macrophages. *Phytochemistry*, 71: 716-725.
- Selvaraj, S., and Natarajan, J. 2011. Microarray Data Analysis and Mining Tools. *Bioinformatics*, 6: 95-99.
- Sheridan, A.M., Force, T., Yoon, H.J., O'Leary, E., Choukroun, G., Taheri M.R., et al. 2001. PLIP, a novel splice variant of Tip60, interacts with group IV cytosolic phospholipase A(2), induces apoptosis, and potentiates prostaglandin production. *Molecular and Cellular Biology*, 21: 4470-4481.
- Smyth, G.K. 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Application in Genetics and Molecular Biology*, 3: 1027-1030.
- Stuart, J.M., Segal, E., Koller, D., and Kim, S.K. 2003. A gene-coexpression network for global discovery of conserved genetic modules. *Science*, 302: 249-255.
- Tusher, V.G., Tibshirani, R., and Chu, G. 2001. Significance analysis of microarrays applied to the ionizing radiation response. *Proceeding of the National Academy of Sciences of the United States of America*, 98: 5116–5121.

The 27th Annual Meeting of the Thai Society for Biotechnology and International Conference

Watson, A., Mazumder, A., Stewart, M., and Balasubramanian, S. 1998. Technology for microarray analysis of gene expression. *Current Opinion in Biotechnology*, 9: 609-614.

November 17-20, 2015 Mandarin Hotel Bangkok by Centre Point, Bangkok, Thailand

757

ภาคผนวก ค

นำเสนอรายงานสืบเนื่องจากงานประชุมทางวิชาการระดับชาติ

การประชุมวิชาการระดับชาติ “วิทยาศาสตร์วิจัย” ครั้งที่ 7

วันที่ 30 – 31 มีนาคม 2558 มหาวิทยาลัยนเรศวร จังหวัดพิษณุโลก ประเทศไทย

เทพพนม อารัญญาณ จรรย์ รุ่งเรือง เหมือนฝัน โวหารกล้า และพิทักษ์ สุตรอนันต์. (2558). เครือข่ายความสัมพันธ์การแสดงออกร่วมของกลุ่มยีนในเซลล์แมโครฟาจ RAW 264.7 เมื่อถูกกระตุ้นด้วย LPS. รายงานสืบเนื่องจากการประชุมวิชาการระดับชาติ “วิทยาศาสตร์วิจัย” ครั้งที่ 7, วันที่ 30-31 มีนาคม 2558, มหาวิทยาลัยนเรศวร จังหวัดพิษณุโลก, BI-P-048, 6 หน้า.

เครือข่ายความสัมพันธ์การแสดงออกร่วมของกลุ่มยีนในเซลล์แมคโครฟาจ RAW 264.7 เมื่อถูกกระตุ้นด้วย LPS

The genes co-expression network of LPS-stimulated macrophage RAW264.7

เทพพนม อารัญญานัน¹, จริญญา รุ่งเรือง¹, เหมือนฝัน โหวนกรัก² และพิทักษ์ สุตรอนันต์^{1*}
Teppanom Aranphanu¹, Jarun Rungrueng¹, Muanfan Wohankla² and Pitak Sootanan^{1*}
¹ ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา
² โปรแกรมวิทยาศาสตร์ชีวภาพ โครงการบัณฑิตศึกษา คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

บทคัดย่อ

เครือข่ายความสัมพันธ์การแสดงออกร่วมของยีนถูกนำมาใช้ในการศึกษากลไกการตอบสนองต่อการอักเสบเมื่อถูกกระตุ้นด้วย LPS ตามช่วงเวลาตั้งแต่ 3-8 ชั่วโมงในเซลล์แมคโครฟาจสายพันธุ์ RAW264.7 สร้างได้จากความสัมพันธ์ระหว่างการแสดงออกของยีนจากข้อมูลไมโครอาร์เรย์ที่อยู่ในฐานข้อมูลสาธารณะ GEO (Gene Expression Onibus) ร่วมกับการคัดเลือกยีนตามพารามิเตอร์ที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบ เพื่ออธิบายความสัมพันธ์ระหว่างยีนในเครือข่ายการแสดงออกร่วมของยีน พบว่าการสร้างเครือข่ายความสัมพันธ์การแสดงออกร่วมของยีนสามารถตรวจพบกลุ่มยีนที่มีความสำคัญต่อการอักเสบที่ไม่พบจากการคัดเลือกยีนด้วยวิธีแบบดั้งเดิม ทำให้สามารถนำมาอธิบายกลไกและกระบวนการที่เกี่ยวข้องกับการตอบสนองการอักเสบได้เข้าใจมากยิ่งขึ้น และนำไปใช้กับการศึกษาที่เกี่ยวข้องต่อไปได้

คำสำคัญ : แมคโครฟาจ / ไลโปโพลีแซคคาไรด์ / ข้อมูลไมโครอาร์เรย์ / เครือข่ายการแสดงออกร่วมของยีน

Abstract

Gene co-expression network is used to study the molecular mechanism response of inflammation in RAW264.7 macrophage cell line. The inflammation of macrophage cell line was induced by LPS over 3 to 8 hrs period. This network can be constructed with correlation of gene pairs in selected inflammatory pathways in microarray data that are retrieved from the public database GEO (Gene Expression Onibus). These results showed that the use of co-expression network can be identified some responsive inflammatory genes that cannot be detected with traditional tools. This help to better understand the mechanisms involved in the inflammatory response of LPS-stimulated RAW264.7 and the further interested studies.

Keywords : Macrophage / Lipopolysaccharide / Microarray data / Co-expression network

*Corresponding author. E-mail : pitak@buu.ac.th

1. บทนำ

การอักเสบ (inflammation) มีบทบาทสำคัญที่เป็นจุดเริ่มต้นในกระบวนการที่ก่อให้เกิดโรคมะเร็ง ต้นแบบในการศึกษาการตอบสนองต่อการอักเสบนั้นจะนิยมใช้เซลล์แมคโครฟาจสายพันธุ์ RAW264.7 ที่ถูกเหนี่ยวนำด้วย LPS ซึ่งเป็นสารพิษที่อยู่ในผนังเซลล์ของแบคทีเรียแกรมลบ เมื่อเซลล์แมคโครฟาจถูกเหนี่ยวนำด้วย LPS จะทำให้มีการตอบสนองผ่านทางทอลลิลไครีเซปเตอร์ 4 (toll-like receptor 4, TLR4) และส่งสัญญาณผ่านวิถีต่างๆ ทำให้เกิดการแสดงออกของยีนที่เกี่ยวข้องกับการอักเสบมากมาย (Comer *et al.*, 2006; Hammer *et al.*, 2010) ดังนั้นการเรียนรู้เพื่อค้นหากลุ่มยีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบในเซลล์แมคโครฟาจ จึงมีส่วนสำคัญที่ทำให้สามารถทำความเข้าใจกลไกการตอบสนองต่อการอักเสบที่เกิดขึ้นได้ (Nilsson *et al.*, 2006) การตรวจวัดปริมาณระดับการแสดงออกของยีนในระดับ mRNA ภายใต้สภาวะที่ศึกษาด้วยเทคโนโลยีไมโครอาร์เรย์จะช่วยให้เข้าใจภาพรวมของกลไกการอักเสบที่เกิดขึ้นภายในเซลล์ (Watson *et al.*, 1998) และมีการรวบรวมข้อมูลไมโครอาร์เรย์จากหลากหลายกลุ่มงานวิจัยไว้ในฐานข้อมูลสาธารณะ Gene Expression Omnibus (GEO) (Edgar *et al.*, 2002) ซึ่งสามารถนำมาใช้ในการศึกษาได้

การศึกษาการตอบสนองต่อการอักเสบที่เกิดขึ้นโดยการวิเคราะห์ข้อมูลไมโครอาร์เรย์ ขั้นตอนการคัดเลือกยีนเป็นขั้นตอนที่มีความสำคัญมาก เนื่องจากข้อมูลไมโครอาร์เรย์นั้นประกอบด้วยยีนจำนวนมาก ทำให้การวิเคราะห์มีความซับซ้อนและอาจส่งผลให้เกิดความผิดพลาดได้สูง จึงต้องมีการคัดเลือกเฉพาะกลุ่มยีนที่มีความสำคัญต่อสภาวะที่สนใจศึกษา (Selvaraj *et al.*, 2011) โดยวิธีการที่นิยมใช้ในการคัดเลือกยีน ได้แก่ SAM (Significance Analysis for Microarray) (Dziuda, 2010) แต่วิธีนี้จะคัดเลือกเฉพาะกลุ่มยีนที่มีการแสดงออกอย่างเด่นชัดเท่านั้น อาจส่งผลทำให้มีการสูญเสียข้อมูลหรือกลุ่มยีนที่มีการแสดงออกน้อยแต่มีความสำคัญต่อกลไกการอักเสบได้ หากมีการคัดเลือกกลุ่มยีนโดยใช้ข้อมูลพารามิเตอร์ที่เกี่ยวข้องกับกลไกการตอบสนองต่อการอักเสบ จะทำให้การคัดเลือกยีนมีความจำเพาะมากยิ่งขึ้น (Sootanan *et al.*, 2012) และนำข้อมูลไมโครอาร์เรย์ของยีนที่คัดเลือกมาวิเคราะห์ร่วมกับการสร้างเครือข่ายการแสดงออกร่วมของยีน (gene co-expression network; GCN) จะช่วยให้เกิดความเข้าใจภาพรวมกลไกการตอบสนองต่อการอักเสบที่เกิดขึ้นได้อย่างเป็นระบบมากยิ่งขึ้น (Mabbott *et al.*, 2013)

ในงานวิจัยครั้งนี้ได้นำข้อมูลการแสดงออกของยีนจากไมโครอาร์เรย์ที่เกี่ยวข้องกับการทดสอบการตอบสนองต่อการอักเสบของเซลล์แมคโครฟาจ RAW264.7 เมื่อกระตุ้นด้วย LPS ตามช่วงเวลา จากฐานข้อมูลสาธารณะ GEO (Edgar *et al.*, 2002) จำนวน 2 ชุดข้อมูล คือ GSE4712 (Comer *et al.*, 2006) และ GSE21841 (Hammer *et al.*, 2010) ซึ่งจะนำมาวิเคราะห์ร่วมกับการคัดเลือกยีนจากข้อมูลพารามิเตอร์ที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบ (Sootanan *et al.*, 2012) และการตรวจวัดค่าความสัมพันธ์ระหว่างยีนมาใช้ในการสร้างเครือข่ายการแสดงออกร่วมของยีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบ (Mabbott *et al.*, 2013) แล้วทำการแปลผลทางชีวภาพโดยการจัดกลุ่มเครือข่ายการแสดงออกร่วมของยีนตามข้อมูลพารามิเตอร์ที่เกี่ยวข้องกับกระบวนการอักเสบ

2. วิธีการ

2.1 การสืบค้นข้อมูลและเตรียมข้อมูลไมโครอาร์เรย์

ดาวน์โหลดชุดข้อมูลไมโครอาร์เรย์จากฐานข้อมูลสาธารณะ GEO (Edgar *et al.*, 2002) ซึ่งเป็นค่าการแสดงออกของยีนที่อยู่ในรูปแบบไฟล์ gene expression matrix ทั้งหมด 2 ชุดข้อมูล ได้แก่ GSE4712 (Comer *et al.*, 2006) และ GSE21841 (Hammer *et al.*, 2010) ที่อยู่ภายใต้แพลตฟอร์มเดียวกันคือ Affymetrix Mouse Genome 430 2.0 โดยชุดข้อมูลไมโครอาร์เรย์ที่ได้นำมาวิเคราะห์นั้น จะคัดเลือกมา 2 กลุ่มการทดลองคือ กลุ่มที่ถูกกระตุ้นด้วย LPS และกลุ่มที่ไม่ถูกกระตุ้นด้วย LPS ตามช่วงเวลา ในเซลล์แมคโครฟาจเซลล์ไลน์สายพันธุ์ RAW264.7 ของหนู ในการทดลองไมโครอาร์เรย์นั้นอาจมีการทำการทดลองซ้ำ ซึ่งจะทำให้ 1 ยีนนั้นมีหลายโพรบ จึงต้องทำการปรับข้อมูลไมโครอาร์เรย์ให้อยู่ในรูปแบบ 1 ยีน 1 โพรบ โดยการเลือกโพรบที่มีค่าเฉลี่ยของการแสดงออกสูงสุด จากในชุดข้อมูลไมโครอาร์เรย์มีทั้งหมด 45,101 โพรบ มีจำนวนยีน 21,678 ยีน จากนั้นทำการปรับมาตรฐานค่าการแสดงออกของยีนในแต่ละชุดข้อมูลให้อยู่ในรูปแบบเดียวกันโดยใช้การคำนวณค่าลอการิทึมฐานสิบ และปรับค่าการแสดงออกของยีนให้อยู่ในรูปของค่าโพลด์เชนจ์ (fold change)

2.2 การคัดเลือกข้อมูลและการจัดกลุ่มการแสดงออกของยีนตามพารามิเตอร์

งานวิจัยนี้จะใช้ข้อมูลพารามิเตอร์ในการคัดเลือกและจัดกลุ่มยีน ดังนั้นจึงได้คัดเลือกพารามิเตอร์ที่เกี่ยวข้องกับกลไกการส่งสัญญาณการอักเสบมาทั้งหมด 8 พารามิเตอร์ ได้แก่ Chemokine signaling pathway, Cytokine-cytokine receptor interaction, Jak-Stat signaling pathway, Nod-Like receptor pathway, Notch signaling pathway, P53 signaling pathway, Toll-like receptor pathway และ MAPK signaling pathway โดยดาวน์โหลดข้อมูลพารามิเตอร์จากฐานข้อมูลสาธารณะ MsigDB (Molecular Signatures Database v3.1, <http://www.broadinstitute.org/gsea/msigdb/index.jsp>) โดยเลือกเฉพาะข้อมูลของ KEGG และจัดกลุ่มการแสดงออกของยีนตามพารามิเตอร์

2.3 การตรวจหาความสัมพันธ์การแสดงผลของยีน

นำข้อมูลไมโครอาร์เรย์ที่เตรียมไว้ข้างต้นมาทำการตรวจหาความสัมพันธ์การแสดงผลของยีนที่เกี่ยวข้องกับการอักเสบ โดยใช้หลักการของ Pearson Correlation Analysis ภายใต้โปรแกรม Cladist (Fukuchi *et al.*, 2009) ซึ่งจะตรวจหาระดับความสัมพันธ์ของการแสดงผลของยีนที่ละคู่ยีนจากยีนที่มีอยู่ทั้งหมด จะได้กลุ่มยีนที่มีความสัมพันธ์ที่มีแบบแผนการแสดงผลร่วมกัน จากนั้นทำการคัดเลือกกลุ่มยีนที่มีค่าความสัมพันธ์ (*r*-value) โดยกำหนดค่า cut-off ที่ระดับ ± 0.9 (Mabboot *et al.*, 2013) และตัดค่าความสัมพันธ์ที่เหลือทิ้งทางเดียว แล้วนำไปสร้างเครือข่ายความสัมพันธ์การแสดงผลของยีนต่อไป

2.4 การสร้างเครือข่ายความสัมพันธ์การแสดงผลของยีนร่วมของยีน

นำข้อมูลความสัมพันธ์การแสดงผลของยีนที่ผ่านการคัดกรองข้อมูลมาแล้ว ร่วมกับค่าการแสดงผลของยีนที่ผ่านการคัดกรองจากข้อมูลพารามิเตอร์ที่เกี่ยวข้องกับการอักเสบ มาสร้างเครือข่ายความสัมพันธ์การแสดงผลของยีนที่เกี่ยวข้องกับการอักเสบโดยใช้โปรแกรม Cytoscape (Cline *et al.*, 2007) จะได้เครือข่ายการแสดงผลของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบในเซลล์แมคโครฟาจ เมื่อมีการกระตุ้นด้วย LPS ตามช่วงเวลา

2.5 การคัดเลือกกลุ่มย่อยของเครือข่ายแสดงความสัมพันธ์การแสดงผลของยีนร่วมของยีน

จากการสร้างเครือข่ายความสัมพันธ์การแสดงผลของยีนจะได้เครือข่ายความสัมพันธ์ของกลุ่มยีนที่มีขนาดใหญ่ เพื่อให้การวิเคราะห์และแปลผลมีความครอบคลุม จึงได้มีการคัดเลือกกลุ่มย่อยของเครือข่ายแสดงความสัมพันธ์ของกลุ่มยีน โดยใช้คำสั่ง JactiveModules (Ideker *et al.*, 2002) ซึ่งเป็นปลั๊กอินที่บรรจุภายในโปรแกรม Cytoscape (Cline *et al.*, 2007) ผลที่ได้จะออกมาเป็นโมดูล (module) ย่อยหลายโมดูลของแต่ละช่วงเวลาที่มีการกระตุ้นด้วย LPS จากนั้นทำการคัดเลือกตัวแทนโมดูลย่อยที่สนใจมา 1 โมดูล โดยคัดเลือกตัวแทนโมดูลที่มีแบบแผนการแสดงผลของยีนอย่างเด่นชัด และเป็นเครือข่ายที่ไม่มีความซับซ้อนมากจะทำให้ง่ายต่อการวิเคราะห์และแปลผล แล้วทำการแปลผลโดยการจัดกลุ่มเครือข่ายการแสดงผลของยีนตามข้อมูลพารามิเตอร์ที่เกี่ยวข้องกับกระบวนการอักเสบ

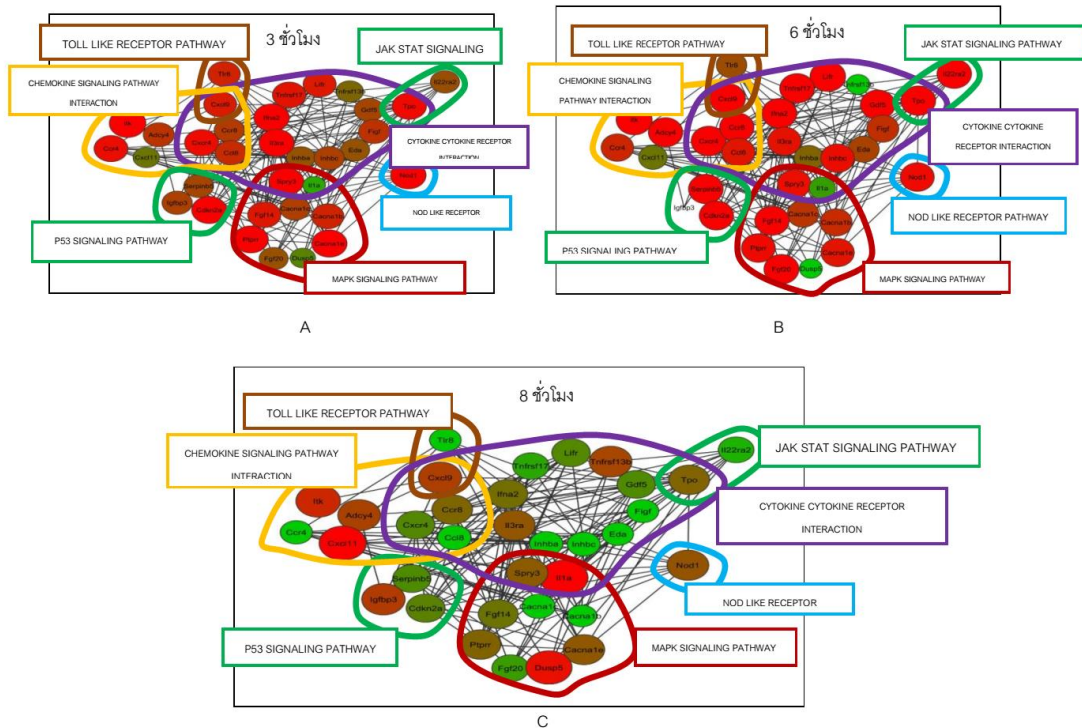
3. ผลและอภิปราย

งานวิจัยนี้ได้นำข้อมูลไมโครอาร์เรย์ การคัดเลือกยีนตามพารามิเตอร์ และการสร้างเครือข่ายความสัมพันธ์การแสดงผลของยีน เพื่อค้นหาและตรวจสอบความสัมพันธ์ระหว่างยีนที่ตอบสนองต่อการอักเสบในเซลล์แมคโครฟาจ RAW 264.7 เมื่อถูกกระตุ้นด้วย LPS ตามช่วงเวลา พบว่าการคัดเลือกกลุ่มยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบ โดยการคัดเลือกตามพารามิเตอร์ที่เกี่ยวข้องกับการตอบสนองการอักเสบทั้งหมด 8 พารามิเตอร์ การคัดเลือกจากค่าความสัมพันธ์ (*r*-value) และการคัดเลือกจากค่าโพลีดีเซนจ์ จะช่วยคัดกรองกลุ่มยีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบได้อย่างครอบคลุม ดังแสดงผลในตารางที่ 1 เนื่องจากการวิเคราะห์ข้อมูลไมโครอาร์เรย์ ขั้นตอนการคัดเลือกยีนเป็นขั้นตอนที่มีความสำคัญมาก เพราะข้อมูลไมโครอาร์เรย์นั้นประกอบด้วยยีนจำนวนมาก ทำให้การวิเคราะห์มีความซับซ้อนและอาจส่งผลให้เกิดความผิดพลาดได้สูง จึงต้องมีการคัดเลือกเฉพาะกลุ่มยีนที่มีความสำคัญต่อสถานะที่สนใจศึกษา (Selvaraj *et al.*, 2011) และจากการสร้างเครือข่ายความสัมพันธ์การแสดงผลของยีน ซึ่งคู่ของยีนที่ถูกคัดเลือกนั้นจะมีแบบแผนความสัมพันธ์การแสดงผลร่วมกัน ทำให้ขนาดของเครือข่ายและจำนวนยีนลดลงซึ่งง่ายต่อการวิเคราะห์และแปลผลการทำงานที่มีร่วมกันของยีนในแต่ละกลไกที่เกี่ยวข้องกับการตอบสนองการอักเสบได้ (Medzhitov and Horng, 2009 และ Nilsson *et al.*, 2006) ซึ่งในบทความวิจัยของ Le และคณะ (2013) ก็ได้ทำการศึกษาและวิเคราะห์เครือข่ายให้มีเครือข่ายที่เล็กลงเพื่อให้ง่ายต่อการแปลผลทางชีวภาพ

ตารางที่ 1 แสดงจำนวนยีน (โหนด) และจำนวนปฏิสัมพันธ์ (เอดจ์/ความสัมพันธ์) ในแต่ละขั้นตอนการทดลอง

ขั้นตอนการทดลอง	จำนวนยีน	จำนวนปฏิสัมพันธ์
ชุดข้อมูลไมโครอาร์เรย์	21,678	469,935,684
คัดเลือกยีนจาก 8 พารามิเตอร์ที่เกี่ยวข้องกับการตอบสนองการอักเสบ	659	434,281
คัดเลือกค่าโพลีดีเซนจ์ > 1.4 และตรวจหาความสัมพันธ์ในโปรแกรม Cladist	48	2,304
คัดเลือกจากค่า <i>r</i> -value 0.9 (+,-)	48	704
ตัดความสัมพันธ์จากไปกลับให้เหลือทิศทางเดียว	48	352
คัดเลือกเครือข่ายย่อยแต่ละช่วงเวลาโดย jActiveModules และตัดยีนที่มีชื่อและความสัมพันธ์ที่ซ้ำกันออก	34	155

จากการคัดเลือกเครือข่ายย่อยโดยใช้คำสั่ง jActiveModules และทำการคัดกรองยีนที่มีความสัมพันธ์ระหว่างคู่ยีนที่ซ้ำกันออกจะได้ยีนที่มีความสัมพันธ์ต่อการตอบสนองการอักเสบในช่วงเวลา 3, 6 และ 8 ชั่วโมง จำนวน 34 ยีน 155 ปฏิสัมพันธ์ (ตารางที่ 1) จากนั้นนำกลุ่มยีนที่ได้จากการคัดเลือกไปแปลผลโดยการจัดกลุ่มตามพารามิเตอร์ที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบทั้ง 8 พารามิเตอร์ เมื่อกระตุ้นด้วย LPS ตามช่วงเวลา 3, 6 และ 8 ชั่วโมง (รูปที่ 1) เครือข่ายความสัมพันธ์การแสดงออกของยีนแบ่งกลุ่มตามข้อมูลพารามิเตอร์ที่เกี่ยวข้องกับกระบวนการการอักเสบเมื่อถูกกระตุ้นด้วย LPS ที่เวลา 3 ชั่วโมง (รูปที่ 1A) ยกตัวอย่าง Jak-Stat signaling pathway จะพบว่ามีอยู่ 2 ยีน ได้แก่ *Tpo* และ *Il22ra2* ที่ถูกจัดกลุ่มให้อยู่ในพารามิเตอร์ หากสนใจศึกษา *Il22ra2* ที่มีการแสดงออกที่ไม่เด่นชัดแต่มีความสัมพันธ์และอยู่ในกลุ่มพารามิเตอร์เดียวกับ *Tpo* ที่มีการแสดงออกที่เพิ่มขึ้นเด่นชัด (up-regulated) ทำให้สามารถทำนายได้ว่ายีนที่สนใจน่าจะมีการทำงานร่วมกับ *Tpo* หรือมีการแสดงออกที่เกี่ยวข้องกับ Jak-Stat signaling pathway เมื่อกระตุ้นด้วย LPS ที่ช่วงเวลา 6 ชั่วโมง (รูปที่ 1B) ในทางกลับกันเมื่อกระตุ้นด้วย LPS ที่ช่วงเวลา 8 ชั่วโมง (รูปที่ 1C) จาก Jak-Stat signaling pathway เมื่อเวลาผ่านไปการแสดงออกของยีน ก็มีการแสดงออกที่เปลี่ยนไป โดยพบว่า *Il22ra2* จากที่ไม่มีการแสดงออกอย่างเด่นชัด แต่ที่เวลา 8 ชั่วโมงกลับมีการแสดงออกที่ลดลงอย่างเด่นชัด (down-regulated) แต่ *Tpo* จากที่มีการแสดงออกเพิ่มขึ้นอย่างเด่นชัดกลับมีการแสดงออกที่ไม่เด่นชัด แต่จากการแสดงออกที่ลดลงอย่างเด่นชัดของ *Il22ra2* ทำให้สามารถที่จะใช้เป็นตัวช่วยในการศึกษา *Tpo* ได้ ดังนั้นการสร้างเครือข่ายความสัมพันธ์ของการแสดงออกพร้อมและจัดกลุ่มการแสดงออกของยีนโดยใช้ข้อมูลพารามิเตอร์ จะช่วยในการอธิบายกระบวนการทางชีวภาพที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบ และช่วยในการคัดเลือกยีนที่มีระดับการแสดงออกการแสดงออกที่ไม่เด่นชัด แต่มีความสัมพันธ์กับกลุ่มยีนที่มีการแสดงออกอย่างเด่นชัดได้จำเพาะกับกลุ่มหน้าที่ทางชีวภาพที่สนใจได้มากขึ้น



รูปที่ 1 ผลของเครือข่ายความสัมพันธ์การแสดงออกของยีนแบ่งกลุ่มตามข้อมูลพารามิเตอร์ที่เกี่ยวข้องกับกระบวนการการอักเสบเมื่อถูกกระตุ้นด้วย LPS ตามช่วงเวลา 3, 6 และ 8 ชั่วโมง

4. บทสรุป

จากการวิเคราะห์ข้อมูลไมโครอาร์เรย์ การคัดเลือกยีนจากข้อมูลพารเวย์ที่เกี่ยวข้องกับการตอบสนองการอักเสบ และวิเคราะห์การแสดงออกร่วมของยีน ในเซลล์แมคโครฟาจ RAW 264.7 เมื่อถูกกระตุ้นด้วย LPS ตามช่วงเวลา พบว่ายีนที่ถูกคัดเลือกมีความสัมพันธ์การแสดงออกที่เด่นชัดจากการคัดเลือกที่มีค่าโฟลด์เชนจ์มากกว่า 1.4 และการคัดเลือกยีนด้วยข้อมูลพารเวย์ที่เกี่ยวข้องกับการตอบสนองการอักเสบ เมื่อนำมาสร้างเครือข่ายความสัมพันธ์การแสดงออกร่วมของยีนจะให้เห็นความสัมพันธ์ระหว่างยีนที่มีการแสดงออกต่อการตอบสนองการอักเสบ เมื่อถูกกระตุ้นด้วย LPS ที่เวลา 3, 6 และ 8 ชั่วโมง โดยใช้การจัดกลุ่มตามพารเวย์ที่เกี่ยวข้องกับการตอบสนองการอักเสบทั้ง 8 พารเวย์ เพื่อช่วยในการอธิบายรายละเอียดหรือช่วยในการคัดเลือกยีนได้จำเพาะมากขึ้น และพบว่า เมื่อกระตุ้นด้วย LPS ที่เวลา 3 และ 6 ชั่วโมง ส่วนมากความสัมพันธ์การแสดงออกร่วมของยีนในพารเวย์จะมีการแสดงออกเพิ่มขึ้น (up-regulated) และเมื่อกระตุ้นด้วย LPS ที่เวลา 8 ชั่วโมง จะมีความสัมพันธ์การแสดงออกร่วมของยีนในพารเวย์จะมีการแสดงออกลดลง (down-regulated) ซึ่งการตอบสนองของแต่ละพารเวย์จะมีการแสดงออกที่แตกต่างกันตามช่วงเวลา

5. กิตติกรรมประกาศ

ขอขอบพระคุณคณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา ที่ให้ทุนสนับสนุนการนำเสนอผลงานในงานวิจัย และงานวิจัยนี้ได้รับทุนอุดหนุนการวิจัย มหาวิทยาลัยบูรพา ประจำปีงบประมาณ พ.ศ.2557 สัญญาเลขที่ 87/2557

6. เอกสารอ้างอิง

- Cline, MS., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., Christmas, R., Campilo, IA., et al. (2007). Integration of biological networks and gene expression data using cytoscape. *Nature Protocols*, 2(10), 2366-2382.
- Comer, JE., Galindo, CL., Zhang, F., Wenglikowski, AM., Bush, KL., Garner, HR., et al. (2006). Murine macrophage transcriptional and functional responses to Bacillus anthracis edema toxin. *Microbial Pathogenesis*, 41(2-3), 96-110.
- Dziuda, DM. (2010). *Data Mining for Genomics and Proteomics: Analysis of Gene and Protein Expression Data*. John Wiley & Sons, Inc. Hoboken, New Jersey. 336 pp.
- Edgar, R., Domrachev, M. and Lash, AE. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, 30(1), 207-210.
- Fukuchi, S., Homma, K., Minezaki, Y., Gojobori, T. and Nishikawa, K. (2009). Development of an accurate classification system of proteins into structured and unstructured regions that uncovers novel structural domains its application to human transcription factors. *BMC Structural Biology*, 9, 26-35.
- Hammer, KDP., Yum, MY., Dixon, PM. And Birt, DF. (2010). Identification of JAK-STAT pathways as important for the anti-inflammatory activity of a *Hypericum perforatum* fraction and bioactive constituents in RAW 264.7 mouse macrophages. *Phytochemistry*, 71(7), 716-725.
- Ideker, T., Ozier, O., Schwikowski, B. and Siegel, AF. (2002). Discovering regulatory and signaling circuits in molecular interaction networks. *Bioinformatics*, 1, S233-S240.
- Le, TD., Liu, L., Liu, B., Tsykin, A., Goodall, GJ., Satou, K. and Li J. (2013). Inferring microRNA and transcription factor regulatory networks in heterogeneous data. *BMC Bioinformatics*, 14, 92-100.
- Mabbott, NA., Baillie, JK., Brown, H., Freeman, TC. and Hume, DA. (2013). An expression atlas of human primary cells inference of gene function from coexpression networks. *BMC Genomics*, 14, 632-645.
- Medzhitov, R. and Horng, T. (2009). Transcriptional control of the inflammatory response. *Nature Reviews Immunology*, 9(10), 692-703.
- Nilsson, R., Bajic, VB., Suzuki, H., Bernardo, DD., Björkegren, J., Katayama, S., et al. (2006). Transcriptional network dynamics in macrophage activation. *Genomics*, 88(2), 133-142.
- Selvaraj, S. and Natarajan, J. (2011). Microarray Data Analysis and Mining Tools. *Bioinformatics*, 6(3), 95-99.
- Sootanan, P., Chan, JH., Prom-on, P. and Meechai, A. (2012). Pathway-based microarray analysis for robust disease classification. *Neural Computing and Applications*, 21(4), 649-660.

Proceedings The 7th National Science Research Conference. 30-31 March 2015. Naresuan University.

Watson, A., Mazumder, A., Stewart, M. and Balasubramanian, S. (1998), Technology for microarray analysis of gene expression. *Curr Opin Biotechnol*, 9(6), 609-614.

ภาคผนวก ง

นำเสนอบทคัดย่อ งานประชุมทางวิชาการระดับนานาชาติ

The 1st International Conference of Pharmaceutical Sciences and Medicines,
22 – 23 January 2015, Burapha University, Chonburi, Thailand.

Wohankla, M. and Sootanan, P. 2015. “Inflammatory network inferred from microarray time series data of LPS-stimulated macrophage”. PP17. The 1st International Conference of Pharmaceutical Sciences and Medicines, 22 – 23 January 2015, Burapha University, Thailand.

PP17

INFLAMMATORY NETWORK INFERRED FROM MICROARRAY TIME SERIES DATA OF LPS-STIMULATED MACROPHAGE

Muanfan Wohankla¹ and Pitak Sootanan^{2*}

¹*Biological Science Program, Faculty of Science, Burapha University, Chonburi, Thailand*

²*Department of Biochemistry, Faculty of Science, Burapha University, Chonburi, Thailand*

*Corresponding author: pitak@buu.ac.th

Abstract

Network analysis is an alternatively effective method to select responsive genes and to systematically explain the inflammatory response mechanisms of LPS-stimulated macrophages. In this study, we analyzed microarray-time-course gene expression datasets involving LPS-response by RAW 264.7 macrophage cell line. These microarray datasets were retrieved from an open-access database and selectively analyzed to identify key inflammatory genes by using data mining techniques and network approaches to construct inflammatory network. Finally, our analysis could help for further studies to better understand the mechanism of inflammatory process in RAW264.7 macrophages cell line as well as providing new target candidates gene for inflammation therapeutics. Validated inflammatory response network may be applied to be a model for studying to discover the mechanism of anti-inflammatory effects of plant extracts and the other related studies.

Keywords: Microarray, RAW264.7 macrophages, Inflammatory network

ภาคผนวก จ.

นำเสนอรายงานสืบเนื่องจากงานประชุมทางวิชาการระดับชาติ

การประชุมวิชาการ การศึกษาทางพันธุศาสตร์และจีโนมิกส์ด้วยเครื่องวิเคราะห์ยุคใหม่

วันที่ 29 – 30 กรกฎาคม พ.ศ. 2557 กรุงเทพฯ ประเทศไทย

วันเฉลิม จินอู๋ และพิทักษ์ สุตรอนันต์. (2557). กลไกการตอบสนองการอักเสบจากการวิเคราะห์ข้อมูลไมโครอาร์เรย์. รายงานการประชุมวิชาการ การศึกษาทางพันธุศาสตร์และจีโนมิกส์ด้วยเครื่องวิเคราะห์ยุคใหม่, วันที่ 29-30 กรกฎาคม พ.ศ. 2557, กรุงเทพฯ, 25 – 31.

กลไกการตอบสนองการอักเสบจากการวิเคราะห์ข้อมูลไมโครอาร์เรย์ Inflammation Response Mechanism from Microarray Data Analysis

วันเฉลิม จินอู๋ และ พิทักษ์ สุตรอนันต์

Wanchalerm Ginou and Pitak Sootanan*

ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา จังหวัดชลบุรี รหัสไปรษณีย์ 20131 ประเทศไทย
Department of Biochemistry, Faculty of Science, Burapha University, Chonburi, Thailand, 20131

*Corresponding author: pitak@buu.ac.th

บทคัดย่อ

เทคโนโลยีไมโครอาร์เรย์ถูกนำมาใช้กันอย่างแพร่หลายในการศึกษากลไกที่เกี่ยวข้องกับการทดสอบฤทธิ์ของสารต้านการตอบสนองการอักเสบ โดยอาศัยเซลล์แมคโครฟาจ RAW264.7 ที่ถูกกระตุ้นด้วยไลโปพอลิแซ็กคาไรด์ (LPS) เป็นต้นแบบของการอักเสบ เราสามารถเข้าถึงข้อมูลไมโครอาร์เรย์เหล่านี้ได้จากฐานข้อมูลสาธารณะ GEO (Gene Expression Omnibus) โดยคัดเลือกเฉพาะกลุ่มตัวอย่าง 2 กลุ่ม คือกลุ่มควบคุม (control) กับกลุ่มที่มีการบ่มด้วย LPS (LPS-treated samples) ตามช่วงเวลาของแต่ละชุดข้อมูลไมโครอาร์เรย์ จากการรวบรวมข้อมูลไมโครอาร์เรย์จำนวน 3 ชุดข้อมูล ที่มีช่วงเวลากการบ่มด้วย LPS แตกต่างกัน 4 ช่วงเวลา ข้อมูลไมโครอาร์เรย์ที่จะนำมาจัดเตรียมและปรับมาตรฐาน ก่อนทำการคัดเลือกยีนที่เกี่ยวข้องกับการตอบสนองซึ่งถูกคัดเลือกโดย 2 วิธี จากทั้งหมด 3 วิธี ที่ใช้ซึ่งอยู่ในเครื่องมือการทำเหมืองข้อมูลออนไลน์ ArrayMining (<http://www.arraymining.net>) คือ SAM (Significance analysis for microarray) eBayes (Empirical Bayes moderated *t*-statistic) และ PLS-CV (Partial least squares cross-validation) แล้วนำยีนที่ได้จากการคัดเลือกมาทำการแปรผลทางชีวภาพผ่านโปรแกรมออนไลน์ DAVID (<http://david.abcc.ncifcrf.gov/>) โดยใช้ข้อมูล Gene ontology (GO) พบว่าฟังก์ชันการทำงานของยีนที่ถูกคัดเลือกมีความสอดคล้องตามช่วงเวลาของการตอบสนองการอักเสบของเซลล์แมคโครฟาจ RAW264.7 ที่ถูกกระตุ้นด้วย LPS จากงานวิจัยครั้งนี้ทำให้ได้เรียนรู้ว่าเราสามารถใช้อุปกรณ์ข้อมูลไมโครอาร์เรย์และโปรแกรมออนไลน์บนเครือข่ายสาธารณะมาช่วยในงานวิจัยทางชีววิทยาเชิงระบบเพื่อทำความเข้าใจกลไกการตอบสนองการอักเสบและกลไกอื่นที่สนใจต่อไปได้

ABSTRACT

Microarray technology is widely used in the study of the mechanisms associated with the effects of anti-inflammatory response. LPS-stimulated macrophage RAW264.7 is a favorite model of inflammation study. These microarray data can be retrieved from the public database GEO (Gene Expression Omnibus). Only two groups of samples such as control and LPS-treated cells in different time were selected. There are three microarray datasets with 4 different incubation times. Selected microarray data were prepared before selection of responsive genes by using 2 of 3 methods which are add-in of online data mining tool, ArrayMining (<http://www.arraymining.net>). These chosen 3 methods are SAM (Significance analysis for microarray), eBayes (Empirical Bayes moderated *t*-statistic) and PLS-CV (Partial least squares cross-validation). The selected genes from different time period were interpreted in biological function by using functional annotation tool, DAVID (<http://david.abcc.ncifcrf.gov/>) with gene ontology (GO) information. We

found that the function of selected genes in each different incubation time was related to the inflammatory response mechanism of LPS-stimulated macrophage. In summary, there are many useful microarray data and online tools to support research in the field of systems biology for more understanding inflammatory response and the other biological mechanisms.

คำสำคัญ: การอักเสบ, เซลล์แมคโครฟาจ, ลิโปพอลิแซ็กคาไรด์, ข้อมูลไมโครอาร์เรย์, เครื่องมือการทำเหมืองข้อมูล

Keywords: Inflammation, Macrophage, Lipopolysaccharide, Microarray data, Data mining tools

บทนำ

การอักเสบ (inflammation) เป็นกระบวนการที่ร่างกายตอบสนองต่อสิ่งที่ทำให้เนื้อเยื่อของร่างกายได้รับบาดเจ็บ เช่น เชื้อโรค การตายของเซลล์ เป็นต้น หากไม่มีกระบวนการอักเสบเกิดขึ้น ร่างกายจะไม่สามารถกำจัดสิ่งแปลกปลอมนั้นออกไปได้ เนื้อเยื่อเกิดการบาดเจ็บโดยที่ไม่มีการซ่อมแซม ทำให้การทำงานของเนื้อเยื่อดังกล่าวนั้นผิดปกติไป แต่หากเกิดการอักเสบมากเกินไป จะทำลายเนื้อเยื่อนั้นทำให้เกิดการผิดปกติได้เช่นกัน (พีรยุทธ, 2552) การศึกษากลไกการตอบสนองการอักเสบอย่างเป็นระบบโดยเทคโนโลยีไมโครอาร์เรย์ (Watson *et al.*, 1998) จึงมีความสำคัญต่อการทำความเข้าใจ การศึกษากลไกการตอบสนองการอักเสบส่วนใหญ่มักนิยมใช้เซลล์ไลน์แมคโครฟาจ RAW264.7 ที่ถูกกระตุ้นด้วยลิโปพอลิแซ็กคาไรด์ (Lipopolysaccharide; LPS) เป็นต้นแบบของการศึกษาการตอบสนองการอักเสบภายในเซลล์ (Shell *et al.*, 2005; Comer *et al.*, 2006; Hammer *et al.*, 2010) โดยที่ข้อมูลไมโครอาร์เรย์เหล่านั้นถูกเก็บไว้ภายในฐานข้อมูลสาธารณะ Gene Expression Omnibus (GEO) (Edgar *et al.*, 2002) ซึ่งผู้ที่สนใจสามารถนำดาวนโหลดออกมาใช้งานได้

การตรวจสอบกลไกการตอบสนองการอักเสบที่เกิดขึ้นสามารถทำได้โดยอาศัยการวิเคราะห์ข้อมูลไมโครอาร์เรย์ โดยการคัดเลือกกลุ่มยีนที่เกี่ยวข้องสามารถทำได้โดยการใช้เครื่องมือการทำเหมืองข้อมูลออนไลน์ ArrayMining (<http://www.arraymining.net/>) (Glaab *et al.*, 2009) ซึ่งเป็นแหล่งรวมวิธีการทำเหมืองข้อมูลอยู่หลากหลายชนิด ตัวอย่างเช่น SAM (Significance analysis for microarray) (Tusher *et al.*, 2001), eBayes (Empirical Bayes moderated *t*-statistic) (Smyth, 2004) และ PLS-CV (Partial least squares cross-validation) (Hall, 2000) โดยการคัดเลือกยีนในแต่ละวิธีก็จะมีกระบวนการในการคำนวณและการพิจารณาตัวแปร หรือความสัมพันธ์ระหว่างข้อมูลของแต่ละยีนที่แตกต่างกันไป ดังนั้นหากกลุ่มยีนถูกคัดเลือกด้วยวิธีการตั้งแต่ 2 วิธีขึ้นไป กลุ่มยีนนั้นก็จะเป็นกลุ่มยีนที่น่าเชื่อถือซึ่งเหมาะสำหรับนำมาใช้ในการอธิบายความหมายเชิงชีวภาพที่เกี่ยวข้องต่อไปได้

ในงานวิจัยครั้งนี้จะทำการวิเคราะห์ข้อมูลไมโครอาร์เรย์การตอบสนองของเซลล์ไลน์แมคโครฟาจ RAW 264.7 ต่อการกระตุ้นของ LPS ซึ่งได้จากฐานข้อมูลสาธารณะ GEO (Edgar *et al.*, 2002) จำนวน 3 ชุดข้อมูล คือ GSE2002 (Shell *et al.*, 2005), GSE4712 (Comer *et al.*, 2006) และ GSE21841 (Hammer *et al.*, 2010) ซึ่งจะถูกจัดแบ่งข้อมูลตามแต่ละช่วงเวลาที่ได้ศึกษาได้แก่ 3, 6, 8 และ 18 ชั่วโมง ตามลำดับ โดยจะทำการคัดเลือกยีนด้วยเครื่องมือการทำเหมืองข้อมูลทั้ง 3 วิธี คือ SAM (Tusher *et al.*, 2001), eBayes (Smyth, 2004) และ PLS-CV (Hall, 2000) ซึ่งเป็นวิธีการคัดเลือกยีนภายใต้เครื่องมือ ArrayMining (<http://www.arraymining.net/>) (Glaab *et al.*, 2009) แล้วนำกลุ่มยีนที่ถูกคัดเลือกอย่างน้อย 2 วิธี ไปทำการแปลผลทางชีวภาพโดยใช้ Gene Ontology (GO) (The Gene Ontology Consortium, 2000) จากฐานข้อมูลสาธารณะ DAVID (<http://david.abcc.ncifcrf.gov/>) (Huang *et al.*, 2009) เพื่อค้นหาว่ากลุ่มยีนที่ถูกคัดเลือกในแต่ละช่วงเวลาเกี่ยวข้องกับการทำงานต่อการตอบสนองการอักเสบอย่างไร สำหรับนำมาใช้ในการอธิบายกลไกที่เกิดขึ้นของการตอบสนองการอักเสบตามช่วงเวลา

อุปกรณ์และวิธีการ

1. การเตรียมข้อมูลไมโครอาร์เรย์

ข้อมูลที่ต้องการนั้นเป็นข้อมูลไมโครอาร์เรย์การแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบที่เกิดจากการกระตุ้นเซลล์แมคโครฟาจ RAW 264.7 ด้วย LPS โดยข้อมูลที่ใช้ในงานวิจัยนี้ประกอบด้วย 3 ชุดข้อมูล คือ GSE2002 (Shell *et al.*, 2005), GSE4712 (Comer *et al.*, 2006) และ GSE21841 (Hammer *et al.*, 2010) ที่ได้มาจากฐานข้อมูลสาธารณะ GEO (<http://www.ncbi.nlm.nih.gov/geo>) (Edgar *et al.*, 2002) ซึ่งสามารถแบ่งข้อมูลออกเป็น 4 กลุ่มตามช่วงเวลาคือ GSE4712(3hr), GSE4712(6hr), GSE21841(8hr) และ GSE2002(18hr) ข้อมูลที่เลือกใช้เป็นแพลตฟอร์มของ Affymetrix ที่ผ่านการทำ pre-process ให้อยู่ในรูปของค่าระดับการแสดงออกแล้ว แต่เนื่องจากระดับการแสดงออกของแต่ละโพรบ (probe) มีความแตกต่างกันมาก จึงจำเป็นต้องปรับระดับค่าการแสดงออกของข้อมูลโดยใช้ลอการิทึมฐานสิบ (log base 10) และปรับมาตรฐานค่าการแสดงออกโดยเปลี่ยนให้เป็น z-score ซึ่งมีค่าเฉลี่ย (mean) และค่าความแปรปรวน (standard deviation) ของทุกตัวอย่างเท่ากับ ศูนย์ และหนึ่ง ตามลำดับ แล้วนำข้อมูลที่ได้มาปรับให้สามารถใช้ได้กับเครื่องมือการทำเหมืองข้อมูลออนไลน์ ArrayMining โดยการใส่ชนิดของกลุ่มตัวอย่าง (outcome) ทั้ง 2 กลุ่ม คือตัวอย่างควบคุมที่ไม่มีการบ่ม (control) กับตัวอย่างที่มีการบ่มด้วย LPS (LPS-treat) เพื่อให้เครื่องมือสามารถเรียนรู้ได้ว่าตัวอย่างที่ให้อยู่ในกลุ่มใด

2. การคัดเลือกโพรบ (ยีน)

เนื่องจากข้อมูลไมโครอาร์เรย์อยู่ในรูปแบบของระดับการแสดงออกของโพรบ ในการคัดเลือกนี้จึงเป็นการคัดเลือกโพรบที่มีระดับการแสดงออกที่สามารถบ่งบอกความแตกต่างของ 2 กลุ่มตัวอย่าง ในทั้ง 4 ชุดข้อมูล ออกจากกันได้ โดยจะเลือกใช้วิธีการทั้งสิ้น 3 วิธี ในเครื่องมือการทำเหมืองข้อมูล ArrayMining (<http://www.arraymining.net/>) (Glaab *et al.*, 2009) ที่ประกอบไปด้วย SAM (Significance analysis for microarray) (Tusher *et al.*, 2001), eBayes (Empirical Bayes moderated *t*-statistic) (Smyth, 2004) และ PLS-CV (Partial least squares cross-validation) (Hall, 2000) จากชุดข้อมูลที่ใช้ทั้ง 4 ชุดข้อมูลคือ GSE4712(3hr), GSE4712(6hr), GSE21841(8hr) และ GSE2002(18hr) จะทำให้ได้กลุ่มยีนที่ถูกคัดเลือกอย่างน้อย 2 วิธี ซึ่งจะมีจำนวนและชนิดแตกต่างกันไปในแต่ละชุดข้อมูล (Table 1) ในที่นี้จะคัดเลือกเพียง 100 ลำดับแรกของโพรบ (top 100) มาใช้ในการวิเคราะห์

3. การแปลผลทางชีวภาพ

จากกลุ่มของโพรบที่ถูกคัดเลือกมาด้วยวิธีคัดเลือกตั้งแต่ 2 วิธี ในแต่ละชุดข้อมูลทั้ง 4 ชุดข้อมูล ไปทำการแปลผลทางชีวภาพโดยใช้ Gene Ontology (GO) (The Gene Ontology Consortium, 2000) จากฐานข้อมูลสาธารณะ DAVID (<http://david.abcc.ncifcrf.gov/>) (Huang *et al.*, 2009) ซึ่งจะทำการแปลผลทางชีวภาพของกลุ่มยีนตาม 3 ลักษณะของ GO คือ Biological Process (BP), Cell Component (CC), และ Molecular Function (MF) ที่จะให้ทราบว่ากลุ่มยีนที่ถูกคัดเลือกในแต่ละช่วงเวลานั้น เกี่ยวข้องกับกระบวนการทางชีวภาพ ตำแหน่งการทำงาน หรือฟังก์ชันระดับโมเลกุลที่เกี่ยวข้องต่อการตอบสนองการอักเสบอย่างไรโดยอาศัยการเปรียบเทียบกับข้อมูลพาธเวย์ของ TLR signaling pathway จากฐานข้อมูล BioCarta (<http://www.biocarta.com/>)

ผลการทดลองและวิจารณ์

1. การคัดเลือกโพรบ

จำนวนของโพรบที่ได้จากการคัดเลือกของทั้ง 3 วิธี จำนวนโพรบที่ถูกคัดเลือกอย่างน้อย 2 วิธี และจำนวนของยีนที่ตรวจสอบได้ฐานข้อมูล DAVID (<http://david.abcc.ncifcrf.gov/>) (Huang *et al.*, 2009) ตามชนิดของชุดข้อมูลไมโครอาร์เรย์ทั้ง 4 ชุดข้อมูล แสดงใน Table 1 การคัดเลือกโพรบโดยใช้เครื่องมือการทำเหมืองข้อมูลออนไลน์ ArrayMining (<http://www.arraymining.net/>) (Glaab *et al.*, 2009) ด้วยวิธี eBayes (Smyth, 2004) และ PLS-CV (Hall, 2000) จะได้จำนวนโพรบที่มากกว่าวิธี SAM (Tusher *et al.*, 2001) เนื่องจากการคัดเลือกโพรบด้วย eBayes

(Smyth, 2004) และ PLS-CV (Hall, 2000) อาศัยการจัดอันดับ (rank) ความสำคัญของโพรบที่ทำให้ได้จำนวนของโพรบเท่ากับ 100 ซึ่งเป็นจำนวนที่กำหนด ส่วนวิธี SAM จะเป็นค้นหาโพรบที่มีการแสดงออกแตกต่างกันอย่างมีนัยสำคัญ (Tusher *et al.*, 2001) และเนื่องจากกระบวนการวิเคราะห์ที่แตกต่างกันของทั้ง 3 วิธี ทำให้จำนวนของโพรบที่ถูกคัดเลือกอย่างน้อย 2 วิธีมีจำนวนที่น้อยลง แต่ก็ทำให้มีความน่าเชื่อถือของโพรบหรือยีนที่ถูกคัดเลือกเพิ่มมากขึ้นเพื่อนำไปใช้ในการแปลผลและค้นหาวิถีทางชีวภาพที่เกี่ยวข้องต่อไป

Table 1 Number of probes selected by 3 different methods (SAM, eBayes, and PLS-CV), number of probes selected by 2 of 3 methods and number of genes found in DAVID (<http://david.abcc.ncifcrf.gov/>)

ชุดข้อมูล	SAM (โพรบ)	eBayes (โพรบ)	PLS-CV (โพรบ)	โพรบที่พบอย่างน้อย 2 วิธี (โพรบ)	Official Gene Symbol จาก DAVID (ยีน)
GSE4712(3hr)	5	100	100	30	26
GSE4712(6hr)	100	100	100	65	57
GSE21841(8hr)	20	100	100	50	39
GSE2002(18hr)	100	100	100	96	96

2. การแปลผลทางชีวภาพ

นำโพรบที่ถูกคัดเลือกอย่างน้อย 2 วิธี มาทำการแปลผลทางชีวภาพโดยใช้ฐานข้อมูลสาธารณะ DAVID (<http://david.abcc.ncifcrf.gov/>) (Huang *et al.*, 2009) ซึ่งจะพบจำนวนยีนตาม **Table 1** ซึ่งเป็นกลุ่มยีนที่ถูกคัดเลือกที่มีความสำคัญต่อการตอบสนองการอักเสบของเซลล์แมคโครฟาจ RAW264.7 ที่ถูกกระตุ้นด้วย LPS และสามารถแปลผลทางชีวภาพดัง **Figure 1** ซึ่งอธิบายกลไกการแสดงออกของยีนในแต่ละช่วงเวลา ได้แก่ ในช่วงเวลา 3 ชั่วโมง ยีนในกลุ่ม cell adhesion ที่อยู่ในชุดของข้อมูล GSE4712(3hr) ที่เกี่ยวข้องกับกลไกที่ชักนำให้เซลล์เม็ดเลือดขาวเข้ามาในเนื้อเยื่อบริเวณที่เกิดการอักเสบจะมีการแสดงออกมาก ส่วนที่เวลา 6 ชั่วโมง จากชุดข้อมูล GSE4712(6hr) กลุ่มของยีนที่พบส่วนใหญ่จะมีผลผลิตโปรตีนที่ทำงานภายใต้พื้นฐานของการถอดรหัสและแปลรหัสของชุดข้อมูล DNA และ RNA ที่เกี่ยวข้องกับการรับส่งสัญญาณในกระบวนการเกิดการอักเสบที่ถูกกระตุ้นด้วย LPS ที่เวลา 8 ชั่วโมง จากชุดข้อมูล GSE21841(8hr) จะพบยีนในกลุ่ม cytokine activity, cytokine-mediated signaling pathway, positive regulation of MAPKKK cascade, regulation of MAPKKK cascade ที่มีความสำคัญในการควบคุมกระบวนการตอบสนองการอักเสบเกิดขึ้น และที่เวลา 18 ชั่วโมง จากชุดข้อมูล GSE2002(18hr) จะพบยีนในกลุ่มของ cell division, defense response, chemokine activity, cell death, death, apoptosis และ DNA repair ที่เกี่ยวข้องกับการรักษาตัวเองของเซลล์เมื่อการอักเสบสิ้นสุดลง สำหรับ **Figure 2** จะแสดงให้เห็นถึงการแปลผลทางชีวภาพที่ได้จากการคัดเลือกกลุ่มยีนในแต่ละชุดข้อมูลตามเวลาที่แตกต่างกัน เปรียบเทียบกับข้อมูลพารเวย์ของ TLR signaling pathway ที่ได้จากฐานข้อมูล BioCarta (<http://www.biocarta.com/>) ที่มีความสอดคล้องและมีแนวโน้มไปในทิศทางเดียวกัน

GSE4712(3hr)	GSE4712(6hr)	GSE21841(8hr)	GSE2002(18hr)
GOTERM_BP	GOTERM_BP	GOTERM_BP	GOTERM_BP
cell adhesion	regulation of transcription	acute-phase response	microtubule cytoskeleton organization
biological adhesion	monovalent inorganic cation transport	regulation of cytokine production	regulation of mitosis
cell-cell adhesion	transport	chemotaxis	inflammatory response
GOTERM_CC	metal ion transport	positive regulation of cell proliferation	blood vessel development
cell-cell adhesion junction	regulation of RNA metabolic process	negative regulation of transport	regulation of apoptotic process
adhesion junction	GOTERM_CC	positive regulation of MAPKKK cascade	negative regulation of cell death
anchoring junction	cation channel complex	regulation of kinase activity	regulation of protein kinase activity
Plasma membrane	ion channel complex	GOTERM_CC	proteolysis involved in cellular protein catabolic process
GOTERM_MF	plasma membrane	external side of plasma membrane	GOTERM_CC
calcium ion binding	integral to plasma membrane	intrinsic to membrane	kinetochore
metal ion binding	intrinsic to plasma membrane	Cytoskeleton	chromosome, centromeric region
cation binding	GOTERM_MF	non-membrane-bounded organelle	spindle
ion binding	metal ion binding	Intracellular non-membrane-bounded organelle	nucleolus
	transition metal ion binding	GOTERM_MF	GOTERM_MF
	zinc ion binding	cytokine activity	microtubule motor activity
	DNA binding	purine ribonucleotide binding	chemokine activity
		nucleotide binding	ATP binding
		Zinc ion binding	protein homodimerization activity
		transition metal ion binding	identical protein binding
			GTPase activity

Figure 1 Biological interpretation of 4 selected responsive genes in different microarray datasets by using GO in functional annotation tools, DAVID (<http://david.abcc.ncifcrf.gov/>).

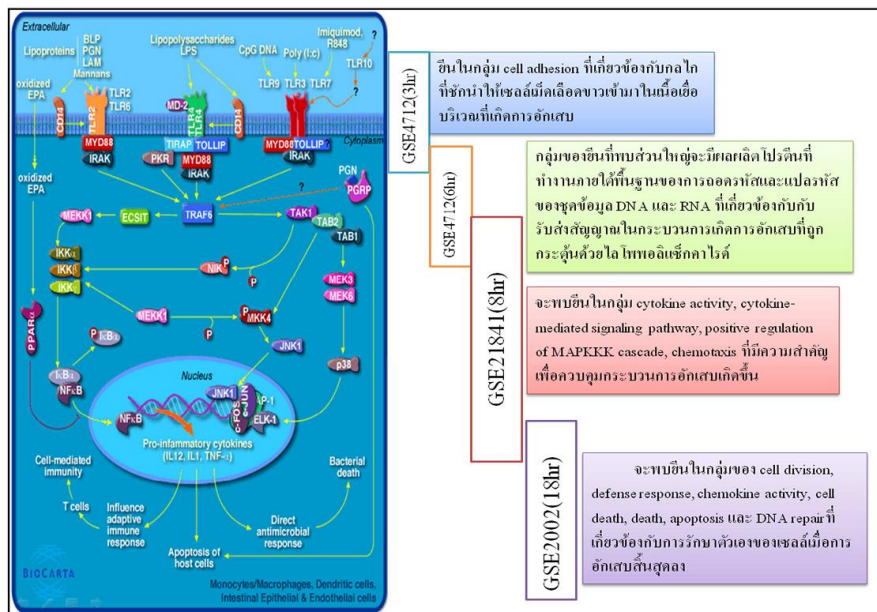


Figure 2 Functional preparation between 4 selected responsive genes in different microarray datasets and TIR signaling pathway from BioCarta (<http://www.biocarta.com/genes/index.asp>)

สรุปผลการทดลอง

การศึกษาการตอบสนองการอักเสบสามารถใช้ข้อมูลไมโครอาร์เรย์ภายใต้สภาวะของเซลล์แมคโครฟาจ RAW264.7 ที่มีและไม่มีการกระตุ้นด้วย LPS โดยใช้วิธีการคัดเลือกยีนที่อยู่ภายในเครื่องมือการทำเหมืองข้อมูลออนไลน์ ArrayMining (<http://www.arraymining.net/>) (Glaab *et al.*, 2009) ได้แก่วิธี SAM (Tusher *et al.*, 2000), eBayes (Smyth, 2004) และ PLS-CV (Hall, 2000) ซึ่งยีนที่ถูกคัดเลือกจาก 2 ใน 3 วิธีข้างต้น เมื่อนำมาแปลผลทางชีวภาพโดยใช้ Gene Ontology (GO) (The Gene Ontology Consortium, 2000) จากฐานข้อมูลสาธารณะ DAVID (<http://david.abcc.ncifcrf.gov/>) (Huang *et al.*, 2009) ซึ่งจะทำให้การแปลผลทางชีวภาพของกลุ่มยีนตาม 3 ลักษณะของ GO คือ Biological Process (BP), Cell Component (CC), และ Molecular Function (MF) ก็จะสามารถนำมาอธิบายกลไกการตอบสนองการอักเสบที่เกิดขึ้นตามช่วงเวลาได้ จากวิธีการข้างต้นสามารถนำไปประยุกต์ใช้ในการอธิบายกลไกระดับโมเลกุลเพื่อการวิเคราะห์ข้อมูลไมโครอาร์เรย์ที่เกี่ยวข้องกับลักษณะทางชีวภาพอื่นที่สนใจต่อไปได้

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนสนับสนุนการวิจัยจากมหาวิทยาลัยบูรพา ประจำปีงบประมาณ 2557 สัญญาเลขที่ 87/2557

เอกสารอ้างอิง

- พีรยุทธ ลิทธิไชยากุล (2552) Acute and Chronic Inflammation. เอกสารประกอบการสอนรายวิชาหลักพยาธิวิทยาและนิติเวชศาสตร์, ภาควิชาพยาธิวิทยาและนิติเวชศาสตร์ คณะแพทยศาสตร์ มหาวิทยาลัยนเรศวร
- Comer JE, Galindo CL, Zhang F, Wenglikowski AM, Bush KL, Garner HR, Peterson JW, Chopra AK (2006) Murine macrophage transcriptional and functional responses to *Bacillus anthracis* edema toxin. *Microb Pathog* 41(2-3):96-110.
- Edgar R, Domrachev M, Lash AE. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30(1):207-10.
- Glaab E, Garibaldi JM, Krasnogor N (2009) ArrayMining: a modular web-application for microarray analysis combining ensemble and consensus methods with cross-study normalization. *BMC Bioinformatics* 10:358.
- Hall MA (2000) Correlation-based feature selection for discrete and numeric class machine learning. *ICML* 359-366
- Hammer KD, Yum MY, Dixon PM, Birt DF (2010) Identification of JAK-STAT pathways as important for the anti-inflammatory activity of a *Hypericum perforatum* fraction and bioactive constituents in RAW 264.7 mouse macrophages. *Phytochemistry* 71(7):716-25.
- Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4(1):44-57.
- Shell SA, Hesse C, Morris SM Jr, Milcarek C (2005) Elevated levels of the 64-kDa cleavage stimulatory factor (CstF-64) in lipopolysaccharide-stimulated macrophages influence gene expression and induce alternative poly(A) site selection. *J Biol Chem* 280(48):39950-61.
- Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3:Article3.

The Gene Ontology Consortium (2000) Gene ontology: tool for the unification of biology. *Nat Genet* 25(1): 25–9.

Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98(9):5116–21.

Watson A, Mazumder A, Stewart M, Balasubramanian S (1998) Technology for microarray analysis of gene expression. *Curr Opin Biotechnol* 9(6):609–616.

ArrayMining, <http://www.arraymining.net> (June 2014)

BioCarta, <http://www.biocarta.com/> (June 2014)

DAVID, <http://david.abcc.ncifcrf.gov/> (June 2014)

Gene Expression Omnibus (GEO), <http://www.ncbi.nlm.nih.gov/geo> (June 2014)

ภาคผนวก ฉ
บทความวิจัย วารสารระดับชาติ

วารสารวิทยาศาสตร์บูรพา ปีที่ 19 ฉบับพิเศษ
การประชุมวิชาการระดับชาติ “วิทยาศาสตร์วิจัย” ครั้งที่ 6

พศิกา ไบยา วิไลพร ศรีตะบุตร และพิทักษ์ สุตรอนันต์. (2557). เครือข่ายควบคุมการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบจากข้อมูลไมโครอาร์เรย์. วารสารวิทยาศาสตร์บูรพา ปีที่ 19 ฉบับพิเศษ การประชุมวิชาการระดับชาติ วิทยาศาสตร์วิจัยครั้งที่ 6, ISBN 0858-7612, 173-184.

เครือข่ายควบคุมการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบ
จากข้อมูลไมโครอาร์เรย์

Gene regulatory network of inflammatory response from microarray data

พศิกา ไบยา, วิไลพร ศรีตะบุตร และ พัทภัสร์ สุตรอนันต์

Pasika Baiya, Wilaipon Sritaboot and Pitak Sootanan

ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

บทคัดย่อ

การอักเสบเป็นกระบวนการที่ร่างกายตอบสนองต่อสิ่งกระตุ้นภายนอกทำให้เกิดโรคในสิ่งมีชีวิต การศึกษากระบวนการควบคุมการแสดงออกของยีนที่เกี่ยวข้องกับการอักเสบ จะช่วยให้เราสามารถเข้าใจกลไกที่เกี่ยวข้องได้ดียิ่งขึ้น การศึกษานี้ผู้วิจัยได้นำข้อมูลไมโครอาร์เรย์ที่เกี่ยวข้องกับการตอบสนองการอักเสบของเซลล์แมคโครฟาจหนู RAW 264.7 ที่ถูกเหนี่ยวนำด้วย LPS จากฐานข้อมูลสาธารณะไปตรวจหากลุ่มยีนและทรานสคริปชันแฟกเตอร์ที่เกี่ยวข้องกับการตอบสนองการอักเสบ โดยอาศัยการสร้างเครือข่ายการควบคุมการแสดงออกของยีนและการคัดกรองข้อมูลที่เกี่ยวข้องโดยการสืบค้นวารสารวิจัย ผลที่ได้คือเครือข่ายที่มียีนจำนวน 5 ยีน คือ *MEF2A*, *NFKB1*, *NFKB2*, *NFATC1* และ *NFATC3* ที่สร้างทรานสคริปชันแฟกเตอร์ MEF2, NF- κ B และ NFAT ซึ่งควบคุมยีนเป้าหมายที่เกี่ยวข้องกับการตอบสนองการอักเสบจำนวน 64 ยีน ซึ่งมีการแสดงออกมากกว่า 1.4 เท่า และพบ 6 ยีนที่เป็นยีนตอบสนองการอักเสบ ซึ่งถูกควบคุมด้วยกลุ่มของทรานสคริปชันแฟกเตอร์ที่แตกต่างกันไปตามแต่ละช่วงของเวลา โดยมีทรานสคริปชันแฟกเตอร์ MEF2 ศูนย์กลางการควบคุมที่สำคัญของการแสดงออกของยีนตอบสนองการอักเสบ 4 ชนิด ในช่วงเวลา 3 และ 6 ชั่วโมง การศึกษานี้แสดงให้เห็นว่าเราสามารถนำข้อมูลไมโครอาร์เรย์มาใช้เพื่อตรวจสอบกลไกการควบคุมการแสดงออกของยีนที่เกี่ยวข้อง ซึ่งจะทำให้เราสามารถเรียนรู้และทำความเข้าใจกลไกการอักเสบในระดับโมเลกุลได้ดียิ่งขึ้น

คำสำคัญ : เครือข่ายควบคุมการแสดงออกของยีน / ไมโครอาร์เรย์ / การอักเสบ / ทรานสคริปชันแฟกเตอร์

*Corresponding author. E-mail : pitak@buu.ac.th

วารสารวิทยาศาสตร์บูรพา ฉบับพิเศษ การประชุมวิชาการระดับชาติ วิทยาศาสตร์วิจัย ครั้งที่ 6 วันที่ 20 – 21 มีนาคม พ.ศ. 2557

Abstract

Inflammation is a process of vascular tissues that response to harmful stimuli such as injury, pathogens or irritants. Study of gene expression regulatory network of inflammation leads to more understand of involved mechanisms. The use of microarray datasets of LPS-stimulated mouse macrophage RAW 256.7 and TF-TG (transcription factors and their target genes) data were used to search for the responsive TFs and their TGs by constructing of gene regulatory network and data filtering with literatures. The results found 5 genes such as *MEF2A*, *NFKB1*, *NFKB2*, *NFATC1* and *NFATC3* that encoded 3 TF proteins such as MEF2, NF-κB and NFAT. These regulated 64 target genes with more than 1.4 fold change. There were 6 inflammatory responsive genes regulated by various TFs according to time-series. TF protein MEF2 is the important hub to regulate the expression of four inflammatory genes at 3 and 6 hr. In addition, public microarray and TF-TG data can be used to discovered gene regulatory network to more understand inflammatory mechanisms.

Keywords : Gene regulatory network / Microarray / Inflammation / Transcription factor

1. บทนำ

การอักเสบ (inflammation) เป็นกระบวนการที่ร่างกายตอบสนองต่อสิ่งที่ทำให้เนื้อเยื่อของร่างกายได้รับบาดเจ็บ จากสิ่ง กระตุ้นภายนอก เซลล์ที่มีบทบาทต่อกระบวนการตอบสนองการอักเสบคือเซลล์แมคโครฟาจ (macrophage) ซึ่งเป็นเซลล์เม็ดเลือดขาวที่สำคัญในระบบภูมิคุ้มกันที่มีมาแต่กำเนิด โดยมีโปรตีนตัวรับจดจำรูปแบบ (pattern recognition receptors, PRRs) ที่จะคอยจดจำโครงสร้างของโมเลกุลสิ่งแปลกปลอม (pathogen-associated molecular pattern, PAMPs) ซึ่ง PRRs ที่สำคัญคือ โปรตีนตัวรับชนิดทอลล์ (toll-like receptors, TLRs) ซึ่งเป็นโปรตีนตัวรับที่ทำหน้าที่จับไลโปพอลิแซคคาไรด์ (lipopolysaccharide; LPS) (Nilsson *et al.*, 2006) การอักเสบเป็นส่วนสำคัญของกระตุ้นและการแสดงออกของยีน ตัวอย่างเช่น กลไกการส่งสัญญาณของตัวรับชนิดทอลล์ในเซลล์แมคโครฟาจที่ประกอบด้วยชุดควบคุมการแสดงออกที่หลากหลายของยีน ซึ่งถูกควบคุมได้โดยการทำงานของทรานสคริปชันแฟคเตอร์ การเปลี่ยนแปลงโครงสร้าง โครมาติน และการเกิดเมธิลเลชันบนสายดีเอ็นเอ (Medzhitov and Horng, 2009) ถ้าหากไม่มีกระบวนการอักเสบเกิดขึ้น ร่างกายจะไม่สามารถกำจัดสิ่งแปลกปลอมนั้นออกไปได้ เนื้อเยื่อจะเกิดการบาดเจ็บโดยที่ไม่มีการซ่อมแซมและทำให้การทำงานของเนื้อเยื่อเหล่านั้นเกิดความผิดปกติที่อาจจะพัฒนาต่อไปเป็นโรคที่เกิดจากการอักเสบเรื้อรังได้ (Hoesel and Schmid, 2013)

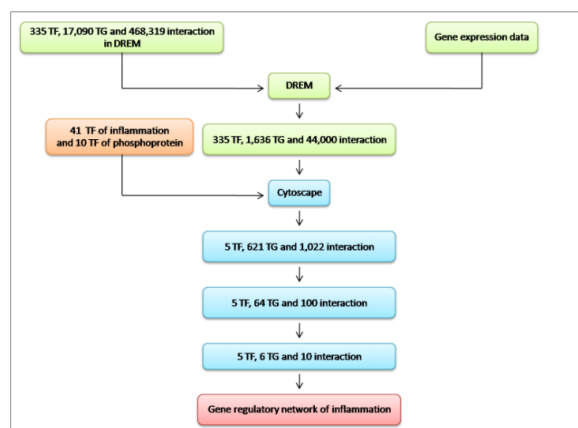
การทำความเข้าใจกลไกที่เกี่ยวข้องจากการตอบสนองการอักเสบที่เกิดขึ้นสามารถช่วยลบลบปัญหาที่เกิดขึ้นจากการอักเสบได้จากเหตุผลทางด้านเทคโนโลยีไมโครอาร์เรย์ได้ถูกนำมาใช้ในการตรวจสอบการแสดงออกของยีนจำนวนมากที่เกี่ยวข้องกับกระบวนการตอบสนองการอักเสบ โดยเฉพาะการตอบสนองของเซลล์แมคโครฟาจ RAW 264.7 ที่มีต่อไลโปพอลิแซคคาไรด์ (Comer *et al.*, 2006 และ Hammer *et al.*, 2010) ซึ่งมักนิยมใช้เป็นต้นแบบของการศึกษาการตอบสนองการอักเสบที่เกิดขึ้น และจากการพัฒนาข้อมูลระดับจีโนมของทรานสคริปชันแฟคเตอร์และยีนเป้าหมายที่ถูกควบคุมการแสดงออก (Kanamori *et al.*, 2004 และ Schulz *et al.*, 2012) ทำให้สามารถนำข้อมูลข้างต้นมาใช้ในการสร้างเครือข่ายการควบคุมการแสดงออกของยีนระดับทรานสคริปชันแฟคเตอร์ที่เกี่ยวข้องกับการตอบสนองการอักเสบ เพื่อเรียนรู้และทำความเข้าใจกลไกการควบคุมการแสดงออกของยีนที่เกี่ยวข้องได้ การตรวจสอบกลไกดังกล่าวขึ้นอยู่กับรูปแบบการศึกษาของข้อมูลไมโครอาร์เรย์ที่เลือกใช้ ทำให้สามารถตรวจสอบการแสดงออกของยีนตามช่วงเวลาหรือตรวจสอบตามโดสของยาที่สนใจ

ศึกษาได้ นอกจากนี้ผลการแสดงออกของยีนที่เพิ่มขึ้นหรือลดลงสามารถนำมาสร้างเครือข่ายการควบคุมการแสดงออกของยีนที่เกี่ยวข้องกับการอักเสบ และสามารถอธิบายกลไกการควบคุมการอักเสบเพื่อช่วยลดกระบวนการอักเสบของเซลล์ได้

ในการศึกษานี้ผู้วิจัยได้นำข้อมูลไมโครอาร์เรย์จากฐานข้อมูลสาธารณะ GEO (Gene Expression Omnibus) (Edgar *et al.*, 2002) ได้แก่ ชุดข้อมูลไมโครอาร์เรย์ GSE4712 (Comer *et al.*, 2006) และ GSE21841 (Hammer *et al.*, 2010) โดยจะทำการคัดเลือกเฉพาะข้อมูลของกลุ่มตัวอย่างที่เกี่ยวข้องกับการศึกษาการตอบสนองการอักเสบของเซลล์แมคโครฟาจที่มีต่อการกระตุ้นด้วย LPS มาผ่านกระบวนการเตรียมและปรับเทียบมาตรฐาน และการตรวจวัดค่าระดับการแสดงออกแบบโฟลด์เชนจ์ (flow change) แล้วนำข้อมูลที่ได้มาใช้ในการคัดเลือกทรานสคริปชันแฟกเตอร์และยีนเป้าหมายที่เกี่ยวข้องโดยการตรวจสอบร่วมกับข้อมูลระดับจีโนมของทรานสคริปชันแฟกเตอร์และยีนเป้าหมายในโปรแกรม Dynamic Regulatory Events Miner (DREM) (Schulz *et al.*, 2012) จากนั้นทำการสร้างเครือข่ายการควบคุมการแสดงออกของยีนของกลุ่มยีนที่มีระดับการแสดงออกมากกว่า 1.4 เท่า ในแต่ละช่วงเวลาเปรียบเทียบ และคัดเลือกเฉพาะทรานสคริปชันแฟกเตอร์ที่เกี่ยวข้องกับการตอบสนองการอักเสบ โดยอาศัยข้อมูลจากบทความวิจัยที่กล่าวถึงการตรวจวัดระดับโฟลด์โปรตีนที่บ่งบอกถึงการส่งผ่านสัญญาณที่เกี่ยวข้องกับการตอบสนองการอักเสบ (Weintz *et al.*, 2010) เพื่อนำมาใช้ในการอธิบายกลไกการตอบสนองการอักเสบที่เกิดขึ้นตามช่วงเวลาการศึกษา

2. วิธีการ

การสร้างเครือข่ายการควบคุมการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบ เพื่อตรวจสอบกลไกควบคุมการอักเสบในระดับการแสดงออกของยีนที่เกี่ยวข้องในเซลล์แมคโครฟาจ RAW 264.7 ที่ถูกเหนี่ยวนำด้วยไลโปพอลิแซคคาไรด์ โดยมีขั้นตอนดังแสดงในภาพที่ 1 ซึ่งประกอบด้วยวิธีการคัดเลือกทรานสคริปชันแฟกเตอร์และยีนเป้าหมายร่วมกับข้อมูลการแสดงออกของยีนโดย DREM (Schulz *et al.*, 2012) แล้วทำการคัดกรองทรานสคริปชันแฟกเตอร์จากงานวิจัยที่เกี่ยวข้องกับการอักเสบ และคัดกรองร่วมกับทรานสคริปชันแฟกเตอร์ที่เป็นโฟลด์โปรตีน จากนั้นทำการสร้างเครือข่ายการควบคุมการแสดงออกของยีนที่เกี่ยวข้องกับการอักเสบโดย Cytoscape (Cline *et al.*, 2007) ซึ่งมีรายละเอียดวิธีการในแต่ละขั้นตอนดังต่อไปนี้



ภาพที่ 1 แผนภาพแสดงขั้นตอนการคัดกรองทรานสคริปชันแฟกเตอร์

2.1 ข้อมูลไมโครอาร์เรย์และการเตรียมข้อมูล

ดาวน์โหลดชุดข้อมูลไมโครอาร์เรย์จากฐานข้อมูลสาธารณะ GEO (Gene Expression Omnibus, <http://www.ncbi.nlm.nih.gov/geo/>, Edgar *et al.*, 2002) ได้แก่ ชุดข้อมูลไมโครอาร์เรย์ GSE4712 (Comer *et al.*, 2006) และ GSE21841 (Hammer *et al.*, 2010) ซึ่งใช้ Affymetrix Mouse Genome 430 2.0 Array (GPL1261) เป็นแพลตฟอร์มของชิปไมโครอาร์เรย์เหมือนกัน

โดยทำการคัดเลือกเฉพาะตัวอย่างที่เกี่ยวข้อง จากชุดข้อมูลไมโครอาร์เรย์ GSE4712 มีจำนวนตัวอย่างทั้งหมด 21 ตัวอย่าง ซึ่งนำมาใช้จำนวน 9 ตัวอย่าง ประกอบไปด้วยตัวอย่างที่ไม่ถูกกระตุ้นด้วยไลโปพอลิแซคคาไรด์ 3 ตัวอย่าง ตัวอย่างที่ถูกกระตุ้นด้วยไลโปพอลิแซคคาไรด์เป็นเวลา 3 ชั่วโมง มีอยู่ 3 ตัวอย่าง และตัวอย่างที่ถูกกระตุ้นด้วยไลโปพอลิแซคคาไรด์เป็นเวลา 6 ชั่วโมง มีอยู่ 3 ตัวอย่าง ส่วนชุดข้อมูลไมโครอาร์เรย์ GSE21841 มีจำนวนตัวอย่างทั้งหมด 18 ตัวอย่าง นำมาใช้จำนวน 6 ตัวอย่าง ได้แก่ ตัวอย่างที่ไม่ถูกกระตุ้นด้วยไลโปพอลิแซคคาไรด์ 3 ตัวอย่าง และตัวอย่างที่ถูกกระตุ้นด้วยไลโปพอลิแซคคาไรด์เป็นเวลา 8 ชั่วโมง 3 ตัวอย่าง แล้วทำการคัดเลือกโพรบที่มีการแสดงออกของยีนเฉลี่ยมากที่สุดเป็นตัวแทนของระดับการแสดงออกของยีน จะได้ยีนจำนวน 21,678 ยีน จากนั้นนำมาตรวจวัดระดับการแสดงออกของยีนโดยอาศัยวิธีการเปรียบเทียบด้วยค่าโพลีแซคคาไรด์ ซึ่งเป็นการตรวจวัดค่าจำนวนเท่าของการแสดงออกระหว่างสภาวะที่มีไลโปพอลิแซคคาไรด์ในแต่ละช่วงเวลากับสภาวะที่ไม่มี ก่อนนำค่าที่ได้มาปรับมาตรฐานด้วยการคำนวณค่าล็อกการิทึมฐานสอง เพื่อแสดงทิศทางของการแสดงออกของยีนในทางบวกและลบ (up- และ down-regulation) เพื่อนำไปใช้สำหรับคัดเลือกยีนที่สร้างทรานสคริปชันแฟคเตอร์และยีนเป้าหมายต่อไป

2.2 การคัดเลือกปฏิสัมพันธ์ระหว่างทรานสคริปชันแฟคเตอร์และยีนเป้าหมาย

จากข้อมูลไมโครอาร์เรย์ที่เตรียมไว้ข้างต้น และข้อมูลปฏิสัมพันธ์ของการควบคุมระหว่างทรานสคริปชันแฟคเตอร์กับยีนเป้าหมาย (TF-TG interaction) ซึ่งในที่นี้จะใช้ข้อมูลของหนูเมาส์ (Mouse; *Mus musculus*) (Ernst *et al.*, 2010) เนื่องจากเป็นสิ่งมีชีวิตต้นกำเนิดของเซลล์แมคโครฟาจ RAW 264.7 โดยข้อมูลนี้ประกอบด้วยทรานสคริปชันแฟคเตอร์จำนวน 335 ชนิด และยีนเป้าหมายจำนวน 17,090 ยีน และมีปฏิสัมพันธ์ของการควบคุมระหว่างทรานสคริปชันแฟคเตอร์กับยีนเป้าหมายจำนวน 468,319 ปฏิสัมพันธ์ ข้อมูลนี้จะถูกบรรจุอยู่ในฐานข้อมูลของโปรแกรม DREM (Schulz *et al.*, 2012) มาผ่านกระบวนการคัดเลือกทรานสคริปชันแฟคเตอร์และยีนเป้าหมาย โดย DREM จากนั้นทำการคัดกรองทรานสคริปชันแฟคเตอร์จากบทความวิจัยที่เกี่ยวข้องกับการศึกษาการตอบสนองการอักเสบ โดยพบทรานสคริปชันแฟคเตอร์จำนวน 41 ชนิด (Medzhitov and Horng, 2009., Nilsson *et al.*, 2006. และ Ravasi *et al.*, 2007) และ ทรานสคริปชันแฟคเตอร์ที่เป็นฟอสโฟโปรตีนจำนวน 10 ชนิด (Weintz *et al.*, 2010) เพื่อให้ง่ายต่อการวิเคราะห์และแปลความหมายทางชีวภาพมากขึ้น แล้วทำการเลือกยีนเป้าหมายที่มีค่าโพลีแซคคาไรด์มากกว่า 1.4 ก่อนนำมาแสดงผลร่วมกันระหว่างเครือข่ายของปฏิสัมพันธ์ของการควบคุมระหว่างทรานสคริปชันแฟคเตอร์กับยีนเป้าหมาย และค่าระดับการแสดงออกของยีนแต่ละชนิด

2.3 การสร้างเครือข่ายการควบคุมการแสดงออกของยีนและการแปลผลทางชีวภาพ

จากข้อมูลปฏิสัมพันธ์ระหว่างทรานสคริปชันแฟคเตอร์และยีนเป้าหมายที่เกี่ยวข้องกับการตอบสนองการอักเสบที่ผ่านการคัดเลือกแล้ว และข้อมูลซึ่งแสดงระดับการแสดงออกของยีนแบบโพลีแซคคาไรด์ นำมาสร้างเครือข่ายการควบคุมการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบ โดยใช้โปรแกรม Cytoscape (<http://www.cytoscape.org>; Cline *et al.*, 2007) โดยข้อมูลการแสดงออกของยีนที่แสดงผลแบบโพลีแซคคาไรด์หากมีการแสดงออกเพิ่มมากขึ้นในสภาวะที่มีไลโปพอลิแซคคาไรด์จะแสดงผลของยีนเป็นสีแดง (fold change > 1.4) และหากมีการแสดงออกที่ลดลงในสภาวะที่มีไลโปพอลิแซคคาไรด์จะแสดงผลของยีนเป็นสีเขียว (fold change < -1.4) และใช้การทดสอบทางสถิติ one-sample-t-test ในการคัดเลือกยีนที่มีระดับการแสดงออกที่เปลี่ยนแปลงตามช่วงเวลาอย่างมีนัยสำคัญ ($P < 0.05$, fold change > 1.4) หลังจากนั้นทำการวิเคราะห์ยีนเป้าหมายด้วยบิงโก (BINGO, Maere *et al.*, 2005) โดยเลือกการจัดกลุ่มยีนตามโมเลกุลาร์ฟังก์ชัน (Ashburner *et al.*, 2000) เพื่อตรวจสอบหน้าที่การทำงานเชิงชีวภาพของยีนเป้าหมาย และจะได้เครือข่ายการควบคุมการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบในเซลล์แมคโครฟาจ ซึ่งเครือข่ายที่ได้นี้จะแสดงให้เห็นถึงกิจกรรมของทรานสคริปชันแฟคเตอร์ที่มีต่อยีนเป้าหมาย และระดับการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบในแต่ละเวลาที่สนใจ

3. ผลและอภิปราย

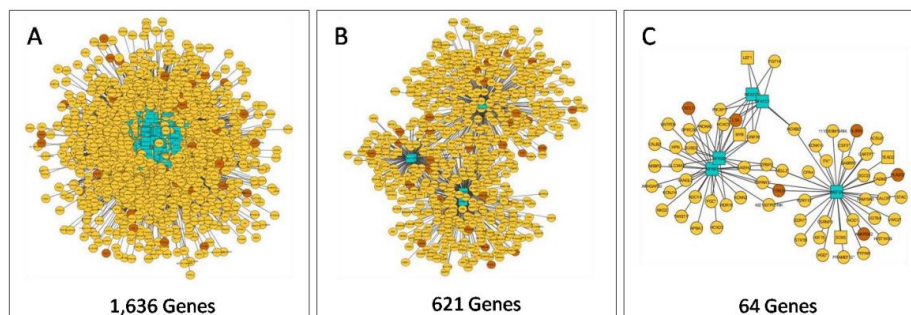
ในงานวิจัยนี้ได้เริ่มต้นโดยการใช้ข้อมูลปฏิสัมพันธ์ของทรานสคริปชันแฟคเตอร์กับยีนเป้าหมายซึ่งเป็นข้อมูลที่บรรจุอยู่ในโปรแกรม DREM (Schulz *et al.*, 2012) ของหนูเมาส์ (*Mus musculus*) (Ernst *et al.*, 2010) ซึ่งประกอบไปด้วยทรานสคริปชันแฟคเตอร์จำนวน 335 ชนิด ยีนเป้าหมายจำนวน 17,090 ยีน และปฏิสัมพันธ์ของการควบคุมระหว่างทรานสคริปชันแฟคเตอร์กับยีนเป้าหมายจำนวน 468,319 ปฏิสัมพันธ์ (ตารางที่ 1) เมื่อนำข้อมูลดังกล่าวมาตรวจสอบกับข้อมูลไมโครอาร์เรย์ที่เกี่ยวข้องกับการตอบสนองของเซลล์แมคโครฟาจที่มีต่อไลโปลิแซคคาไรด์ในช่วงเวลา 3, 6 และ 8 ชั่วโมง (Comer *et al.*, 2006 และ Hammer *et al.*, 2010) จะได้ทรานสคริปชันแฟคเตอร์จำนวน 335 ชนิด ยีนเป้าหมายจำนวน 1,636 ยีน และปฏิสัมพันธ์จำนวน 44,000 ปฏิสัมพันธ์ (ตารางที่ 1) จากผลการทดลองจะเห็นว่ากลไกที่เกี่ยวข้องกับการตอบสนองการอักเสบของกลุ่มยีนที่ถูกคัดเลือกโดย DREM นั้นจะถูกควบคุมโดยทรานสคริปชันแฟคเตอร์เป็นจำนวนมาก คือจำนวนทั้งหมดที่มีจากฐานข้อมูล แต่ยีนเป้าหมายที่เกี่ยวข้องกับการอักเสบจะมีจำนวนลดลงมาเหลือเพียงร้อยละ 15.27 ภายหลังการ คัดกรองทรานสคริปชันแฟคเตอร์จากงานวิจัยที่เกี่ยวข้อง (Medzhitov *et al.*, 2009., Nilsson *et al.*, 2006., Ravasi *et al.*, 2007 และ Weintz *et al.*, 2010) จะได้ทรานสคริปชันแฟคเตอร์จำนวน 5 ชนิด ยีนเป้าหมายจำนวน 621 ยีน และปฏิสัมพันธ์จำนวน 1,022 ปฏิสัมพันธ์ และเมื่อคัดเลือกยีนเป้าหมายที่มีค่าโพลด์เซนจ์มากกว่า 1.4 ทั้งในทิศทางบวกและลบ คือมีการแสดงออกเพิ่มมากขึ้นและลดลงตามลำดับ จะทำได้ ทรานสคริปชันแฟคเตอร์จำนวน 5 ชนิด ยีนเป้าหมายจำนวน 64 ยีน และปฏิสัมพันธ์จำนวน 100 ปฏิสัมพันธ์ (ตารางที่ 1) การเลือกใช้ข้อมูล ทรานสคริปชันแฟคเตอร์ที่ได้จากการตรวจวัดระดับฟอสโฟโปรตีน เพื่อเป็นการยืนยันว่าทรานสคริปชันแฟคเตอร์ชนิดนั้นมีการทำงานเกิดขึ้นจริง (Weintz *et al.*, 2010) ทำให้เราสามารถลดจำนวนของยีนเป้าหมายลงได้มากเพื่อให้ง่ายต่อการแปลผล และไม่ใช่ทุกยีนที่มีการเปลี่ยนแปลงระดับการแสดงออกหรือปริมาณ mRNA ดังนั้นเมื่อคัดกรองจากค่าโพลด์เซนจ์ของยีนจึงสามารถลดจำนวนยีนเป้าหมายลงให้เหลือเฉพาะยีนที่มีการเปลี่ยนแปลงชัดเจนอย่างมีนัยสำคัญ ($P < 0.05$, fold change > 1.4) และเมื่อพิจารณาเฉพาะยีนที่เคยรายงานมาก่อนว่าเกี่ยวข้องกับการอักเสบจะเหลือยีนเป้าหมายจำนวน 6 ยีน และปฏิสัมพันธ์จำนวน 10 ปฏิสัมพันธ์ ในขณะที่จำนวนทรานสคริปชันแฟคเตอร์ยังคงเดิม ก็จะทำให้การแปลผลที่เกี่ยวข้องกับการตอบสนองตามช่วงเวลาตรวจสอบได้ง่ายขึ้น

ตารางที่ 1 จำนวนทรานสคริปชันแฟคเตอร์ ยีนเป้าหมาย และปฏิสัมพันธ์ของการควบคุมระหว่างทรานสคริปชันแฟคเตอร์กับยีนเป้าหมาย ภายหลังจากวิเคราะห์ผลในแต่ละขั้นตอน

ผลการวิเคราะห์	ทรานสคริปชันแฟคเตอร์ (ชนิด)	ยีนเป้าหมาย (ยีน)	ปฏิสัมพันธ์ของการควบคุม (ปฏิสัมพันธ์)
ข้อมูลปฏิสัมพันธ์ของการควบคุมระหว่างทรานสคริปชันแฟคเตอร์กับยีนเป้าหมายหนูเมาส์	335	17,090	468,319
ผลจากการวิเคราะห์ข้อมูลการแสดงออกของยีนในชุดข้อมูลไมโครอาร์เรย์	335	1,636	44,000
ผลจากการคัดเลือกทรานสคริปชันแฟคเตอร์จากงานวิจัยที่เกี่ยวข้องกับการอักเสบและฟอสโฟโปรตีน	5	621	1,022
ผลจากการคัดเลือกยีนที่มีค่าโพลด์เซนจ์มากกว่า 1.4 ซึ่งถือเป็นยีนที่เกี่ยวข้องกับการอักเสบ	5	64	100
ผลจากการคัดเลือกยีนที่ตอบสนองการอักเสบจากการแบ่งกลุ่มฟังก์ชันของยีน	5	6	10

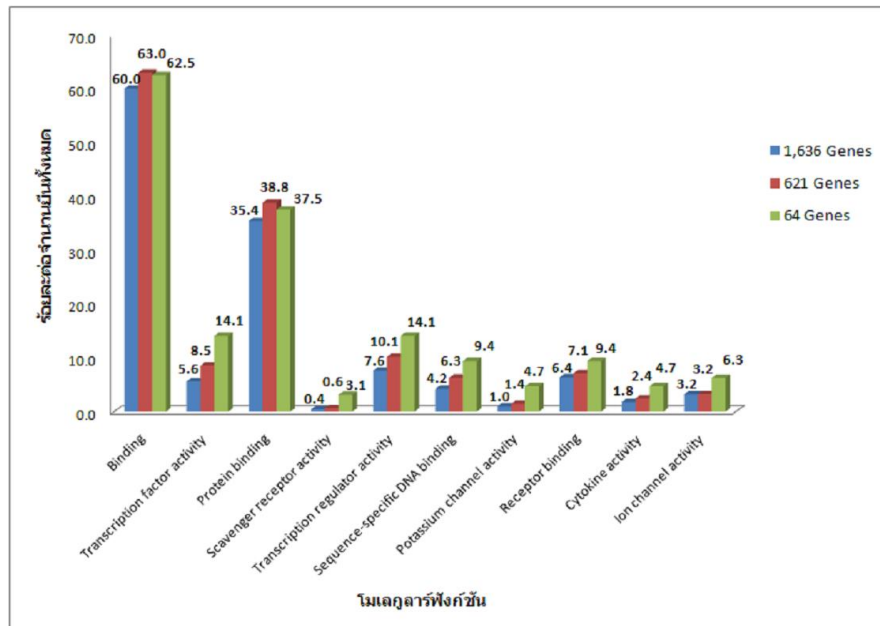
จากผลการคัดเลือกทรานสคริปชันแฟคเตอร์ ยีนเป้าหมาย และปฏิสัมพันธ์ของการควบคุมระหว่างทรานสคริปชันแฟคเตอร์กับยีนเป้าหมาย จะสามารถแสดงผลเครือข่ายได้ดังภาพที่ 2 ซึ่งภาพ 2A เป็นเครือข่ายที่มาจากผลการวิเคราะห์ข้อมูลการแสดงออกของยีน

ในชุดข้อมูลไมโครอาร์เรย์ ภาพ 2B เป็นเครือข่ายที่มาจากผลการคัดเลือกจากงานวิจัยที่เกี่ยวข้องกับการอักเสบ และภาพ 2C เป็นเครือข่ายที่มาจากผลการคัดเลือกยีนเป้าหมายที่มีค่าโพลีเดนจิมมากกว่า 1.4 ซึ่งถือว่าเป็นยีนที่เกี่ยวข้องกับการอักเสบ โดยสีเหลืองมีฟ้าแสดงทรานสคริปชันแฟคเตอร์ วงกลมสีน้ำตาลแสดงยีนเป้าหมายซึ่งเป็นที่ยอมรับของการอักเสบ (inflammatory response genes) จากการแบ่งกลุ่มฟังก์ชันของยีน (Ashburner *et al.*, 2000) สำหรับวงกลมสีส้มจะแสดงยีนเป้าหมายอื่น จากภาพที่ 2C แสดงให้เห็นถึงเครือข่ายที่มีการแสดงภาพกลุ่มยีนที่ตรวจสอบแล้วว่าเกี่ยวข้องกับการอักเสบในวงกลมสีน้ำตาลจำนวนทั้งสิ้น 6 ยีน ซึ่งประกอบไปด้วย *ANKRD42*, *CXCL11*, *IL5RA*, *LTA*, *PLA2G7* และ *TYRO3* (Ashburner *et al.*, 2000) ซึ่งจะถูกนำไปอธิบายลักษณะการแสดงออกในแต่ละช่วงเวลาอีกครั้งดังภาพที่ 4 จากภาพแสดงให้เห็นว่าการคัดกรองข้อมูลทั้งจากทรานสคริปชันแฟคเตอร์ที่เกี่ยวข้องและยีนเป้าหมายที่แสดงออกสอดคล้องกับการตอบสนองการอักเสบ ช่วยลดความซับซ้อนและปริมาณของข้อมูลเพื่อให้เหมาะสมสำหรับนำไปใช้ในการแปลผลทางชีวภาพที่เกี่ยวข้องต่อไปได้ ซึ่งยีนเป้าหมายที่เป็นที่ยอมรับของการอักเสบจะถูกลดจำนวนลงจากการคัดกรองข้อมูล และการแสดงผลภาพเครือข่ายจากข้อมูลปฏิสัมพันธ์โดยอาศัยโปรแกรม Cytoscape (Cline *et al.*, 2007) ก็ช่วยทำให้การตรวจสอบผลทำได้ง่ายขึ้น



ภาพที่ 2 เครือข่ายและจำนวนยีนที่ได้จากผลการวิเคราะห์ข้อมูลการแสดงออกของยีนในชุดข้อมูลไมโครอาร์เรย์จำนวน 1,636 ยีน (A) เครือข่ายที่มาจากผลการคัดเลือกจากงานวิจัยที่เกี่ยวข้องกับการอักเสบจำนวน 621 ยีน (B) และเครือข่ายที่มาจากผลการคัดเลือกยีนเป้าหมายที่มีค่าโพลีเดนจิมมากกว่า 1.4 จำนวน 64 ยีน (C)

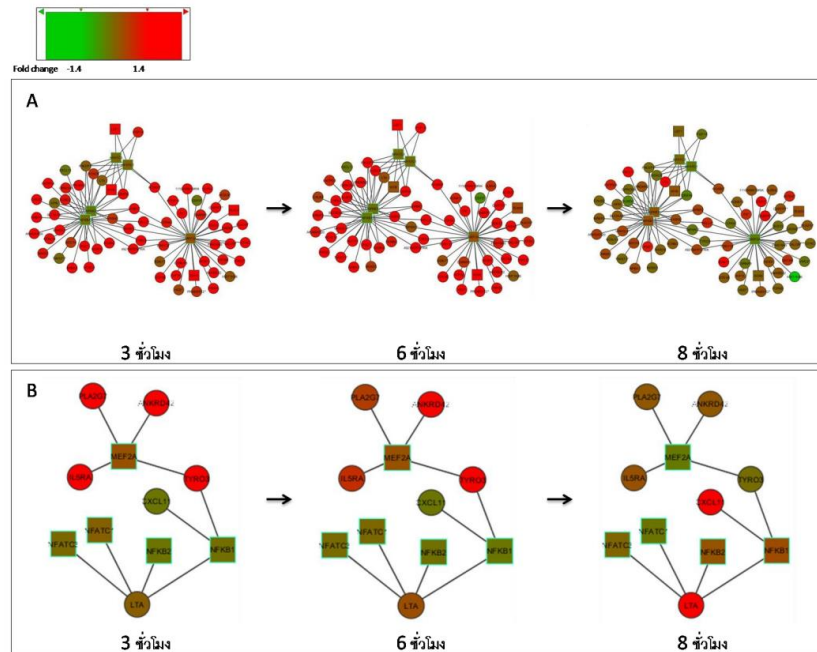
จากภาพที่ 2A จะเห็นว่าเครือข่ายมีขนาดใหญ่ซึ่งประกอบด้วยยีนจำนวนมากที่ถูกควบคุมโดยทรานสคริปชันแฟคเตอร์ โดยจะพบยีนเหล่านี้ในโมเลกุลาร์ฟังก์ชันต่างๆ ดังภาพที่ 3 ซึ่งแสดงถึงร้อยละของจำนวนยีนที่พบในโมเลกุลาร์ฟังก์ชันแต่ละชนิดเมื่อเทียบกับจำนวนยีนทั้งหมดของเครือข่าย 3 ชนิด ที่มีจำนวน 1,636 ยีน, 621 ยีน และ 64 ยีน ซึ่งโมเลกุลาร์ฟังก์ชันที่พบมีความเกี่ยวข้องกับกลไกการตอบสนองการอักเสบที่ประกอบไปด้วยกิจกรรมของ ทรานสคริปชันแฟคเตอร์ (Transcription factor activity, Transcription regulator activity และ Sequence-specific DNA binding) กิจกรรมของไซโตไคน์ (Cytokine activity) การจับกันของโปรตีน (Protein binding) โปรตีนตัวรับ (Receptor binding และ Scavenger receptor activity) โปรตีนขนส่งไอออนและเกลือแร่ (Potassium channel activity และ Ion channel activity) และเมื่อมีการลดขนาดของเครือข่ายที่วิเคราะห์ให้เล็กลงดังภาพที่ 2B และ 2C ก็ยังคงพบยีนสมาชิกในแต่ละกลุ่มข้างต้นที่ยังคงสัดส่วนของร้อยละของจำนวนยีนที่พบคงเดิม ก็แสดงว่ากลุ่มโมเลกุลาร์ฟังก์ชันเหล่านี้มีความสำคัญต่อกลไกการตอบสนองการอักเสบที่เกิดขึ้นของเซลล์แมคโครฟาจหนู RAW 264.7 ที่มีต่อการกระตุ้นโดย LPS เนื่องมาจากกระบวนการตอบสนองการอักเสบเกิดจากการทำงานร่วมกันจากการแสดงออกของยีน เพื่อควบคุมกระบวนการอักเสบและช่วยลดการอักเสบของเซลล์ได้ (Ashburner *et al.*, 2000) ซึ่งให้ผลในทิศทางเดียวกับที่มีการศึกษาการตอบสนองการอักเสบของมนุษย์ (Chen *et al.*, 2008)



ภาพที่ 3 ร้อยละของจำนวนยีนที่ปรากฏในแต่ละกลุ่มโมเลกุลาร์ฟังก์ชันที่ถูกคัดเลือกจากเครือข่ายเมื่อเทียบกับยีนทั้งหมดจำนวน 1,636 ยีน, 621 ยีน และ 64 ยีน ซึ่งแกน X แสดงถึงโมเลกุลาร์ฟังก์ชัน และแกน Y แสดงร้อยละต่อจำนวนยีนทั้งหมดที่ถูกพบในแต่ละโมเลกุลาร์ฟังก์ชัน

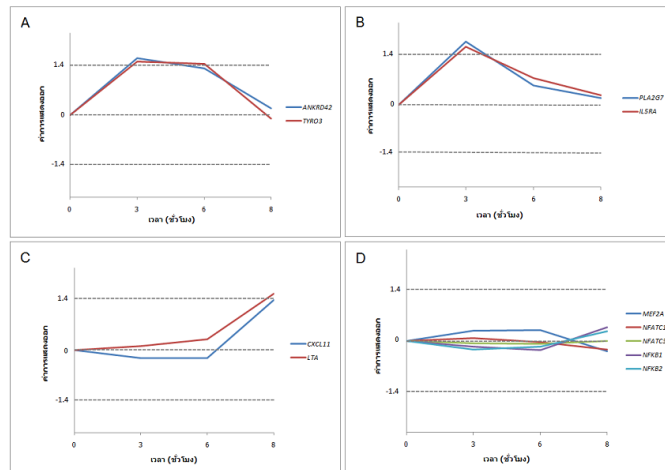
จากภาพที่ 4A เป็นเครือข่ายการควบคุมการแสดงออกของยีนในเวลา 3, 6 และ 8 ชั่วโมง ซึ่งมาจากผลการคัดเลือกยีนที่มีค่าโพลดีเซนจ์มากกว่า 1.4 จะเห็นว่ายีนมีระดับการแสดงออกในเวลา 3 และ 6 มากที่สุด และในเวลา 8 ชั่วโมง ยีนนั้นมีระดับการแสดงออกที่ลดลงที่มีการควบคุมโดยทรานสคริปชันแฟกเตอร์ในสภาวะที่ถูกกระตุ้นด้วยไลโปลิดแซคคาไรด์ จากเครือข่ายนี้จะมียีนที่เกี่ยวข้องกับการตอบสนองการอักเสบดังแสดงในภาพที่ 4B ซึ่งเป็นเครือข่ายการควบคุมการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบที่มีการแสดงออกเวลาที่ 3, 6 และ 8 ชั่วโมง ซึ่งเครือข่ายนี้ประกอบด้วยทรานสคริปชันแฟกเตอร์จำนวน 5 ชนิด ยีนเป้าหมายจำนวน 6 ยีน และปฏิสัมพันธ์ของการควบคุมระหว่างทรานสคริปชันแฟกเตอร์กับยีนเป้าหมายจำนวน 10 ปฏิสัมพันธ์ เมื่อพิจารณาระดับการแสดงออกที่เปลี่ยนแปลงตามช่วงเวลา 3, 6 และ 8 ชั่วโมง ของกลุ่มยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบ (ภาพที่ 4B) จะพบว่ายีนที่เกี่ยวข้องกับการตอบสนองการอักเสบจะถูกควบคุมการแสดงออกโดยทรานสคริปชันแฟกเตอร์ที่แตกต่างกัน โดยที่ยีน *LTA* จะถูกควบคุมการแสดงออกโดยทรานสคริปชันแฟกเตอร์ *NFAT* (*NFATC1* และ *NFATC3*) และ *NF-κB* (*NFKB1* และ *NFKB2*) จะมีการแสดงออกที่เวลา 8 ชั่วโมง ยีน *TYRO3* จะถูกควบคุมการแสดงออกโดย *MEF2* (*MEF2A*) และ *NF-κB* (*NFKB1*) จะมีการแสดงออกที่เวลา 3 และ 6 ชั่วโมง ยีน *ANKRD42*, *IL5RA* และ *PLA2G7* จะถูกควบคุมการแสดงออกโดย *MEF2* (*MEF2A*) โดยที่ยีน *ANKRD42* จะมีการแสดงออกที่เวลา 3 และ 6 ชั่วโมง ส่วนยีน *IL5RA* และ *PLA2G7* จะมีการแสดงออกที่เวลา 3 ชั่วโมงเท่านั้น และยีน *CXCL11* จะถูกควบคุมการแสดงออกโดย *NF-κB* (*NFKB1*) จะมีการแสดงออกที่เวลา 8 ชั่วโมง และหากพิจารณาที่ศูนย์กลางการควบคุมการแสดงออกของยีน จากภาพที่ 4B พบว่าทรานสคริปชันแฟกเตอร์ *MEF2* (*MEF2A*) เป็นศูนย์กลางการควบคุมการแสดงออกของยีนในช่วงเวลา 3 และ 6 ชั่วโมง โดยควบคุมให้เกิดการแสดงออกที่เพิ่มมากขึ้นของยีน *TYRO3*, *ANKRD42*, *IL5RA* และ *PLA2G7* ในเวลา 3 ชั่วโมง และจะเหลือการควบคุมเพียงยีน *TYRO3* และ *ANKRD42* ในช่วงเวลา 6 ชั่วโมง ส่วนในช่วงเวลา 8 ชั่วโมง *NF-κB* (*NFKB1* และ *NFKB2*) จะกลับมามีบทบาทในการควบคุม

การแสดงออกที่เพิ่มมากขึ้นของยีน *CXCL11* และ *LTA* ซึ่งยีน *LTA* ก็จะถูกควบคุมได้ด้วยทรานสคริปชันแฟคเตอร์ NFAT (*NFATC1* และ *NFATC3*) เช่นกัน ซึ่งแสดงให้เห็นถึงแบบแผนของการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบที่ถูกควบคุมด้วยกลุ่มของทรานสคริปชันแฟคเตอร์ที่แตกต่างกันไปตามแต่ละช่วงของเวลา



ภาพที่ 4 เครือข่ายการควบคุมการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบที่มาจากผลการคัดเลือกยีนเป้าหมายที่มีค่าโพลดีเซนจ์มากกว่า 1.4 (A) และเมื่อคัดเลือกเฉพาะกลุ่มยีนเป้าหมายที่เป็นยีนตอบสนองการอักเสบจำนวน 6 ยีน (B) ซึ่งจะแสดงระดับการแสดงออกของยีนตามค่าโพลดีเซนจ์จากการแสดงออกที่ลดลง (สีเขียว) ไปยังค่าการแสดงออกที่เพิ่มขึ้น (สีแดง) สีเหลี่ยมแสดงทรานสคริปชันแฟคเตอร์ วงกลมแสดงยีนเป้าหมาย

จากเครือข่ายการควบคุมยีนที่เป็นยีนตอบสนองการอักเสบจำนวน 6 ยีน (ภาพที่ 4B) จะพบยีนเป้าหมายที่มีการแสดงออกตามช่วงเวลาที่แตกต่างกันไป โดยที่ยีน *ANKRD42* และ *TYRO3* จะมีการแสดงออกไปในทิศทางเดียวกันและมีการแสดงออกมากในเวลา 3 และ 6 ชั่วโมง (ภาพที่ 5A) ยีน *PLA2G7* และ *IL5RA* จะมีการแสดงออกมากในเวลา 3 ชั่วโมง และมีการแสดงออกที่ลดลงในเวลา 6 และ 8 ชั่วโมง (ภาพที่ 5B) ยีน *CXCL11* และ *LTA* จะมีการแสดงออกมากในเวลา 8 ชั่วโมง (ภาพที่ 5C) สำหรับยีน *MEF2A*, *NFATC1*, *NFATC3*, *NFKB1* และ *NFKB2* ซึ่งเป็นยีนที่สร้างโปรตีนทรานสคริปชันแฟคเตอร์ MEF2, NFAT และ NF-κB ซึ่งควบคุมการแสดงออกของยีนเป้าหมาย (ภาพที่ 5D) แสดงให้เห็นว่ายีนที่เกี่ยวข้องกับการตอบสนองการอักเสบจะมีระดับการแสดงออกของยีนที่แตกต่างกันไปตามช่วงเวลาการศึกษา โดยระดับการแสดงออกของยีนที่สร้างทรานสคริปชันแฟคเตอร์จะมีการแสดงออกที่ต่ำ ภายใต้สภาวะที่ถูกกระตุ้นด้วยไลโปพอลิแซคคาไรด์ เนื่องจากการทำงานของทรานสคริปชันแฟคเตอร์จะเกิดขึ้นในระบับการเปลี่ยนแปลงของโปรตีนหรือการเกิดกระบวนการฟอสโฟรีเลชันที่ส่งผลต่อการเกิดฟอสโฟโปรตีน ที่ทำให้เกิดกระบวนการส่งผ่านข้อมูลหรือสัญญาณในระดับโมเลกุล (Weintz *et al.*, 2010)

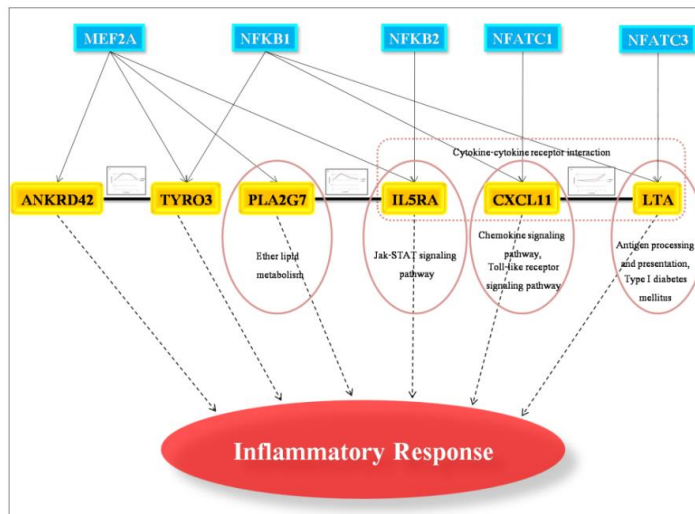


ภาพที่ 5 กราฟการแสดงออกของยีนเป้าหมายที่เป็นยีนตอบสนองการอักเสบจำนวน 6 ยีน (A, B และ C) และยีนที่สร้างโปรตีนทรานสคริปชันแฟกเตอร์ (D) ซึ่งแกน X แสดงถึงเวลาที่ทรานสคริปชันแฟกเตอร์และยีนเป้าหมายนั้นมีการแสดงออก แกน Y แสดงระดับการแสดงออกของยีนแบบค่าโฟลด์เช่นจากการแสดงออกที่ลดลง (fold change < -1.4) ไปยังค่าการแสดงออกที่เพิ่มขึ้น (fold change > 1.4)

เครือข่ายการควบคุมการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบ (ภาพที่ 4B) ซึ่ง *MEF2A* เป็นยีนที่สร้างโปรตีน *MEF2* มีความสัมพันธ์กับ *NF-κB* ที่พบในเครือข่ายของยีนที่เกี่ยวข้องกับไซโตไคน์ ที่กระตุ้นให้เกิดการอักเสบ (Chen *et al.*, 2008) ยีน *NFKB1* จะสร้างโปรตีน *NF-κB* ซึ่งมีความเกี่ยวข้องกับวิถี *MAPK* ที่เหนี่ยวนำให้เกิดการอักเสบทำให้ออกฤทธิ์ของ *NFKB1* ยังมีความสัมพันธ์กับ *NFKB2* ในการควบคุมการถอดรหัสที่เกี่ยวข้องกับการอักเสบและโรคเมะเร็ง (Nair *et al.*, 2008) ยีน *NFATC1* และ *NFATC3* จะสร้างโปรตีน *NFAT* ที่สามารถทำให้เกิดการอักเสบในกล้ามเนื้อของหนู ซึ่งการแสดงออกที่สูงของ *NFATC1* สามารถพบในเซลล์เม็ดเลือดขาวของผู้ป่วยโรคไขข้ออักเสบ (Yarilina *et al.*, 2011) และมีความสามารถที่จะกระตุ้นทริปซิโนเจนแล้วก่อให้เกิดการอักเสบในเนื้อเยื่อของตับอ่อนที่ระยะตัวอ่อนของหนู (Awla *et al.*, 2012) นอกจากนี้ยังมีเป้าหมายที่เกี่ยวข้องกับการตอบสนองการอักเสบ โดยยีน *ANKRD42* จะสร้างโปรตีน 8 ANK-containing protein ที่มีความคล้ายคลึงกับ *IκBβ* ซึ่งโปรตีนที่ยีน *ANKRD42* สร้างขึ้นมาจะอาศัยอยู่ในนิวเคลียสและจับกับ *p50* ของ *NF-κB* (Chiba *et al.*, 2013) สำหรับยีน *CXCL11* สามารถแสดงออกในเซลล์เม็ดเลือดขาวจากการถูกกระตุ้นด้วยแบคทีเรีย (Liu *et al.*, 2011) ส่วนยีน *IL5RA* มีบทบาทสำคัญในการผลิตอีโอซิโนฟิล (Eosinophil) และ *IgE* ต่อแอนติเจนที่นำไปสู่การอักเสบของผิวหนังและโรคหอบหืด (Cheong *et al.*, 2005) ยีน *LTA* มีผลต่อระดับของการอักเสบและทำให้เกิดโรคหัวใจขาดเลือด (Tanaka and Ozaki, 2006) เป็นส่วนหนึ่งที่ทำให้เกิดการอักเสบโดยตรง *PLA2G7* เป็นทั้งยีนที่ก่อให้เกิดและต้านการอักเสบ (Pro- and Anti-inflammatory) ที่เกี่ยวข้องกับกระบวนการอักเสบที่ก่อให้เกิดโรคหลอดเลือดและหัวใจ (Sutton *et al.*, 2009) และยีน *TYRO3* มีผลกระทบบต่อ *Gas6* ในกระบวนการฟาโกไซโทซิสและการตอบสนองการอักเสบที่เกี่ยวข้องกับการกระตุ้นการทำงานของโปรตีน *Mer* และ *Axl* (Zheng *et al.*, 2012)

การวิเคราะห์เป้าหมายที่เกี่ยวข้องกับการตอบสนองการอักเสบด้วย *DAVID* (Huang *et al.*, 2007) โดยอาศัยฐานข้อมูลของ *KEGG pathway* (Kanehisa *et al.*, 2010) ดังแสดงในภาพที่ 6 จะตรวจพบกลไกควบคุมการแสดงออกของยีน *ANKRD42*, *TYRO3*, *PLA2G7*, *IL5RA*, *CXCL11* และ *LTA* ซึ่งเป็นเป้าหมายที่เกี่ยวข้องกับการตอบสนองการอักเสบที่ถูกควบคุมการแสดงออกโดยยีน *MEF2A*, *NFKB1*, *NFKB2*, *NFATC1* และ *NFATC3* ซึ่งเป็นยีนที่สร้างโปรตีนทรานสคริปชันแฟกเตอร์ *MEF2*, *NF-κB* และ *NFAT* ที่ควบคุม

การแสดงออกของยีนเป้าหมาย โดยที่ยีน *ANKRD42* มีการแสดงออกที่สัมพันธ์กันหรือมีการแสดงออกไปในทิศทางเดียวกันกับยีน *TYRO3* สำหรับยีน *PLA2G7* จะมีการแสดงออกที่สัมพันธ์กับยีน *IL5RA* ซึ่งยีน *PLA2G7* เป็นยีนที่มีความเกี่ยวข้องกับกระบวนการสร้างและสลายลิพิด และยีน *IL5RA* มีความเกี่ยวข้องกับการส่งสัญญาณผ่านวิถี Jak-STAT ยีน *CXCL11* จะมีการแสดงออกที่สัมพันธ์กับยีน *LTA* ซึ่งยีน *CXCL11* มีความเกี่ยวข้องกับการส่งสัญญาณผ่านวิถีเคโมไคน์ (Chemokine) และการส่งสัญญาณผ่านวิถี Toll-like receptor และยีน *LTA* มีความเกี่ยวข้องกับการควบคุมการระบอบภูมิคุ้มกันในการผลิตและนำเสนอแอนติเจนให้กับ T-lymphocytes และเป็นยีนที่เกี่ยวข้องกับโรคเบาหวานชนิดที่ 1 นอกจากนี้ยีน *IL5RA*, *CXCL11* และ *LTA* ยังมีความเกี่ยวข้องกับปฏิสัมพันธ์การจับกันระหว่างไซโตไคน์ โดยที่ไซโตไคน์เป็นสารละลายโปรตีนที่อยู่ภายนอกเซลล์หรือเป็นไกลโคโปรตีน ซึ่งอยู่ระหว่างเซลล์มีหน้าที่ควบคุมการทำงานของเซลล์ในภูมิคุ้มกันแต่กำเนิด เช่น การป้องกันการอักเสบของเซลล์เจ้าบ้าน การเจริญของเซลล์ และการตายของเซลล์ เป็นต้น จากภาพแสดงให้เห็นว่ายีน *ANKRD42*, *TYRO3*, *PLA2G7*, *IL5RA*, *CXCL11* และ *LTA* เป็นยีนที่มีความเกี่ยวข้องกับการตอบสนองการอักเสบที่ถูกควบคุมการแสดงออก โดยทรานสคริปชันแฟกเตอร์ซึ่งยีนทั้ง 6 ยีน ไม่มีนัยสำคัญ ($P > 0.05$) ของการเปลี่ยนแปลงที่ระดับการแสดงออก ดังนั้นจึงจำเป็นที่จะต้องมีการยืนยันผลเพิ่มเติมโดยการเพิ่มจำนวนชุดข้อมูลที่ใช้ในการศึกษาและเพิ่มจำนวนข้อมูล TF-TG จากฐานข้อมูลเพื่อใช้ในการคัดเลือกปฏิสัมพันธ์และกลุ่มยีนที่เกี่ยวข้อง ก่อนที่จะทำการตรวจสอบในระดับห้องปฏิบัติการต่อไป



ภาพที่ 6 กลไกการควบคุมการแสดงออกของยีนเป้าหมายที่เกี่ยวข้องกับการตอบสนองการอักเสบที่ถูกควบคุมโดยทรานสคริปชันแฟกเตอร์ ซึ่งกล่องสี่เหลี่ยมสีฟ้าแทนทรานสคริปชันแฟกเตอร์ กล่องสี่เหลี่ยมสีเหลืองแทนยีนเป้าหมาย เส้นสีดำแทนยีนเป้าหมายที่มีรูปแบบการแสดงออกทิศทางเดียวกัน ลูกศรสีดำแทนปฏิสัมพันธ์ของการควบคุมระหว่างทรานสคริปชันแฟกเตอร์และยีนเป้าหมาย ลูกศรสีดำที่เป็นเส้นประแทนยีนเป้าหมายที่เกี่ยวข้องกับการตอบสนองการอักเสบ วงรีและกล่องสี่เหลี่ยมสีชมพูแทนวิถีที่พบในแต่ละยีนเป้าหมาย

4. บทสรุป

ผลจากการตรวจสอบเครือข่ายปฏิสัมพันธ์ของการควบคุมระหว่างทรานสคริปชันแฟกเตอร์กับยีนเป้าหมาย จะได้เครือข่ายที่มาจากผลการวิเคราะห์ข้อมูลการแสดงออกของยีนในชุดข้อมูลไมโครอาร์เรย์ เครือข่ายที่มาจากผลการคัดเลือกทรานสคริปชันแฟกเตอร์จากงานวิจัยที่เกี่ยวข้องกับการอักเสบ เครือข่ายที่มาจากผลการคัดเลือกยีนเป้าหมายที่มีค่าโฟลด์เชนจ์มากกว่า 1.4 และเครือข่ายที่มาจากผลการคัดเลือกยีนที่เกี่ยวข้องกับการอักเสบ ซึ่งเครือข่ายนี้ประกอบด้วยยีน *ANKRD42*, *CXCL11*, *IL5RA*, *LTA*, *PLA2G7* และ *TYRO3* ที่มีการควบคุมการ

แสดงออกโดยทรานสคริปชันแฟคเตอร์ MEF2 (*MEF2A*), NFAT (*NFATC1* และ *NFATC3*) และ NF- κ B (*NFKB1* และ *NFKB2*) จะเห็นว่าในเวลา 3, 6 และ 8 ชั่วโมง จะมีระดับการแสดงออกของยีนที่แตกต่างกันในสภาวะที่ถูกกระตุ้นด้วยไลโปพอลิแซคคาไรด์ ซึ่งแสดงให้เห็นถึงแบบแผนของการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบ ที่ถูกควบคุมด้วยกลุ่มของทรานสคริปชันแฟคเตอร์ที่แตกต่างกันไปตามแต่ละช่วงของเวลา การวิเคราะห์หน้าที่ของยีนเป้าหมายในโมเลกุลาร์ฟังก์ชัน พบว่ามีความเกี่ยวข้องกับการแสดงออกของยีนในกระบวนการตอบสนองการอักเสบ การสร้างเครือข่ายควบคุมการแสดงออกของยีนจากข้อมูลไมโครอาร์เรย์โดยอาศัยข้อมูลปฏิสัมพันธ์การควบคุมของทรานสคริปชันแฟคเตอร์ที่มีต่อยีนเป้าหมาย สามารถนำมาช่วยในการแปลผลเชิงชีวภาพของสภาวะที่สนใจได้ดีมากยิ่งขึ้น และการเพิ่มจำนวนชุดข้อมูลไมโครอาร์เรย์ที่ประกอบไปด้วยช่วงเวลาการศึกษาที่เพิ่มขึ้น การเพิ่มขนาดของฐานข้อมูล TF-TG และการปรับปรุงกระบวนการในขั้นตอนการเตรียมและการปรับมาตรฐานข้อมูล รวมถึงวิธีการคัดกรองกลุ่มยีนจากบทความวิจัย ก็อาจทำให้ได้เครือข่ายการควบคุมการแสดงออกที่มีความชัดเจนมากยิ่งขึ้น ที่พร้อมจะนำไปตรวจสอบในระดับห้องปฏิบัติการต่อไป ซึ่งจะช่วยให้ได้เครือข่ายควบคุมการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบของเซลล์แมคโครฟาจ RAW 264.7 ที่มีต่อการกระตุ้นของ LPS ที่สามารถนำไปพัฒนาเป็นเครือข่ายต้นแบบสำหรับใช้ในการศึกษากลไกการออกฤทธิ์ด้านทานการตอบสนองการอักเสบของสารสกัดชีวภาพหรือสารสังเคราะห์ต่อไปในอนาคต

5. กิตติกรรมประกาศ

ขอขอบคุณคณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา ที่ให้ทุนสนับสนุนสำหรับการนำเสนองานวิจัยในการประชุมวิชาการระดับชาติ "วิทยาศาสตร์วิจัย" ครั้งที่ 6 และงานวิจัยนี้ได้รับทุนสนับสนุนการวิจัยจากมหาวิทยาลัยบูรพา ประจำปีงบประมาณ 2557 สัญญาเลขที่ 87/2557

6. เอกสารอ้างอิง

- Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., *et al.* (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1), 25-29.
- Awla, D., Zetterqvist, A.V., Abdulla, A., Camello, C., Berglund, L.M., Spegel, P., *et al.* (2012). NFATc3 regulates trypsinogen activation, neutrophil recruitment and tissue damage in acute pancreatitis in mice. *Gastroenterology*, 143(5), 1352-1360.
- Chen, B.-S., Yang, S.-K., Lan, C.-Y. and Chuang, Y.-J. (2008). A systems biology approach to construct the gene regulatory network of systemic inflammation via microarray and databases mining. *BMC Medical Genomics*, 1, 46.
- Cheong, H.S., Kim, L. H., Park, B.L., Choi, Y.H., Park, H.-S., Hong, S.-J., *et al.* (2005). Association analysis of interleukin 5 receptor alpha subunit (IL5RA) polymorphisms and asthma. *Journal of Human Genetics*, 50(12), 628-634.
- Chiba, T., Inoko, H., Kimura, M. and Sato, T. (2013). Role of nuclear I κ Bs in inflammation regulation. *Biomolecular Concepts*, 4(2), 187-196.
- Cline, M.S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., *et al.* (2007). Integration of biological networks and gene expression data using Cytoscape. *Nature Protocols*, 2(10), 2366-2382.
- Comer, J.E., Galindo, C.L., Zhang, F., Wenglikowski, A.M., Bush, K.L., Garner, H.R., *et al.* (2006). Murine macrophage transcriptional and functional responses to *Bacillus anthracis* edema toxin. *Microbial Pathogenesis*, 41(2-3), 96-110.
- Edgar, R., Domrachev, M. and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, 30(1), 207-210.
- Ernst, J., Plasterer, H.L., Simon, I. and Bar-Joseph, Z. (2010). Integrating multiple evidence sources to predict transcription factor binding in the human genome. *Genome Research*, 20, 526-536.
- Hammer, K.D.P., Yum, M.-Y., Dixon, P.M. and Birt, D.F. (2010). Identification of JAK-STAT pathways as important for the anti-inflammatory activity of a *Hypericum perforatum* fraction and bioactive constituents in RAW 264.7 mouse macrophages. *Phytochemistry*, 71(7), 716-725.
- Hoesele, B. and Schmid, J.A. (2013). The complexity of NF- κ B signaling in inflammation and cancer. *Molecular Cancer*, 12, 86.

- Huang, D.W., Sherman, B.T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M.W., Lane, H.C. and Lempicki, R.A. (2007). DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Research*, 35: W169–W175.
- Kanamori, M., Konno, H., Osato, N., Kawai, J., Hayashizaki, Y. and Suzuki, H. (2004). A genome-wide and nonredundant mouse transcription factor database. *Biochemical and Biophysical Research Communications*, 322(3), 787–793.
- Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M. and Hirakawa, M. (2010). KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Research*, 38: D355–D360.
- Liu, z., Chen, X., Wang, X., Chen, X., Song, C.-H., Du, Y., et al. (2011). Chemokine CXCL11 links microbial stimuli to intestinal inflammation. *Clinical & Experimental Immunology*, 164(3), 396–406.
- Maere, S., Heymans, K. and Kuiper M. (2005). BINGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. *Bioinformatics*, 21(16), 3448–3449.
- Medzhitov R. and Hornig T. (2009). Transcriptional control of the inflammatory response. *Nature Reviews Immunology*, 9(10), 692–703.
- Nair, S., Doh, S.T., Chan, J.Y., Kong, A-N. and Cai, L. (2008). Regulatory potential for concerted modulation of Nrf2-and Nfka1-mediated gene expression in inflammation and carcinogenesis. *British Journal of Cancer*, 99, 2070–2082.
- Nilsson, R., Bajic, V.B., Suzuki, H., Bernardo, D.D., Björkegren, J., Katayama, S., et al. (2006). Transcriptional network dynamics in macrophage activation. *Genomics*, 88(2), 133–142.
- Ravasi, T., Wells, C.A. and Hume D.A. (2007). Systems biology of transcription control in macrophages. *Bioessays*, 29(12), 1215–1226.
- Schulz, M.H., Devanny, W.E., Gitter, A., Zhong, S., Ernst, J. and Bar-Joseph, Z. (2012). DREM 2.0: Improved reconstruction of dynamic regulatory networks from time-series expression data. *BMC Systems Biology*, 6, 104.
- Sutton, B.S., Crosslin, D.R., Shah, S.H., Nelson, S.C., Bassil, A., Hale A.B., et al. (2009). Comprehensive genetic analysis of the platelet activating factor acetylhydrolase (PLA2G7) gene and cardiovascular disease in case-control and family datasets. *Human Molecular Genetics*, 17(9), 1318–1328.
- Tanaka, T. and Ozaki, K. (2006). Inflammation as a risk factor for myocardial infarction. *Journal of Human Genetics*, 51, 595–604.
- Weintz, G., Olsen, J.V., Fruhauf, K., Niedzielska, M., Amit, I., Jantsch, J., et al. (2010). The phosphoproteome of toll-like receptor-activated macrophage. *Molecular System Biology*, 6, 371.
- Yarilina, A., Xu, K., Chen, J. and Ivashikiv, L.B. (2011). TNF activates calcium-nuclear factor of activated T cells (NFAT)c1 signaling pathways in human macrophages. *Proceedings of the National Academy of Sciences*, 108(4), 1573–1578.
- Zheng, Y., Wang, Q., Xiao, B., Lu, Q., Wang, Y. and Wang, X. (2012). Involvement of receptor tyrosine kinase Tyro3 in amyloidogenic APP processing and β -amyloid deposition in Alzheimer's disease models. *PLoS ONE*, 7(6), e39035.

ภาคผนวก ข

นำเสนอรายงานสืบเนื่องจากงานประชุมทางวิชาการระดับชาติ

การประชุมวิชาการระดับชาติ “วิทยาศาสตร์วิจัย” ครั้งที่ 6

วันที่ 20 – 21 มีนาคม 2557 มหาวิทยาลัยบูรพา จังหวัดชลบุรี ประเทศไทย

เหมือนฝัน โวหารกล้า และพิทักษ์ สุตรอนันต์. (2557). การวิเคราะห์การแสดงออกของยีนในข้อมูลไมโครอาร์เรย์ของเซลล์แมคโครฟาจที่ถูกกระตุ้นด้วยไลโปโพลีแซคคาไรด์. รายงานสืบเนื่องจากการประชุมวิชาการระดับชาติ “วิทยาศาสตร์วิจัย” ครั้งที่ 6, วันที่ 20 – 21 มีนาคม 2557, มหาวิทยาลัยบูรพา จังหวัดชลบุรี, 195–200.

การวิเคราะห์การแสดงออกของยีนในข้อมูลไมโครอาร์เรย์ของเซลล์แมคโครฟาจที่ถูกกระตุ้นด้วยไลโปโพลีแซคคาไรด์

Gene expression analysis in microarray data of LPS-stimulated macrophage

เหมื่อนฝั้น โหวฮากลา¹และ พิทักษ์ สูตรอนันต์^{2*}

Muanfan Wohankla¹ and Pitak Sootanan^{2*}

¹ ปรแกรมวิทยาศาสตรชีวภาพ โครงการบัณฑิตศึกษา คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

² ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

บทคัดย่อ

เมื่อเซลล์แมคโครฟาจถูกกระตุ้นด้วย LPS จะมีการตอบสนองผ่านทางทอลล์ไลค์รีเซพเตอร์ 4 โดยจะส่งผลให้มีการหลั่งของสารกลุ่มไซโตไคน์ (cytokine) เคโมไคน์ (chemokine) รวมถึงสารตัวกลางอื่นที่เกี่ยวข้องกับกระบวนการอักเสบมากมาย ในงานวิจัยนี้จึงได้นำข้อมูลการแสดงออกของยีนจากไมโครอาร์เรย์ที่เกี่ยวข้องกับการทดสอบการตอบสนองการอักเสบของเซลล์แมคโครฟาจ RAW264.7 เมื่อถูกกระตุ้นด้วย LPS จากฐานข้อมูลสาธารณะ มาวิเคราะห์เพื่อคัดเลือกยีนที่มีความสำคัญในการตอบสนองการอักเสบ โดยใช้วิธีการคัดเลือกที่เกี่ยวข้องกับการจำแนกประเภท คือ SVM (support vector machine) การเลือกข้อมูลโดยอาศัยการเปรียบเทียบโฟลด์เชนจ์ (fold change) ที่ระดับ 1.4 และการจัดกลุ่มข้อมูลของยีนตามแบบแผนการแสดงออกโดยใช้ SOM (self organizing map) มาใช้ร่วมกันในการคัดเลือกยีนที่เกี่ยวข้องกับการตอบสนองต่อการอักเสบ พบว่ามียีน 4 ยีน คือ *Cr2*, *H2-D1*, *Trpv1* และ *Gap43* ซึ่งเกี่ยวข้องกับกลไกการป้องกันการตอบสนองต่อสิ่งแปลกปลอมของร่างกายในระบบภูมิคุ้มกัน การนำข้อมูลไมโครอาร์เรย์จากฐานข้อมูลสาธารณะมาวิเคราะห์จะช่วยลดค่าใช้จ่ายในการศึกษา และหากมีการใช้วิธีการที่หลากหลายมาคัดเลือกอาจทำให้สามารถค้นพบกลุ่มยีนใหม่ที่น่าสนใจได้

คำสำคัญ : แมคโครฟาจ / LPS / ข้อมูลไมโครอาร์เรย์ / การคัดเลือกยีน / โฟลด์เชนจ์ / SVM / SOM

Abstract

Macrophage cells are stimulated by LPS response through toll-like receptor 4. Cytokine, chemokine and inflammatory mediators are produced to begin inflammation. In this study, gene expression microarray datasets involving inflammatory response by RAW 264.7 macrophages were retrieved from an open-access database. Classification with SVM (support vector machine), fold-change approach and clustering with SOM (self organizing map) were used to find responsive genes of inflammation. We found 4 genes such as *Cr2*, *H2-D1*, *Trpv1* and *Gap43* are associated non-specific immune response. Using of public microarray datasets lead to pay less cost and the use of combination of difference gene selection approaches can be discovered a novel set of responsive genes.

Keywords : macrophage / LPS / microarray data / gene selection / fold change / SVM / SOM

*Corresponding author. E-mail : Pitak@buu.ac.th

1. บทนำ

การอักเสบเป็นปฏิกิริยาการตอบสนองที่ซับซ้อนของเนื้อเยื่อต่ออันตรายหรือเนื้อเยื่อที่เสียหายหรือตายลง ซึ่งเป็นความพยายามของร่างกายที่จะนำสิ่งแปลกปลอมนั้นออกไปและซ่อมแซมเนื้อเยื่อที่ถูกทำลาย ถ้าหากไม่มีการอักเสบเกิดขึ้นเชื้อโรคจะไม่ถูกกำจัดออกไปจะทำให้เกิดความเสียหายของเนื้อเยื่อมากขึ้นจนอันตรายถึงชีวิตได้ แต่ทั้งนี้การอักเสบที่มีมากเกินไปก็สามารถทำให้เกิดโรคต่างๆ ได้ เช่น ใช้หลอดฟาง โรคท่อเลือดแดง และหลอดเลือดแดงแข็ง และโรคข้ออักเสบรูมาตอยด์ (Cousens *et al.*, 2002) เซลล์แมคโครฟาจเป็นเซลล์เม็ดเลือดขาวที่สำคัญมีหน้าที่เป็นเซลล์หน้าด่านในระบบภูมิคุ้มกันที่ติดตัวมาแต่กำเนิด (innate immune response) ในการกำจัดสิ่งแปลกปลอมที่เข้ามาในร่างกาย เมื่อเซลล์แมคโครฟาจถูกกระตุ้นด้วย LPS (lipopolysaccharide) ซึ่งเป็นส่วนประกอบของผนังเซลล์แบคทีเรียแกรมลบ จะทำให้เซลล์แมคโครฟาจมีการตอบสนองผ่านทางทอลลิลีเคอร์ 4 (toll-like receptor 4, TLR4) ที่อยู่บนเซลล์เมมเบรนซึ่งจะจับกับโมเลกุลโครงสร้างบนผิวสิ่งแปลกปลอมอย่างจำเพาะ แล้วส่งผลให้มีการหลั่งของสารกลุ่มไซโตไคน์ (cytokine) เคโมไคน์ (chemokine) รวมถึงสารตัวกลางอื่นๆ ที่เกี่ยวข้องกับกระบวนการอักเสบมากมาย (Lee *et al.*, 2012) การเรียนรู้เพื่อค้นหากลุ่มยีนที่เกี่ยวข้องกับการตอบสนองการอักเสบ จึงมีส่วนสำคัญที่ทำให้สามารถทำความเข้าใจกลไกการตอบสนองการอักเสบที่เกิดขึ้นได้

ในปัจจุบันได้มีการศึกษาการแสดงออกของยีนโดยใช้เทคโนโลยีไมโครอาร์เรย์ ซึ่งเป็นเทคนิคที่นิยมใช้ในการตรวจสอบการแสดงออกของยีนจำนวนมากในเวลาเดียวกัน โดยจะตรวจวัดปริมาณระดับการแสดงออกของยีนในระดับ เมสเซนเจอร์อาร์เอ็นเอ (mRNA) ภายใต้สภาวะที่ศึกษา (Watson *et al.*, 1998) และได้มีการรวบรวมข้อมูลไมโครอาร์เรย์จากหลากหลายกลุ่มงานวิจัยไว้ในฐานข้อมูลสาธารณะ เช่น Gene Expression Omnibus (GEO) ใน NCBI (Edgar *et al.*, 2002; Gene Expression Omnibus; <http://www.ncbi.nlm.nih.gov/geo/>) และฐานข้อมูล Stanford Microarray (Golub *et al.*, 2003; SMD; <http://genome-www.stanford.edu/microarray/>) เป็นต้น โดยเฉพาะข้อมูลไมโครอาร์เรย์ที่เกี่ยวข้องกับการตอบสนองต่อเซลล์แมคโครฟาจเซลล์ไลน์สายพันธุ์ RAW264.7 ของหนู ที่ถูกกระตุ้นด้วย LPS ซึ่งเป็นเซลล์มาตรฐานที่นิยมใช้เป็นโมเดลในการศึกษาการอักเสบ ในการวิเคราะห์ข้อมูลไมโครอาร์เรย์ขั้นตอนการคัดเลือกยีนเป็นขั้นตอนที่มีความสำคัญมาก เนื่องจากข้อมูลไมโครอาร์เรย์นั้นประกอบด้วยยีนจำนวนมาก ทำให้การวิเคราะห์มีความซับซ้อนและอาจส่งผลให้เกิดความผิดพลาดได้สูง จึงต้องมีการคัดเลือกเฉพาะกลุ่มยีนที่มีความสำคัญต่อสภาวะที่สนใจศึกษาอย่างแท้จริง โดยวิธีการที่นิยมใช้ในการคัดเลือกยีน ได้แก่ วิธีการทางสถิติ SAM (Selvaraj *et al.*, 2011) ในการคัดเลือกยีนโดยใช้วิธีการทั่วไปนั้น อาจไม่สามารถคัดเลือกยีนที่มีการแสดงออกที่ชัดเจนได้ครอบคลุม จึงได้มีการนำเทคนิคการทำเหมืองข้อมูล (data mining) ได้แก่ การจำแนกประเภท (classification) การจัดกลุ่มข้อมูล (clustering) และการเลือกข้อมูล (feature selection) มาใช้ในการคัดเลือกยีน และหากมีการใช้วิธีการเปรียบเทียบอัตราส่วนค่าระดับการแสดงออกของยีนที่แสดงในรูปของจำนวนเท่าหรือโฟลด์เชนจ์ (fold change) (Tarca *et al.*, 2006) มาช่วยในการคัดเลือกยีนที่มีความสำคัญที่เกี่ยวข้องกับการตอบสนองต่อ LPS ได้ครอบคลุมมากยิ่งขึ้น

ในงานวิจัยนี้จึงได้นำข้อมูลการแสดงออกของยีนจากไมโครอาร์เรย์ที่เกี่ยวข้องกับการทดสอบการตอบสนองการอักเสบของเซลล์แมคโครฟาจ RAW264.7 ที่มีต่อ LPS จากฐานข้อมูลสาธารณะ GEO มาวิเคราะห์เพื่อคัดเลือกยีนที่มีความสำคัญในการตอบสนองการอักเสบ โดยใช้วิธีการคัดเลือกที่เกี่ยวข้องกับการจำแนกประเภท คือ SVM (support vector machine; Guyon *et al.*, 2002) การจัดกลุ่มข้อมูล คือ SOM (self-organizing map; Nikkilä, 2002) และการเลือกข้อมูลโดยอาศัยการเปรียบเทียบโฟลด์เชนจ์ มาคัดเลือกยีนที่เกี่ยวข้องกับการอักเสบหรือตอบสนองต่อ LPS และนำยีนที่ถูกคัดเลือกไปค้นหาหน้าที่ความสำคัญในกระบวนการทางชีวภาพ ซึ่งจะช่วยให้สามารถเข้าใจกลไกการตอบสนองต่อ LPS ของเซลล์แมคโครฟาจได้มากยิ่งขึ้น ซึ่งอาจจะเป็นการค้นพบยีนใหม่ที่มีความสำคัญที่เกี่ยวข้องกับการอักเสบหรือมีการตอบสนองต่อ LPS ที่ยังไม่ได้มีการศึกษา เพื่อเป็นการช่วยลดเวลาและลดต้นทุนในการศึกษาก่อนที่จะนำไปตรวจสอบในห้องปฏิบัติการต่อไป

2. วิธีการ

ภาพรวมของงานวิจัยนี้ประกอบด้วยวิธีการคือ การรวบรวมข้อมูลไมโครอาร์เรย์ การเตรียมข้อมูลไมโครอาร์เรย์ การคัดเลือกยีนที่มีการตอบสนองต่อการอักเสบ การวิเคราะห์กลุ่มยีนที่มีการแสดงออกร่วมกัน การวิเคราะห์และแปลผลของยีนที่ถูกคัดเลือก

2.1 การรวบรวมข้อมูลไมโครอาร์เรย์

ดาวน์โหลดชุดข้อมูลไมโครอาร์เรย์จากฐานข้อมูลสาธารณะ GEO (Edgar *et al.*, 2002, Gene Expression Omnibus, <http://www.ncbi.nlm.nih.gov/geo/>) ซึ่งเป็นค่าการแสดงออกของยีนที่อยู่ในรูปแบบไฟล์ series expression matrix ทั้งหมด 2 ชุดข้อมูล ได้แก่ ชุดข้อมูลไมโครอาร์เรย์ GSE4712 และ GSE21841 ที่อยู่ภายใต้แพลตฟอร์มเดียวกันคือ Affymetrix Mouse Genome 430 2.0 โดยชุดข้อมูลไมโครอาร์เรย์ที่ได้นำมาวิเคราะห์นั้น จะคัดเลือกมา 2 กลุ่มการทดลองคือ กลุ่มที่ถูกกระตุ้นด้วย LPS และกลุ่มที่ไม่ถูกกระตุ้นด้วย LPS ตามช่วงเวลา ในเซลล์แมคโครฟาจเซลล์ไลน์สายพันธุ์ RAW264.7 ของหนู ดังที่ได้แสดงรายละเอียดของชุดข้อมูลในตารางที่ 1

ตารางที่ 1 แสดงรายละเอียดชุดข้อมูลไมโครอาร์เรย์

GSE ID	Title	References	Samples	Sources of LPS
GSE4712	Macrophage derived cell lines (RAW264.7)	Comer <i>et al.</i> , 2006	9 samples EdTx _LPS treated_3h=3 EdTx _LPS-treated_6h=3 EdTx _untreated=3	LPS (0.01ng/ml; Sigma Aldrich)
GSE21841	Macrophage derived cell lines (RAW264.7)	Hammer <i>et al.</i> , 2010	6 samples Media+DMSO_untreated=3 Media+DMSO+LPS_8h=3	LPS (1ug/ml; Sigma Aldrich)

2.2 การเตรียมข้อมูลไมโครอาร์เรย์

เนื่องจากชุดข้อมูลไมโครอาร์เรย์ทั้งสองชุดข้อมูลที่ใช้ในงานวิจัยนี้มาจากแหล่งที่มาแตกต่างกัน จึงต้องมีการนำข้อมูลไมโครอาร์เรย์ไปผ่านกระบวนการจัดเตรียมข้อมูล (data preprocessing) และผ่านกระบวนการนอร์มอลไลเซชัน (normalization) เพื่อปรับมาตรฐานค่าของตัวเลขให้อยู่ในมาตรฐานเดียวกัน ซึ่งเป็นการช่วยลดผลกระทบและความผิดพลาดต่อการวิเคราะห์เนื่องมาจากหน่วยวัดในการทดลองไมโครอาร์เรย์นั้นอาจมีการทำการทดลองซ้ำ ซึ่งจะทำให้ 1 ยีนนั้นมีหลายโพรบ จึงต้องทำการปรับข้อมูลไมโครอาร์เรย์ให้อยู่ในรูปแบบ 1 ยีน 1 โพรบ โดยการเลือกโพรบที่มีค่าเฉลี่ยของค่าการแสดงออกสูงที่สุด ซึ่งเป็นการช่วยลดขนาดของข้อมูลเพื่อให้ง่ายต่อการวิเคราะห์และเป็นวิธีที่ทำให้การสูญเสียข้อมูลไปน้อยที่สุด จากในชุดข้อมูลไมโครอาร์เรย์มีทั้งหมด 45,101 โพรบ มีจำนวนยีน 21,678 ยีน จากนั้นทำการปรับมาตรฐานค่าการแสดงออกของยีนในแต่ละชุดข้อมูลให้อยู่ในรูปแบบเดียวกัน โดยใช้การคำนวณค่าลอการิทึมฐานสิบ

2.3 การคัดเลือกยีนที่มีการตอบสนองต่อการอักเสบ

ในการคัดเลือกยีนที่มีการตอบสนองต่อการอักเสบในเซลล์แมคโครฟาจจะใช้วิธีการคัดเลือกทั้งหมด 2 วิธี คือ

2.3.1 วิธี SVM (Support Vector Machines) เป็นเทคนิคที่นำมาใช้กันอย่างกว้างขวางในการจำแนกรูปแบบ (classification) ซึ่งสามารถนำมาใช้ในการคัดเลือกกลุ่มยีน (gene selection) โดยจะเรียงลำดับความสำคัญในการจัดจำแนกกลุ่มข้อมูลจากมากไปหาน้อยโดยอาศัยหลักการการจำแนกข้อมูลออกเป็นสองกลุ่มออกจากกันโดยใช้การแบ่งแบบเชิงเส้น ในงานวิจัยนี้จะทำการคัดเลือกยีนด้วยวิธี SVM ซึ่งเป็นโปรแกรมย่อยภายในโปรแกรม WEKA เวอร์ชัน 3.6.2 (Hall *et al.*, 2009) โดยคัดเลือกยีนที่มีความสำคัญ 100 ลำดับแรกจากทั้งหมด 21,678 ยีน

2.3.2 วิธีการเปรียบเทียบโฟลด์เชนจ์ (fold change) โดยการนำค่าการแสดงออกของยีนในชุดข้อมูลไมโครอาร์เรย์มาเปรียบเทียบอัตราส่วนระดับการแสดงออกของยีนที่ถูกกระตุ้นด้วย LPS ตามช่วงเวลา เปรียบเทียบกับระดับการแสดงออกของยีนเดียวกันเมื่อไม่มีการกระตุ้นด้วย LPS ค่าที่ได้จะแสดงในรูปของจำนวนเท่าของโฟลด์เชนจ์ ซึ่งจะแสดงถึงการเปลี่ยนแปลงระดับการแสดงออกของยีนที่เพิ่มขึ้นหรือลดลงในแต่ละยีนภายใต้สภาวะและช่วงเวลาการศึกษา โดยกำหนดค่าโฟลด์เชนจ์ที่มากกว่า 1.4 และใช้การทดสอบทางสถิติ one-sample *t*-test ในการคัดเลือกยีนที่มีระดับการแสดงออกของยีนที่เปลี่ยนแปลงตามช่วงเวลาอย่างมีนัยสำคัญ ($P < 0.05$, > 1.4 fold change)

2.4 การวิเคราะห์กลุ่มยีนที่มีแนวโน้มการแสดงออกพร้อมกัน

นำยีนที่คัดเลือกได้โดยวิธี SVM และการเปรียบเทียบโพลด์เซนจ์ มาทำการจัดกลุ่มรูปแบบการแสดงออกของยีนโดยใช้ SOM (self-organizing map; Nikkilä, 2002) ที่บรรจุอยู่ในของการจัดกลุ่มข้อมูลของโปรแกรม MEV (Multi Experiment Viewer, Eleanor *et al.*, 2011) ซึ่งเป็นเทคนิคที่ใช้ในการแบ่งกลุ่มข้อมูลที่มีรูปแบบข้อมูลที่มีลักษณะเดียวกันไว้ในกลุ่มเดียวกัน แล้ววิเคราะห์รูปแบบการแสดงออกของยีนในแต่ละกลุ่ม แล้วคัดเลือกกลุ่มยีนที่มีความน่าสนใจและต้องการศึกษา

2.5 การวิเคราะห์และแปลผลทางชีวภาพของยีนที่ถูกคัดเลือก นำยีนที่ได้จากการคัดเลือกที่มีรูปแบบการแสดงออกของยีนในรูปแบบเดียวกันและมีการตอบสนองต่อการอักเสบ อย่างมีนัยสำคัญไปวิเคราะห์หาหน้าที่และความสำคัญในกระบวนการทางชีวภาพโดยใช้ฐานข้อมูลสาธารณะ DAVID (The Database for Annotation, Visualization and Integrated Discovery; Huang *et al.*, 2009)

3. ผลและอภิปราย

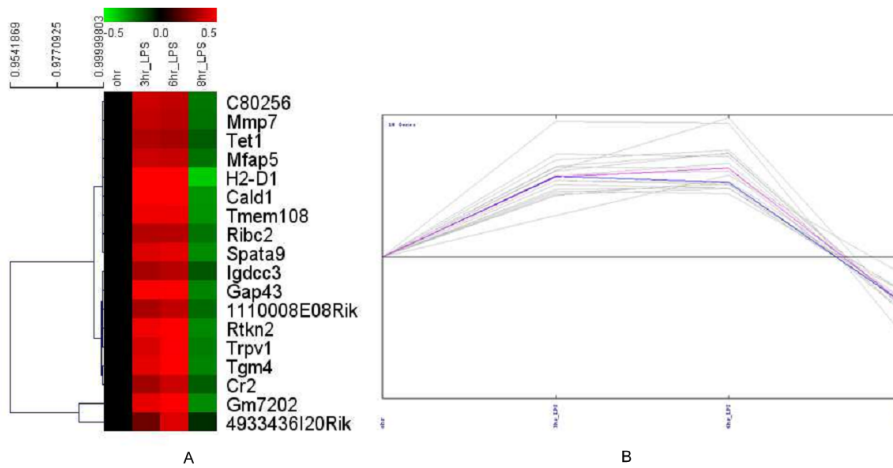
ในงานวิจัยนี้มีวัตถุประสงค์เพื่อคัดเลือกยีนที่เกี่ยวข้องกับการอักเสบหรือยีนที่มีการตอบสนองต่อ LPS ของชุดข้อมูลไมโครอาร์เรย์จากฐานข้อมูลสาธารณะ ในการคัดเลือกยีนที่มีการตอบสนองต่อ LPS โดยใช้วิธีการคัดเลือกทั้งหมด 3 วิธี คือ SVM การเปรียบเทียบโพลด์เซนจ์ และการจัดกลุ่มข้อมูลโดย SOM โดยอาศัยการคัดเลือกด้วย SVM และโพลด์เซนจ์ ในการคัดเลือกช่วงแรก เนื่องจากการคัดเลือกโดยวิธี SVM นั้น เป็นวิธีที่นิยมใช้ในการเลือกรูปแบบหรือข้อมูลที่มีความแม่นยำ และเป็นเทคนิคที่สำคัญในการคัดเลือกและจัดกลุ่มทางชีวภาพของการแสดงออกของยีนในชุดข้อมูลไมโครอาร์เรย์ (Furey *et al.*, 2000) และ มีงานวิจัยพบว่าการวิเคราะห์ข้อมูลไมโครอาร์เรย์ด้วยค่าโพลด์เซนจ์ มีประสิทธิภาพในการวิเคราะห์มากกว่าการวิเคราะห์โดยใช้ *t*-statistic ซึ่งสามารถระบุยีนที่มีการแสดงออกที่แตกต่างกันอย่างมีนัยสำคัญที่เพิ่มมากขึ้น และยังพบว่าการวิเคราะห์ด้วยค่าโพลด์เซนจ์นี้เหมาะแก่การศึกษากลไกทางชีวภาพมากกว่าการวิเคราะห์ด้วย *t*-statistic อีกด้วย (Witten and Tibshirani, 2007) ผลจากการคัดเลือกยีนที่มีการตอบสนองต่อ LPS โดยวิธี SVM และการเปรียบเทียบโพลด์เซนจ์ ที่ระดับ 1.4 จะมียีนที่มีการตอบสนองต่อ LPS ทั้งหมด 1,475, 1,611 และ 321 ยีนตามลำดับเมื่อมีการกระตุ้นด้วย LPS ที่เวลา 3, 6 และ 8 ชั่วโมง (ตารางที่ 2)

ตารางที่ 2 แสดงจำนวนยีนที่ถูกคัดเลือกในแต่วิธี

ชุดข้อมูล	จำนวนยีนที่คัดเลือก (ยีน)		
	SVM	การเปรียบเทียบโพลด์เซนจ์	SVM และ การเปรียบเทียบโพลด์เซนจ์
3hr_LPS (GSE4712)	1,475	1,475	55
6hr_LPS (GSE4712)	1,611	1,611	85
8hr_LPS (GSE21841)	321	321	5

เมื่อนำยีนที่ได้จากการคัดเลือก (gene selection) ที่มีการตอบสนองต่อ LPS โดยวิธี SVM และการเปรียบเทียบโพลด์เซนจ์ มาเปรียบเทียบยีนที่มีรูปแบบการแสดงออกพร้อมกันในแต่ละช่วงเวลาที่ถูกกระตุ้นด้วย LPS พบว่า มียีนที่มีรูปแบบการแสดงออกพร้อมกันเมื่อกระตุ้นด้วย LPS ที่เวลา 3 ชั่วโมง ทั้งหมด 55 ยีน เมื่อกระตุ้นด้วย LPS ที่เวลา 6 ชั่วโมงทั้งหมด 85 ยีน และเมื่อกระตุ้นด้วย LPS ที่เวลา 8 ชั่วโมงทั้งหมด 5 ยีน (ตารางที่ 2) เมื่อนำยีนที่มีรูปแบบการแสดงออกพร้อมกันที่มีการกระตุ้นด้วย LPS ทั้ง 3 ช่วงเวลา มาทำการจัดกลุ่มยีนโดยใช้วิธี SOM (self-organizing map) ซึ่งเป็นการวิเคราะห์ข้อมูลโดยอาศัยการแบ่งกลุ่มข้อมูล จะอาศัยความเหมือนหรือความใกล้ชิดกันของข้อมูลโดยการวัดระยะห่างระหว่างเวกเตอร์ของข้อมูล ซึ่งวิธี SOM นี้มีความเหมาะสมกับการวิเคราะห์ข้อมูลที่มีขนาดใหญ่ เพราะจะทำให้มองเห็นภาพ (visualization) และการตีความหรือแปลผลได้ง่าย (Törönen *et al.*, 1999) จากการจัดกลุ่มโดยวิธี SOM มีรูปแบบการแสดงออกของยีนที่สนใจคือ เมื่อมีการกระตุ้นด้วย LPS ที่เวลา 3 และ 6 ชั่วโมงจะมีระดับการแสดงออกของยีนเพิ่มขึ้น แต่เมื่อมีการกระตุ้นด้วย LPS ที่เวลา 8 ชั่วโมง จะมีระดับการแสดงออกของยีนที่ลดลง ซึ่งมีทั้งหมด 18 ยีน คือ *C80256, Mmp7, Tet1, Mfap5, H2-D1, Cald1, Tmem108, Ribc2, Spata9, Igdcc3, Gap43, 1110008E08Rik, Rtkn2, Trpv1, Tgm4, Cr2, Gm7202* และ *4933436I20Rik* (รูปที่ 1) อาจเกิดจากการที่ยีนในกลุ่มนี้ส่วนใหญ่แล้วอยู่ในกลุ่มที่มีการตอบสนองต่อ LPS ได้เร็ว เมื่อเวลาผ่านไปจึงทำให้มีการแสดงออกที่ลดลง เมื่อนำยีนในกลุ่มนี้ไปค้นหา

หน้าที่ทางชีวภาพพบว่ายีนกลุ่มนี้มีหน้าที่ที่เกี่ยวข้องกับการตอบสนองของภูมิคุ้มกันต่อสิ่งแปลกปลอม (non-specific immune response) มีจำนวน 3 ยีน คือ *Cr2*, *Gap43* และ *Trpv1* และกระบวนการตอบสนองต่อการบาดเจ็บ (response to wounding) มีจำนวน 3 ยีน คือ *Cr2*, *H2-D1* และ *Trpv1* โดยจะพบว่ายีน *Cr2* (complement receptor 2) และ *Trpv1* (transient receptor potential cation channel, subfamily V, member 1) จะเกี่ยวข้องทั้งกระบวนการตอบสนองของภูมิคุ้มกันต่อสิ่งแปลกปลอมและกระบวนการตอบสนองต่อการบาดเจ็บ ซึ่งยีน *Cr2* จะเกี่ยวข้องกับการเหนี่ยวนำให้เกิดโรคเนื้องอก (follicular dendritic cell tumor) และมะเร็ง (follicular dendritic cell sarcoma) โดยจะกระตุ้นระบบคอมพลีเมนต์ในกลไก Lectin pathway และวิถีการแข็งตัวของเลือด (Barel *et al.*, 1998) และยีน *Trpv1* เป็นยีนที่จะเหนี่ยวนำให้เกิดโรคกระดูกประสาทไขสันหลัง โดยจะเกี่ยวข้องกับการกระตุ้น CREB pathway (cAMP response element-binding protein) และกระตุ้น protein kinase PKA ในกลไกการส่งสัญญาณกลูคาγον (Gavva *et al.*, 2004) จากงานวิจัยนี้เห็นได้ว่าการใช้ข้อมูลไมโครอาร์เรย์ที่เกี่ยวข้องกับการทดสอบการตอบสนองต่อ LPS มาคัดเลือกยีนที่เกี่ยวข้องกับการอักเสบหรือตอบสนองต่อ LPS ได้ ซึ่งจะทำให้เข้าใจภาพรวมกลไกการตอบสนองต่อ LPS ได้ดียิ่งขึ้น ในการใช้ข้อมูลไมโครอาร์เรย์จากฐานข้อมูลไมโครอาร์เรย์มาใช้ในการวิเคราะห์นี้ช่วยประหยัดค่าใช้จ่าย และควรมีการใช้วิธีการที่หลากหลายมาใช้ในการคัดเลือกอาจจะทำให้ค้นพบกลุ่มยีนใหม่ที่มีการตอบสนองต่อ LPS ได้



รูปที่ 1 Heat map รูปแบบการแสดงออกของยีน (A) และกราฟแสดงแนวโน้มรูปแบบการแสดงออกของยีน (B) เมื่อถูกกระตุ้นด้วย LPS ที่เวลา 3, 6 และ 8 ชั่วโมง จากการจัดกลุ่มตามรูปแบบการแสดงออกของยีนโดยวิธี SOM (self-organizing map)

4. บทสรุป

จากการใช้ข้อมูลไมโครอาร์เรย์จากฐานข้อมูลสาธารณะที่เกี่ยวข้องกับการตอบสนองต่อ LPS ของเซลล์แมคโครฟาจ โดยใช้วิธีการการคัดเลือกยีนโดยวิธี SVM (support Vector Machines) การเปรียบเทียบโพลดีเซนจ์ และการจัดกลุ่มโดยใช้วิธี SOM (self organization map) ร่วมกันในการคัดเลือกยีนที่เกี่ยวข้องกับการอักเสบหรือการตอบสนองต่อ LPS มีกลุ่มยีนที่มีรูปแบบการแสดงออกที่สนใจ ประกอบด้วย 18 ยีน ที่มีรูปแบบการแสดงออกของยีนมีระดับการแสดงผลเพิ่มขึ้น เมื่อกระตุ้นด้วย LPS ที่เวลา 3 และ 6 ชั่วโมง และมีระดับการแสดงผลของยีนลดลง เมื่อมีการกระตุ้นด้วย LPS ที่เวลา 8 ชั่วโมง มี 3 ยีน คือ *Cr2* *H2-D1* และ *Trpv1* ซึ่งเกี่ยวข้องกับกลไกการป้องกันเชื้อโรคของร่างกาย และมี 3 ยีน คือ *Cr2* *Gap43* และ *Trpv1* ซึ่งเกี่ยวข้องกับกลไกการตอบสนองต่อการบาดเจ็บ โดยทั้งสองกลไกนี้ต่างเกี่ยวข้องกับกระบวนการอักเสบ จะเห็นได้ว่าการใช้ข้อมูลการแสดงออกของยีนจากไมโครอาร์เรย์มาศึกษาการตอบสนองต่อ LPS ในเซลล์แมคโครฟาจ โดยการคัดเลือกยีนที่เกี่ยวข้องกับการอักเสบหรือการตอบสนองต่อ LPS จะทำให้สามารถเข้าใจภาพรวมของกลไกการตอบสนองต่อ LPS ได้มากขึ้น และควรมีการนำกลุ่มยีนที่คัดเลือกได้นี้ไปทำการศึกษาตรวจสอบเพื่อยืนยันผลในการแสดงออกตามช่วงเวลาที่ถูกกระตุ้นด้วย LPS ตามช่วงเวลาที่มีการแสดงออกในห้องปฏิบัติการต่อไป

5. กิตติกรรมประกาศ

ขอขอบพระคุณคณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา ที่ให้ทุนสนับสนุนการนำเสนอผลงานในงานวิจัย และงานวิจัยที่ได้รับทุนอุดหนุนการวิจัย มหาวิทยาลัยบูรพา ประจำปีงบประมาณ พ.ศ.2557 สัญญาเลขที่ 87/2557

6. เอกสารอ้างอิง

- Barel, M., Balbo, M. and Frade, R. (1998). Evidence for a new transcript of the Epstein-Barr virus/C3d receptor (CR2, CD21) which is due to alternative exon usage. *Mol Immunol*, 35(16), 1025-1031.
- Comer, J.E., Galindo, C.L., Zhang, F., Wenglikowski, A.M., Bush, K.L., Garner, H.R., et al. (2006). Murine macrophage transcriptional and functional responses to Bacillus anthracis edema toxin. *Microb Pathog*, 41(2-3), 96-110.
- Coussens, L. M. and Werb, Z. (2002). Inflammation and cancer. *Nature*, 420(6917), 860-867.
- Edgar, R., Domrachev, M. and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, 30, 207-210.
- Eleanor, A.H., Raktim, S., Daniel, S. and John, Q. (2011). RNA-Seq analysis in MeV. *Bioinformatics*, 27(22), 3209-3210.
- Furey, T.S., Cristianini, N., Duffy, N., Bednarski, D.W., Schummer, M. and Haussler, D. (2000). Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics*, 16(10), 906-914.
- Gavva, N.R., Klionsky, L., Qu, Y., Shi, L., Tamir, R., Edenson, S., Zhang, T.J., et al. (2004). Molecular determinants of vanilloid sensitivity in TRPV1. *Journal of Biological Chemistry*, 279(19), 20283-20295.
- Gollub, J., Ball, C.A., Binkley, G., Demeter, J., Finkelstein, D.B., Hebert, J.M., et al. (2003). The Stanford Microarray Database: data access and quality assessment tools. *Nucleic Acids Research*, 31(1), 94-96.
- Guyon, I., Weston, J., Barnhill, S. and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, 46, 389-422.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Elkan, H.W. (2009). The WEKA Data Mining Software: An Update. *SIGKDD Explorations*, 11(1), page 10-18.
- Hammer, K.D., Yum, M.Y., Dixon, P.M. and Birt, D.F. 2010. Identification of JAK-STAT pathways as important for the anti-inflammatory activity of a Hypericum perforatum fraction and bioactive constituents in RAW 264.7 mouse macrophages. *Phytochemistry*, 71(7), 716-725.
- Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nature Protoc*, 4(1), 44-57.
- Lee, H-S., Ryu D-S., Lee, G-S. and Lee, D-S. (2012). Anti-inflammatory effects of dichloromethane fraction from *Orostachys japonica* in RAW 264.7 cells: Suppression of NF- κ B activation and MAPK signaling. *Journal of Ethnopharmacology*, 140, 271-276.
- Nikkilä, J., Törönen, P., Kaski, S., Venna, J., Castrén, E. and Wong, G. (2002). Analysis and visualization of gene expression data using self-organizing maps. *Neural Networks*, 15(8-9), 953-966.
- Selvaraj, S. and Natarajan, J. (2011). Microarray Data Analysis and Mining Tools. *Bioinformatics*, 6(3), 95-99.
- Tarca, A.L., Romero, R. and Draghici, S., (2006). Analysis of microarray experiments of gene expression profiling. *American Journal of Obstetrics and Gynecology*, 195(2), 373-388.
- Törönen, P., Kolehmainen, M., Wong, G. and Castrén, E. (1999). Analysis of gene expression data using self-organizing maps. *FEBS Letters*, 451(2), 142-146.
- Watson, A., Mazumder, A., Stewart, M. and Balasubramanian S. (1998). Technology for microarray analysis of gene expression. *Curr Opin Biotechnol*, 9(6), 609-614.
- Witten, D. and Tibshirani, R. (2007). Volume A comparison of fold-change and the t-statistic for microarray data analysis. *Analysis*, 1: 1-17.

ภาคผนวก ซ

นำเสนอบทคัดย่อ งานประชุมทางวิชาการระดับนานาชาติ

The 4th International Biochemistry and Molecular Biology Conference,
2 – 3 April 2014, Bangkok, Thailand.

Sootanan, P. 2014. “Responsive-gene finding of LPS-stimulated macrophage via network approaches in microarray data”. FO-02. The 4th International Biochemistry and Molecular Biology Conference, April 2-3, 2014, Bangkok, Thailand.



FO-02

Responsive-gene Finding of LPS-stimulated Macrophage via Network Approaches in Microarray Data

Pitak Sootanan*

Department of Biochemistry, Faculty of Science, Burapha University, Chonburi 20131, Thailand

* Corresponding Author: pitak@buu.ac.th

Macrophage Raw 264.7 stimulated with LPS is a model to learn about the mechanism of inflammation. It was used by several research groups for testing anti-inflammatory compounds. Microarray technology can detect the response on whole cell in interesting condition by measurement of the expression of mRNA levels. There was a huge of submitted microarray datasets of LPS-stimulated macrophage in public database. These data can be used for analysis with several approaches. Network based analysis approaches of microarray data were interesting method because some interested responsive genes, that can not be found with conventional methods can be detected. In this study, three difference kinds of the network approaches such as protein-protein interaction, co-expression and regulatory networks were used to find the responsive genes in microarray data of LPS-stimulated macrophage. The results demonstrated that the responsive genes which are encoded transcription factors, such as MEF2, NFkB and NFAT, were found with their low expression but important for regulation of the several high differentially expressed genes that responded to inflammation. In conclusion, the network based analysis of microarray data can be used to discovered more knowledge which are concealed under a huge of data that cannot be found by conventional methods. This leads to find a novel responsive gene from the traditional public microarray data that make a new model for understanding the mechanism of inflammatory response.

Keywords: co-expression network, microarray, protein-protein interaction network, regulatory network

ภาคผนวก ฅ

นำเสนอบทคัดย่อ งานประชุมทางวิชาการระดับชาติ

การประชุมวิชาการวิทยาศาสตร์และเทคโนโลยีแห่งประเทศไทย ครั้งที่ 39

วันที่ 21 – 23 ตุลาคม 2556 กรุงเทพฯ ประเทศไทย

Wohankla, M. and Sootanan, P. 2556. “Comparative pathway-based analysis of gene expression profiles between two macrophages stimulated with LPS”. F_F0024. การประชุมวิชาการวิทยาศาสตร์และเทคโนโลยีแห่งประเทศไทย ครั้งที่ 39, ศูนย์นิทรรศการและการประชุมไบเทค, กรุงเทพฯ, ประเทศไทย, 21 – 23 ตุลาคม 2556, หน้า 176. (poster present)

Arunphanu, T., Rungrueng, J. and Sootanan, P. 2556. “Coexpression network approach to explain the response of LPS-stimulated RAW264.7”. F_F0026. การประชุมวิชาการวิทยาศาสตร์และเทคโนโลยีแห่งประเทศไทย ครั้งที่ 39, ศูนย์นิทรรศการและการประชุมไบเทค, กรุงเทพฯ, ประเทศไทย, 21 – 23 ตุลาคม 2556, หน้า 177. (oral present)

Ginou, W., Wichanprecha, K. and Sootanan, P. 2556. “Finding differentially expressed genes with data mining tools in microarray data of LPS-stimulated macrophage RAW264.7”. F_F0049. การประชุมวิชาการวิทยาศาสตร์และเทคโนโลยีแห่งประเทศไทย ครั้งที่ 39, ศูนย์นิทรรศการและการประชุมไบเทค, กรุงเทพฯ, ประเทศไทย, 21 – 23 ตุลาคม 2556, หน้า 186. (poster present)

Abstract: Cadmium contaminated soil was collected from paddy soil in the Mae Sod area, Tak Province, Thailand, and generally ranged from 0.5-284 mg kg⁻¹ soil. The aim of this study is to study the ability of biochar and microorganism additions for cadmium contaminated soil in order to stabilize cadmium soil. Rice plants were grown on no cadmium/free cadmium soil (N), cadmium soil (Cds), and cadmium soil amended with 1% w/w sawdust fly ash (SDFA), bagasse fly ash (BGFA), rice husk ash (RHA), 2% v/v *Beauveria bassiana* and *Pseudomonas aeruginosa*. The results showed that the addition of biochars and microorganisms enhanced Cd reduction in rice plants. *Pseudomonas aeruginosa* can increase the dry weight of the rice plant up to 20.01 g/pot compared to other conditions. *Pseudomonas aeruginosa* was the best for decreasing Cd accumulation in the roots, shoots, husks and grains in rice plants. It was found that after the microorganisms were grown in the system, the increased pH resulted in the stabilization of Cd in the soil. Therefore, less Cd was taken up by the rice plants. Sawdust fly ash is one of the best biochars for decreasing Cd accumulation in the roots, shoots, husks and grains in rice plants. This might be because sawdust fly ash itself can adsorb Cd higher than other biochars. (full paper available on CD)

F_F0024: COMPARATIVE PATHWAY-BASED ANALYSIS OF GENE EXPRESSION PROFILES BETWEEN TWO MACROPHAGES STIMULATED WITH LPS

Muanfan Wohankla,¹ Pitak Sootanan^{2,*}

¹Biological Science Program, Faculty of Science, Burapha University, Chonburi 20131, Thailand

²Department of Biochemistry, Faculty of Science, Burapha University, Chonburi 20131, Thailand

*e-mail: pitak@buu.ac.th

Abstract: The use of public microarray datasets to analyze the LPS-stimulated macrophages can reveal systematic inflammatory responses. Diverse microarray datasets of macrophages were publicly available. The macrophages are usually divided into two types, primary and cell line macrophages. In this study, microarray datasets involving LPS response by murine primary bone-marrow-derived and RAW 264.7 macrophages were retrieved from an open-access database. These datasets were comparatively analyzed for their differential expression according to the response to LPS. Pathway-based microarray analysis was used to explain the mechanism of LPS responses by both macrophages. The responses to LPS of both cell types show similar profiles but different level of expression in the most of selected pathways. Finally, our analysis could help for further studies to better understanding the mechanism of inflammatory process in different macrophages. (abstract only)

F_F0025: DNA ORIGAMI COMPARATIVE DESIGN BY TWO PROGRAMS: CADNANO AND NANOENGINEER-1

Pajaree Chaovadee, Pitak Sootanan, Anuttara Udomprasert*

Department of Biochemistry, Faculty of Science, Burapha University, Chonburi 20131, Thailand

*e-mail: anuttara@buu.ac.th

Abstract: DNA molecule has been widely used in nanotechnology, which is useful in many applications such as cosmetic and medical applications. One of the fascinating properties of DNA molecule is self-assembly due to its specific complementary base pairing. DNA origami is a popular technique used to create nanoscale structures with DNA molecules by using a long ssDNA from viral genome M13 (7.2 kb) as a scaffold and hundreds of short synthetic ssDNA as staples. The design process of this technique is labor and time-consuming if every step needs to be done by hands. Nowadays few computer programs were already developed to make this design step easier and faster. However, each program has its own advantages and disadvantages for designing different DNA nanostructures such as planar structures and 3D structures. In this study, two different programs, Cadnano and

Nanoengineer-1 (free software), used in the structural design step for DNA origami technique were compared. These include the accessibility, limitation, and flexibility of the software on designing various DNA origami nanostructures. It is necessary to learn about the limitations of each program before deciding which program is suitable for the designing step of any desirable DNA nanostructures. (abstract only)

F_F0026: COEXPRESSION NETWORK APPROACH TO EXPLAIN THE RESPONSE OF LPS-STIMULATED RAW264.7

Teppanom Arunphanu, Jarun Rungrueng, Pitak Sootanan*

Department of Biochemistry, Faculty of Science, Burapha University, Chonburi 20131, Thailand

*e-mail: pitak@buu.ac.th

Abstract: Inflammation is the response of the various cell types to the foreign matter that caused various kinds of harmful diseases. Macrophage RAW 264.7 cell line was often used as an indicator in the study of inflammatory response especially with LPS stimulation. Microarray technology can be used to systematically monitor the mechanism of the response in LPS-stimulated RAW264.7. This study examined the inflammatory response mechanism in public microarray data studied on LPS-stimulated RAW264.7. The approach to identify differences in the coexpression genes between phenotypes was used. These results showed that the use of coexpression network can be identified differentially expressed genes that cannot be identified with traditional tools. These gene-gene associations and coexpression networks have been correlated to inflammatory response that verified by functional annotation tools and literature reviews. In conclusion, public microarray data analysis by construction of coexpression network help to better understand the mechanisms involved in the inflammatory response of LPS-stimulated RAW264.7 and the further interested studies. (abstract only)

F_F0027: IDENTIFICATION OF A NEWLY ISOLATED *Bacillus* sp. H15 FROM BATONG HOT SPRING AND ITS THERMOSTABLE PROTEASE PRODUCTION

Ubon Wawwanjut, Montira Leelakriangsak*

Biology division, Department of Science, Faculty of Science and Technology, Prince of Songkla University, Pattani 94000, Thailand

*e-mail: montira-l@bunga.pn.psu.ac.th

Abstract: Thermostable proteases producing Bacteria from Betong hot spring, Yala province were isolated. The total of 35 isolates identified by morphological and biochemical tests, were members of *Bacillus* spp. Isolate H15 determined by measuring the clear zones on skim milk agar, from soil sample showed the best Relative Caseinolytic Activity (RCA) of 2.25 ± 0.10 cm. Therefore, H15 was chosen to determine the optimum conditions on thermostable proteases production in nutrient broth+1% Skim milk culture medium under 200 rpm shaking condition. Based on the morphological and biochemical characteristics, the H15 isolate is identified as *B. pantothenicus*. Protease activity was measured by using azocasein as a substrate. One unit of enzyme activity was defined as the amount of crude extract causing an increase in A_{420} of 0.001 got 1 h of incubation at 60°C. Protein concentration was determined by Bradford Protein assay (Bio-Rad) with BSA as the standard. Optimization of growth conditions for protease production was carried out within various incubation periods (8, 12, 24, 48, 72 and 96 h) in nutrient broth supplemented with 1% skim milk. The effect of temperature and pH on enzyme production was studied at 25, 37, 60 and 70°C and in pH of 6, 7, 8, 9 and 11. The maximum protease production (30.51 ± 0.92 Unit/mg protein) was observed at 48 h of incubation period, 60°C and pH value of 8. The thermostable protease from H15 was considered to be extracellular proteases which mainly were serine protease due to the inhibition by a specific protease inhibitor (PMSF). Furthermore, H15 will be identified by 16S rDNA gene sequencing for further taxonomic purpose. (abstract only)

F

⁴Department of Biochemistry, Faculty of Medical Science, Naresuan University, Phitsanulok 56000, Thailand

⁵College of Life science, Ritsumeikan University, Japan

*e-mail: jaruwanth@nu.ac.th

Abstract: Acrylamide is carcinogenic in animal and humans. Several researchers have established that the main pathway of acrylamide formation in foods is linked to, and, in particular, the amino acid asparagines Acrylamide formation in food is naturally during a series of reactions between an amino acid, primarily asparagine, and a reducing sugar such as fructose or glucose, that reaction is known a Maillard reaction at temperatures around 120 °C. Thus, acrylamide is formed during frying, roasting, and baking. 43 isolated of asparaginase-producing bacteria were isolated from Thai fermented food and screened asparaginase producing by modified M9 medium agar adding 0.15% asparagines and phenol red. TNL5.1 was the strain that shown the highest asparaginase enzyme activity was 9.46 Units. The study of bacterial identification by Gram stain and cell morphology, the data showed TNL5.1 would belong to the genus *Bacillus* sp., because they were Gram positive, spore-forming and bacilli bacteria. Also, to eliminate pathogenic bacteria, *Bacillus cereus*, the isolated were determined by negative lecithinase production strain on Egg Yolk agar and alpha-hemolytic bacteria. Optimization of asparaginase producing condition illustrated that the TNL5.1 produced the highest asparaginase (10.09 Units) in Potato broth in 18 hours. In the study of partial enzyme purification by gel filtration chromatography presented 20 fold times purity higher than crude enzyme and 19 % yield recovery. Stability test of the enzyme for industrial application explored by Freeze dry storage test that showed the enzyme had no activity in powder. (abstract only)

F_F0049: FINDING DIFFERENTIALLY EXPRESSED GENES WITH DATA MINING TOOLS IN MICROARRAY DATA OF LPS-STIMULATED MACROPHAGE RAW264.7

Wanchalerm Ginou, Kittisak Wichanprecha, Pitak Sootanan*

Department of Biochemistry, Faculty of Science, Burapha University, Chonburi 20131, Thailand

*e-mail: pitak@buu.ac.th

Abstract: Macrophage RAW264.7 is cell line that commonly used to study the inflammatory response and the effect of extracts on the inhibition of inflammation. There are several microarray experiments in public database studied on the response of RAW264.7, especially the response involved LPS stimulation. In this study, microarray expression data from public database were retrieved. Only samples that represent control macrophage RAW264.7 and LPS-stimulated cells were selected. Data mining tools were used for finding differentially expressed genes involved with LPS response in macrophage RAW264.7. Biological meanings of these genes were interpreted with web-based tools and literature reviews. These results showed several sets of differentially expressed genes, selected from different microarray datasets, which associated with signaling pathways and the secretion of inflammatory mediators. Thus, a variety of microarray dataset on public databases can be useful as preliminary data for gene selection in order to determine the response of RAW264.7 stimulated with LPS or other interested studies. (abstract only)

F_F0051: EFFECT OF CYP2C19 AND PONI POLYMORPHISMS ON CLOPIDOGREL RESPONSE IN CORONARY ARTERY DISEASE PATIENTS

Ruttakarn Kamkaew,^{1,*} Atip Likidilid,¹ Bhoom Suktitipat,¹ Damras Tresukosol²

¹Department of Biochemistry, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand

²Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand

*e-mail: rkmahidol@yahoo.com