เครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนของการสังเคราะห์กรดจัสโมนิก ในอะราบิดอพซิส ทาเลียนา และการจัดกลุ่มยีนออนโทโลยี

Protein-Protein Interaction Networks of Jasmonic Acid Synthesis in *Arabidopsis thaliana* and Gene Ontology Clustering

อิษฎา ชาลีเครือ และ พิทักษ์ สูตรอนันต์*

Issada Charleekrua and Pitak Sootanan*

ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

Biochemistry Department, Faculty of Science, Burapha University

Received : 30 May 2018

Revised : 9 August 2018

Accepted : 20 December 2018

บทคัดย่อ

เครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนสามารถสร้างได้จากฐานข้อมูลสาธารณะ และเครื่องมือทางชีวสารสนเทศที่เปิดให้ใช้ฟรี วัตถุประสงค์ของงานวิจัยนี้จะทำการสร้างเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนของโฮโมโลกัสโปรตีนในวิถีการสังเคราะห์กรดจัสโมนิกของอะราบิดอพซิส ทาเลียนา เพื่อตรวจสอบข้อมูลโฮโมโลกัสโปรตีนและเครือข่ายที่ได้เทียบกับเครือข่ายการแสดงออกร่วมของยีนจากงานวิจัยของ van Verk *et al*. (2011) ซึ่งพบว่าเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนที่ได้มีความแตกต่างออกไปมากเนื่องมาจากความแตกต่างของข้อมูลปฏิสัมพันธ์ทางพันธุกรรมและทางกายภาพจากฐานข้อมูลที่ใช้ในการสร้างเครือข่ายนอกเหนือไปจากข้อมูลการแสดงออกร่วมของยีน แต่การแปลผลทางชีวภาพของทั้งสองเครือข่ายด้วยยีนออนโทโลยีไม่มีความแตกต่างกัน ดังนั้นการสร้างเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนจึงไม่สามารถนำมาใช้ได้กับวิถีการสังเคราะห์กรดจัสโมนิก และเมื่อทำการจัดกลุ่มยีนออนโทโลยีก็พบว่ามีส่วนช่วยในการแปลผลทางชีวภาพได้ดียิ่งขึ้น แต่ถ้าหากมีการพัฒนาเครื่องมือการจัดกลุ่มและปรับปรุงฐานข้อมูลยีนออนโทโลยีให้สอดคล้องกับวัตถุประสงค์ของการใช้งานได้ก็จะทำให้การแปลผลทางชีวภาพมีความถูกต้องและตรงตามความต้องการมากยิ่งขึ้น

คำสำคัญ : เครือข่ายปฏิสัมพันธ์ระหว่างโปรตีน, วิถีการสังเคราะห์กรดจัสโมนิก, อะราบิดอพซิส ทาเลียนา, การจัดกลุ่มยีนออนโทโลยี

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

Abstract

Protein-protein interaction (PPI) networks can be reconstructed from public database and open-source bioinformatics tools. The aim of this work is to establish an PPI networks of homologous protein of jasmonic acid synthesis pathway in *Arabidopsis thaliana* from an existing co-expression network of van Verk *et al*. (2011) and compared them. We found that the extended PPI networks, comprising genetic and physical interactions, were different to the initial co-expression network, where the genes were linked by the expression patterns. However, the results of both biological interpretation on gene ontology are not different. Therefore, PPI network construction can not be applied to the case of jasmonic acid synthesis pathway. Gene ontology clustering was found to better contribute for biological interpretation. If the tools for gene ontology clustering and its database are developed in accordance with the purpose of the use, it will make biological interpretation more accurate and meet the demand.

Keywords : protein-protein interaction networks, jasmonic acid synthesis, *Arabidopsis thaliana*, gene ontology clustering

บทนำ

พืชตอบสนองต่อสิ่งแปลกปลอมโดยอาศัยกลไกการป้องกันตนเองในระดับโมเลกุลซึ่งพบว่ามีความเกี่ยวข้องกับการสร้างกรดซาลิไซลิก (Salicylic acid, SA) และเอธิลีน (Ethylene, ET) ในวิถีเมแทบอลิซึม และการส่งสัญญาณผ่านกรดจัสโมนิก (Jasmonic acid, JA) หรือฮอร์โมนจัสโมเนต (Pieterse *et al*., 2009) เนื่องจากวิถีการสังเคราะห์กรดจัสโมนิกมีบทบาทหน้าที่ในระดับโมเลกุลที่เกี่ยวข้องกับการยับยั้งการเจริญของราก มีการชักนำให้มีการสร้างโปรตีนพืชบางชนิดที่มีคุณสมบัติต้านทานการย่อยของแมลงศัตรูพืช และสามารถเพิ่มการแสดงออกของยีนที่ควบคุมการก่อตัวของท่อลำเลียง และการพัฒนาของดอก รวมถึงการตอบสนองต่อการเกิดแผล (wounding) ซึ่งสารที่ปลดปล่อยออกมาจะเข้าทำลายโรคของพืช เกี่ยวข้องกับการตอบสนองทางสรีรวิทยาหลายประการในพืช เช่น การสร้างเอธิลีน การสร้างคลอโรฟิลล์ และการสร้างเบตาคาร์โรทีน (Wasternack *et al*., 2013) แต่อย่างไรก็ตามการศึกษากลไกการป้องกันตนเองในพืชอะราบิดอพซิส ทาเลียนา (*Arabidopsis thaliana*) ผ่านวิถีการสังเคราะห์กรดจัสโมนิกอย่างเป็นระบบ ยังคงต้องศึกษาในรายละเอียดระดับโมเลกุลที่เกี่ยวข้องต่อไป

จากงานวิจัยของ van Verk *et al.* (2011) ซึ่งเป็นงานวิจัยที่เกี่ยวข้องกับการศึกษากลไกระดับโมเลกุลโดยอาศัยการวิเคราะห์ข้อมูลไมโครอาร์เรย์ด้วยวิธีการหาเครือข่ายการแสดงออกร่วมของยีน (Gene co-expression network) ในวิถีการควบคุมกระบวนการทรานสคริปชันของกลไกการป้องกันตนเองของพืชอะราบิดอพซิส ทาเลียนา แต่เนื่องจากกลไกระดับโมเลกุลดังกล่าวอาศัยการสร้างเครือข่ายจากข้อมูลการแสดงออกของยีนที่เกี่ยวข้องกับการตอบสนองของพืชภายใต้สภาวะเครียด ซึ่งเป็นข้อมูลระดับ mRNA ของยีนเท่านั้น ทำให้เครือข่ายที่ได้ยังคงขาดข้อมูลในระดับโมเลกุลอื่น ตัวอย่างเช่น ข้อมูลปฏิสัมพันธ์ระดับโปรตีน ที่จะทำให้การแสดงผลเครือข่ายของกลไกที่เกี่ยวข้องมีความถูกต้อง และสมบูรณ์มากยิ่งขึ้น (Pellegrini *et al*., 2004) ภายในสิ่งมีชีวิตโดยทั่วไปจะมีความหลากหลายของข้อมูลลำดับกรดอะมิโนของโปรตีนที่เป็น

โฮโมโลกัส อันเนื่องมาจากความผันแปรในการถ่ายทอดข้อมูลทางพันธุกรรม ที่ทำให้เกิดการจำลองตัวของข้อมูลทางพันธุกรรมและเกิดวิวัฒนาการของข้อมูลลำดับกรดอะมิโนเหล่านี้ทั้งชนิดเดียวกัน และต่างชนิดของสิ่งมีชีวิตซึ่งเรียกได้ว่าเป็นพาราโลกัสและออโธโลกัส ตามลำดับ ซึ่งโฮโมโลกัสโปรตีนเหล่านี้หลายชนิดยังคงมีโดเมนหลักเช่นเดียวกันกับโฮโมโลกัสโปรตีนในกลุ่มเดียวกัน แต่อาจจะมีบทบาทและหน้าที่ที่แตกต่างกันไปในเซลล์และเนื้อเยื่อแต่ละประเภทที่แตกต่างกัน (Fitche, 1970) และจากการศึกษาในกลุ่มยีนของวิถีการสังเคราะห์กรดซาลิไซลิก (Duangnapa *et al*., 2015) พบความหลากหลายที่เพิ่มมากขึ้นของชนิดและหน้าที่ของโปรตีนที่เกี่ยวข้อง ดังนั้นการเรียนรู้เครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนของกลุ่มโฮโมโลกัสโปรตีนที่เกี่ยวข้องกับวิถีพาธเวย์ของการสังเคราะห์กรดจัสโมนิก ก็อาจจะทำให้สามารถค้นพบบทบาทและหน้าที่อื่นที่มีความสำคัญต่อวิถีพาธเวย์ของการสังเคราะห์กรดจัสโมนิกนี้ได้ โดยสามารถประยุกต์ใช้ฐานข้อมูลและเครื่องมือทางชีวสารสนเทศ ที่มีการเผยแพร่และเปิดให้ใช้งานเป็นสาธารณะได้

ฐานข้อมูลสาธารณะทางชีวสารสนเทศที่รู้จักกันเป็นอย่างดี ได้แก่ ฐานข้อมูล NCBI (National Center for Biotechnology Information, Sayers *et al*., 2012) ซึ่งเป็นฐานข้อมูลสาธารณะที่เริ่มต้นให้บริการทางด้านข้อมูลลำดับนิวคลีโอไทด์ ลำดับกรดอะมิโน ข้อมูลโปรตีน ข้อมูลระดับจีโนม และข้อมูลบทความวิจัย เป็นต้น ซึ่งได้มีการปรับปรุงและพัฒนาฐานข้อมูลมาอย่างต่อเนื่องเพื่อให้สอดคล้องและรองรับกับการเพิ่มขึ้นของข้อมูลทางชีวภาพที่มีปริมาณมาก และมีความหลากหลายสูง นอกจากนี้ก็ยังมีฐานข้อมูลที่ถูกสร้างขึ้นมาเพื่อให้มีความจำเพาะต่อการใช้งานของผู้เข้าใช้มากยิ่งขึ้น ตัวอย่างเช่น ฐานข้อมูล TAIR (The Arabidopsis Information Resource, Lamesch *et al*., 2012) ซึ่งเป็นการรวบรวมข้อมูลระดับโมเลกุลในหลากหลายระดับที่เกี่ยวข้องกับอะราบิดอพซิส ทาเลียนา เพื่อดัดแปลงข้อมูล (Data) เหล่านี้ให้อยู่ในรูปของสารสนเทศ (Information) และองค์ความรู้ (Knowledge) เพื่อให้เกิดประโยชน์ต่อผู้ใช้งานมากที่สุด

สำหรับข้อมูลที่เกี่ยวข้องกับบทบาทและหน้าที่การทำงานของโปรตีน ได้แก่ข้อมูลโดเมน การจัดกลุ่มโปรตีน ก็มีการจัดเก็บอย่างเป็นระบบไว้ในฐานข้อมูล Pfam (Finn *et al*., 2014) ที่สามารถเข้าใช้เพื่อการตรวจสอบหน้าที่ของกลุ่มโปรตีนที่สนใจโดยอาศัยข้อมูลโดเมนที่เป็นส่วนแสดงถึงบทบาทและการทำงานที่สำคัญของโปรตีนแต่ละชนิดได้ สำหรับฐานข้อมูลปฏิสัมพันธ์ระหว่างโปรตีนที่มีการจัดเก็บข้อมูลปฏิสัมพันธ์ที่ได้จากการทดลองในห้องปฏิบัติการและผ่านการตรวจสอบโดยผู้เชี่ยวชาญเรียบร้อยแล้ว ได้แก่ ฐานข้อมูล DIP (Database of Interacting Proteins, Salwinski *et al*., 2004) และฐานข้อมูล STRING (Szklarczyk *et al*., 2015) ซึ่งเป็นฐานข้อมูลที่แสดงเครือข่ายปฏิสัมพันธ์ของโปรตีนที่ได้จากวิธีการทำนาย (Active Prediction Methods) ได้แก่ Neighborhood, Gene Fusion, Co-occurrence, Co-expression, Experiments, Databases และ Text mining และสามารถดาวน์โหลดข้อมูลของเครือข่ายปฏิสัมพันธ์เพื่อมาใช้วิเคราะห์ต่อไปได้ โดยฐานข้อมูลที่เกี่ยวข้องกับการแปลผลทางชีวภาพของกลุ่มยีนและโปรตีนในสิ่งมีชีวิตแต่ละชนิดได้ถูกจัดเก็บไว้ในรูปแบบของยีนออนโทโลยี (Gene ontology) ที่มีการปรับปรุงและพัฒนาฐานข้อมูลให้ทันสมัยเพื่อใช้ในการแปลผลโมเดลของระบบทางชีวภาพตั้งแต่ระดับโมเลกุล ไปสู่ระดับวิถีพาธเวย์ขนาดใหญ่ ระดับเซลล์ และระดับสิ่งมีชีวิต (Ashburner *et al*., 2000; GO Consortium, 2017) ที่นักวิจัยสามารถนำมาใช้ในการแปลผลข้อมูลทางชีวภาพของกลุ่มยีนหรือโปรตีนที่สนใจ เพื่อค้นหาบทบาทและหน้าที่ทางชีวภาพที่สำคัญต่อไปได้ (Maere *et al*., 2005; Duangnapa *et al*., 2015)

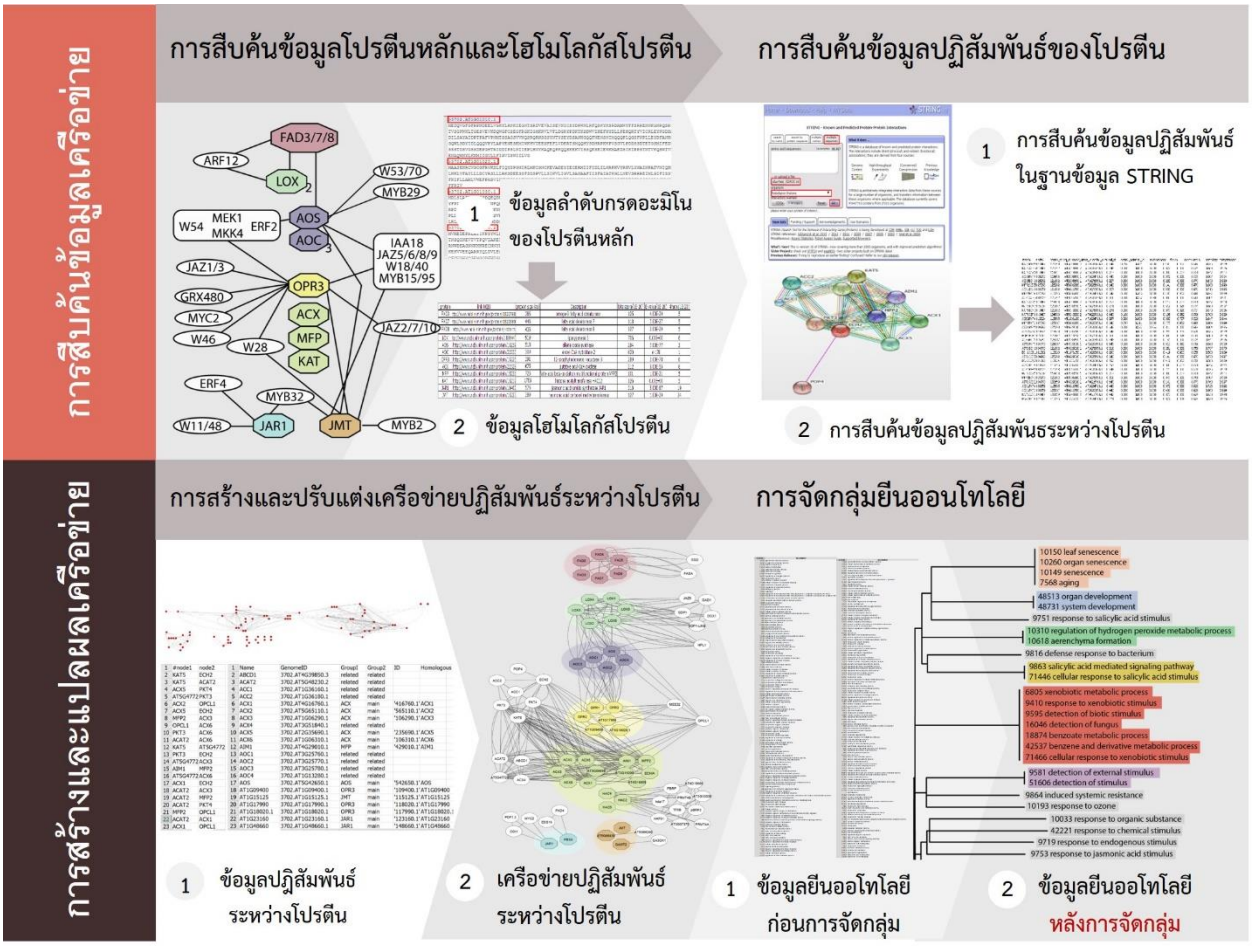
นอกจากฐานข้อมูลทางชีวสารสนเทศแล้ว ยังมีโปรแกรมทางชีวสารสนเทศที่นักวิจัยสามารถนำมาใช้ได้ฟรีโดยไม่มีค่าใช้จ่าย ตัวอย่างเช่น โปรแกรม BioEdit (Hall, 1999) ที่มีชุดเครื่องมือสำหรับการจัดการข้อมูลลำดับนิวคลีโอไทด์

และลำดับกรดอะมิโนหลากหลายประเภท ซึ่งมีเครื่องมือ local BLASTP ที่สามารถนำมาใช้ในการค้นหาโฮโมโลกัสโปรตีนในจีโนมของสิ่งมีชีวิตเป้าหมายได้ สำหรับการสร้าง แสดงผล และวิเคราะห์เครือข่ายปฏิสัมพันธ์ของโปรตีนสามารถทำได้โดยการใช้โปรแกรม Cytoscape (Shannon *et al*., 2003) เนื่องจากเป็นโปรแกรมที่มีชุดเครื่องมือของปลั๊กอินที่ได้มีการพัฒนาโดยกลุ่มผู้วิจัยเป็นจำนวนมากและมีการพัฒนาออกมาอย่างต่อเนื่อง โดยมีเครื่องมือ BiNGO (Maere *et al*., 2005) ซึ่งเป็นปลั๊กอินของ Cytoscape (Shannon *et al*., 2003) ที่สามารถทำนายบทบาทของโปรตีนที่เป็นสมาชิกในเครือข่ายปฏิสัมพันธ์ของโปรตีนที่สนใจโดยการแปลผลด้วยข้อมูลยีนออนโทโลยี ด้วยข้อมูลทางด้านกระบวนการทางชีวภาพ (Biological process) (Ashburner *et al*., 2000; GO Consortium, 2017) แต่เนื่องจากการแปลผลด้วย BiNGO จะทำให้ได้ผลลัพธ์ออกมาเป็นจำนวนมากและยากต่อการแปลผล การจัดกลุ่มยีนออนโทโลยีด้วยหลักการของการวิเคราะห์กลุ่มแบบขั้นตอน (Hierarchical clustering analysis) จึงถูกนำมาประยุกต์ใช้ (Duangnapa *et al*., 2015) จากฐานข้อมูลสาธารณะและโปรแกรมทางชีวสารสนเทศซึ่งเปิดให้ใช้ฟรีที่นักวิจัยและบุคคลทั่วไปสามารถเข้าถึงและใช้งานได้ จึงเป็นโอกาสในการตรวจสอบข้อมูลของเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนของวิถีการสังเคราะห์กรดจัสโมนิก เพื่อศึกษาถึงบทบาทและหน้าที่ทางชีวภาพที่เกี่ยวข้องต่อไปได้

ในการศึกษาครั้งนี้จะทำการตรวจสอบกลไกระดับโมเลกุลของระบบการป้องกันตนเองผ่านวิถีการสังเคราะห์กรดจัสโมนิกของพืชต้นแบบอะราบิดอพซิส ทาเลียนา โดยการวิเคราะห์เครือข่ายปฏิสัมพันธ์ระหว่างโปรตีน เปรียบเทียบกับเครือข่ายการแสดงออกร่วมของยีนจากงานวิจัยของ van Verk *et al*. (2011) และแสดงผลการจัดกลุ่มยีนออนโทโลยีเพื่อนำมาช่วยในการแปรผลทางชีวภาพของเครือข่ายข้างต้น ซึ่งจะทำให้ทราบบทบาทและหน้าที่ที่สำคัญของกลุ่มโฮโมโลกัสโปรตีนจากวิถีการสังเคราะห์กรดจัสโมนิกในพืชอะราบิดอพซิส ทาเลียนา ได้

วิธีดำเนินการวิจัย

การสร้างเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนของวิถีการสังเคราะห์กรดจัสโมนิกของอะราบิดอพซิส ธาเลียน่า จะเริ่มต้นด้วยการสืบค้นข้อมูลของกลุ่มโปรตีนหลักที่เกี่ยวข้องกับการสังเคราะห์กรดจัสโมนิกจากข้อมูลบทความวิจัยของ van Verk *et al*. (2011) ก่อนนำรายชื่อไปสืบค้นภายใต้ฐานข้อมูลสาธารณะ NCBI (Sayers *et al*., 2012), TAIR (Lamesch *et al*., 2012) เพื่อสืบค้นข้อมูลโปรตีนที่เกี่ยวข้องและทำการยืนยันผลด้วยการตรวจสอบโดเมนภายใต้ฐานข้อมูล Pfam (Finn *et al*., 2014) แล้วค้นหาลำดับความเหมือนของกรดอะมิโนด้วยชุดเครื่องมือ local BLASTP โดยคัดเลือกจากค่า E-value ≤ $1x10^{-20}0$ ที่จะทำให้ได้ข้อมูลลำดับเบสของกลุ่มโฮโมโลกัสโปรตีนของโปรตีนหลักแต่ละชนิด ก่อนที่จะนำข้อมูลลำดับเบสที่ได้ไปทำการสืบค้นข้อมูลปฏิสัมพันธ์ของโปรตีนบนฐานข้อมูล STRING (Szklarczyk *et al*., 2015) แล้วนำมาสร้างและปรับแต่งเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนของวิถีการสังเคราะห์กรดจัสโมนิก ด้วยโปรแกรม Cytoscape (Shannon *et al*., 2003) ต่อจากนั้นจะทำการแปลผลทางชีวภาพในกลุ่มกระบวนการทางชีวภาพ (Biological Process) ของยีนออนโทโลยี (Gene Ontology:GO) ด้วยปลั๊กอิน BiNGO (Maere *et al*., 2005) ของโปรแกรม Cytoscape แล้วนำผลยีนออนโทโลยีมาทำการจัดกลุ่มเพื่อการแปลผลทางชีวภาพ ขั้นตอนของวิธีการทดลองแสดงดังภาพที่ 1 และรายละเอียดของวิธีการทดลองในแต่ละขั้นตอนจะถูกแสดงภายในเนื้อหาต่อไป



ภาพที่ 1 ภาพรวมแสดงขั้นตอนของวิธีการทดลองในงานวิจัยนี้

1. การสืบค้นข้อมูลเครือข่าย

1.1 การสืบค้นข้อมูลโปรตีนหลักและการค้นหาโฮโมโลกัสโปรตีน

การสืบค้นรายชื่อโปรตีนที่เกี่ยวข้องกับระบบการป้องกันตัวเองของพืชผ่านวิถีการสังเคราะห์กรดจัสโมนิก ของอะราบิดอพซิส ทาเลียนา (*Arabidopsis thaliana*) โดยอาศัยการสืบค้นจากบทความวิจัยของ van Verk *et al*. (2011) ซึ่งพบโปรตีนหลักทั้งหมดจำนวน 12 โปรตีน ได้แก่ FAD3, FAD7, FAD8, LOX, AOS, AOC, OPR3, ACX, MFP, KAT, JAR1 และ JMT ก่อนนำรายชื่อโปรตีนเหล่านี้ไปสืบค้นข้อมูลลำดับกรดอะมิโนภายใต้ฐานข้อมูลสาธารณะ NCBI (Sayers *et al*., 2012) และฐานข้อมูล TAIR (Lamesch *et al*., 2012) จากนั้นนำข้อมูลลำดับกรดอะมิโนที่ได้มาตรวจสอบโดเมนด้วยฐานข้อมูลสาธารณะ Pfam (Finn *et al*., 2014) เพื่อยืนยันผลข้อมูลโปรตีนหลักแต่ละชนิด ก่อนจะนำข้อมูลลำดับกรดอะมิโนของโปรตีนหลักมาค้นหาโฮโมโลกัสโปรตีนด้วยชุดเครื่องมือ local BLASTP ภายในโปรแกรม BioEdit (Hall, 1999) โดยการใช้ลำดับกรดอะมิโนของโปรตีนหลักเป็นลำดับตั้งต้น (Query sequence) และค้นหาลำดับเป้าหมาย (Target sequence) ที่อยู่ภายใน

ฐานข้อมูลลำดับกรดอะมิโนของโปรตีนทั้งหมดในจีโนมของอะราบิดอพซิส ทาเลียนา ที่สามารถดาวน์โหลดได้จากฐานข้อมูล STRING (Szklarczyk *et al*., 2015) โดยกำหนดค่า E-value ≤ $1x10^{-20}$ แล้วนำรายชื่อโปรตีนที่ได้มาดึงข้อมูลลำดับกรดอะมิโนจากโปรตีนทั้งหมดที่มีอยู่ภายในจีโนมของอะราบิดอพซิส ทาเลียนา ซึ่งจะทำให้ได้ข้อมูลลำดับกรดอะมิโนของโฮโมโลกัสโปรตีน เพื่อนำไปใช้ในการสืบค้นข้อมูลปฏิสัมพันธ์ของโปรตีนต่อไป

1.2 การสืบค้นข้อมูลปฏิสัมพันธ์ระหว่างโปรตีน

การสืบค้นข้อมูลปฏิสัมพันธ์ระหว่างโปรตีนจะดำเนินการสืบค้นภายใต้ฐานข้อมูล STRING (Szklarczyk *et al*., 2015) โดยทำการสืบค้นด้วยลำดับกรดอะมิโนของโฮโมโลกัสโปรตีนจากข้อ 1.1 โดยเลือกการสืบค้นข้อมูลด้วย multiple sequences ภายใต้สิ่งมีชีวิต *Arabidopsis thaliana* ซึ่งกำหนดค่าการแสดงผลข้อมูลเครือข่ายปฏิสัมพันธ์ของโปรตีน ให้เลือกทุกวิธีการทำนาย (Active Prediction Methods) ได้แก่ Neighborhood, Gene Fusion, Co-occurrence, Co-expression, Experiments, Databases และ Text mining พร้อมกำหนดค่าความเชื่อมั่น (required confidence (score)) ที่ระดับ 0.900 (highest confidence) และแสดงผลโปรตีนในเครือข่ายมากกว่า 10 โปรตีน (more than 10 interactors) ก่อนทำการเก็บข้อมูลด้วยการดาวน์โหลด tabdelimited.txt (interaction) และ proteins_desc.txt (description) สำหรับใช้ในการสร้างและการปรับแต่งเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนในขั้นตอนต่อไป

2. การสร้างและการแปลผลเครือข่าย

2.1 การสร้างและการปรับแต่งเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีน

การสร้างและการปรับแต่งเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนสามารถทำได้โดยการใช้โปรแกรม Cytoscape (Shannon *et al*., 2003) โดยการนำข้อมูล tabdelimited.txt (interaction) จากข้อ 1.2 มาใช้ในการสร้างเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนด้วยการกำหนด Source interaction (node1) และ Target interaction (node2) ก่อนที่จะนำข้อมูล protein_desc.txt (description) ที่ผ่านการกำหนดกลุ่มการแสดงผลโปรตีนแล้ว มากำหนดเป็นข้อมูลคุณลักษณะ (attribute) สำหรับข้อมูลของโปรตีนหรือโหนด (node) แต่ละตัวภายในเครือข่ายปฏิสัมพันธ์ เพื่อนำมาใช้ในการปรับแต่งเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนให้ได้รูปแบบตามที่ต้องการ โดยอาศัยเครื่องมือ Group Attribute Layout และ Style (Shape, Fill color) เพื่อทำให้ลักษณะของเครือข่ายที่ได้มีความเหมาะสมสำหรับการแสดงผลเพื่อการเปรียบเทียบและการแปลผลทางชีวภาพต่อไป

2.2 การจัดกลุ่มยีนออนโทโลยีเพื่อการแปลผลทางชีวภาพ

จากเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนที่ผ่านการปรับแต่งแล้วจากข้อ 2.1 จะถูกนำมาแปลผลทางชีวภาพของโปรตีนทั้งหมดที่เป็นสมาชิกในเครือข่าย โดยอาศัยเครื่องมือ BiNGO (Maere *et al*., 2005) ที่เป็นปลั๊กอินของโปรแกรม Cytoscape (Shannon *et al*., 2003) โดยทำการคัดเลือกกระบวนการแปลผลทางชีวภาพด้วยยีนออนโทโลยีในกลุ่มกระบวนการทางชีวภาพ (Select ontology file: GO_Biological_Process) และคัดเลือกสิ่งมีชีวิตเป็น อะราบิดอพซิส ทาเลียนา (Select organism/annotation:) ก่อนเริ่มต้นการแปลผลทางชีวภาพด้วย BiNGO ที่ทำให้ได้ผลการวิเคราะห์เป็นจำนวนทั้งสิ้น 174 ยีนออนโทโลยี ซึ่งมีจำนวนมากและยากต่อการแปลผล จึงนำข้อมูลยีนออนโทโลยีที่ได้มาทำการจัดกลุ่มด้วยหลักการของการวิเคราะห์กลุ่มแบบขั้นตอน (Hierarchical clustering analysis) ของโปรตีนสมาชิกในแต่ละชุดของยีนออนโทโลยี เพื่อทำการ

จัดกลุ่มและลดจำนวนของยีนออนโทโลยีที่มีปริมาณและความซับซ้อนที่มากให้ลดน้อยลง เพื่อให้ง่ายต่อการแปลผลทางชีวภาพมากยิ่งขึ้น (Duangnapa *et al*., 2015)

ผลการวิจัยและวิจารณ์ผล

การตอบสนองของพืชต่อสิ่งเร้าถูกควบคุมผ่านระบบการป้องกันตนเองของพืชที่ซับซ้อน ซึ่งเกี่ยวข้องกับวิถีการสังเคราะห์สามวิถี คือ กรดซาลิไซลิก (Salicylic acid, SA) กรดจัสโมนิก (Jasmonic acid, JA) และเอธิลีน (Ethylene, ET) (van Verk *et al*., 2011) โดยวิถีการสังเคราะห์กรดจัสโมนิกที่มีบทบาทหน้าที่เกี่ยวข้องกับการยับยั้งการเจริญเติบโตของพืช และสามารถเพิ่มการแสดงออกของยีนที่ควบคุมเกี่ยวกับการตอบสนองทางสรีรวิทยาหลายประการในพืช เช่น การสร้างเอธิลีน การสร้างคลอโรฟิลล์ และ การสร้างเบตาคาร์โรทีน อีกทั้งกรดจัสโมนิกยังมีบทบาทสำคัญเกี่ยวกับความต้านทานโรคของพืชในอะราบิดอพซิส ทาเลียนา (*Arabidopsis thaliana*) เป็นต้น (Wasternack *et al*., 2013) โดยในการวิจัยครั้งนี้จะให้ความสำคัญกับโปรตีนที่ได้จากการแสดงออกของยีนที่เกี่ยวข้องกับวิถีการสังเคราะห์กรดจัสโมนิก ซึ่งพบว่ามีโปรตีนหลัก (Major proteins) ได้แก่ FAD3, FAD7, FAD8, LOX, AOS, AOC, OPR3, ACX, MFP, KAT, JAR1, JMT ที่มีความสำคัญต่อกระบวนการสังเคราะห์กรดจัสโมนิกในพืชซึ่งเป็นข้อมูลที่ได้มาจากงานวิจัยของ van Verk *et al*. (2011) ดังแสดงรายชื่อโปรตีนในตารางที่ 1 ที่ได้แสดงผลการตรวจสอบโดเมนจากฐานข้อมูล Pfam (Finn *et al*., 2014) ขนาดของโปรตีน (กรดอะมิโน) คำอธิบายของโปรตีน และจำนวนโฮโมโลกัสโปรตีนที่ได้จากการสืบค้นภายใต้ข้อมูลจีโนมของอะราบิดอพซิส ทาเลียนา จากฐานข้อมูล STRING (Szklarczyk *et al*., 2015) ที่ผ่านการสืบค้นโฮโมโลกัสโปรตีนของกลุ่มโปรตีนหลักโดยอาศัยชุดเครื่องมือ local BLASTP ภายในโปรแกรม BioEdit (Hall, 1999) ก่อนที่จะนำข้อมูลลำดับกรดอะมิโนของโฮโมโลกัสโปรตีนทั้งหมดไปสืบค้นข้อมูลปฏิสัมพันธ์ระหว่างโปรตีนภายใต้ฐานข้อมูล STRING (Szklarczyk *et al*., 2015) แล้วนำข้อมูลที่ได้มาสร้างและปรับแต่งเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนด้วยโปรแกรม Cytoscape (Shannon *et al*., 2003) ดังแสดงในภาพที่ 2B

จากการตรวจสอบข้อมูลโปรตีนหลักที่เกี่ยวข้องกับการวิถีการสังเคราะห์กรดจัสโมนิกจำนวน 12 โปรตีน และโฮโมโลกัสโปรตีนแสดงดังตารางที่ 1 ซึ่งพบว่าโปรตีนในแฟมิลีเดียวกันจะมีส่วนประกอบของโดเมนร่วมกัน ดังเช่นกรณีของ FAD3, FAD7 และ FAD8 ซึ่งโปรตีนทั้งสามชนิดจะมีโดเมนหลักสองประเภทคือ DUF3474 (PF11960) และ FA_desaturase (PF00487) ซึ่งเป็นโดเมนที่มีบทบาทยังไม่แน่ชัด กับบทบาทของการเร่งปฏิกิริยาดีแซเทอเรสของกรดไขมัน (Fatty acid desaturase) ตามลำดับ จากการตรวจสอบลำดับกรดอะมิโนของโปรตีนด้วยฐานข้อมูล Pfam (Finn *et al*., 2014) โปรตีนทั้งสามชนิดเกี่ยวข้องกับการสังเคราะห์กรดอัลฟาลิโนเลนิก (α-linolenic acid, 18:3) ซึ่งเป็นสารตั้งต้นที่สำคัญของวิถีการสังเคราะห์กรดจัสโมนิก (Wasternack *et al*., 2013 และ van Verk *et al*., 2011) โฮโมโลกัสโปรตีนที่ตรวจสอบพบเพิ่มเติมมีจำนวนทั้งสิ้น 3 โฮโมโลกัสโปรตีน ได้แก่ FAD2, FAD5 และ FAD6 ซึ่งเป็นโปรตีนที่มีส่วนเร่งปฏิกิริยาดีแซเทอเรสของกรดไขมันเช่นเดียวกัน แต่ทำให้ได้ผลผลิตของกรดไขมันไม่อิ่มตัวที่แตกต่างกัน คือ กรดลิโนเลอิก (Linoleic acid, 18:2) (Lamesch *et al*., 2012) สำหรับโปรตีนในแฟมิลีไลปอกซีจีเนส (Lipoxygenase, LOX) ที่ทำหน้าที่ในการเร่งปฏิกิริยาการเปลี่ยนแปลงกรดไขมันไม่อิ่มตัวไปเป็นไฮโดรเปอร์ออกไซด์ (Hydroperoxide) ในกระบวนการที่เรียกว่าไฮโดรเปอร์ออกซิเดชัน (Hydroperoxidation) ของกรดไขมัน ซึ่งพบโฮโมโลกัสโปรตีนภายในจีโนมของอะราบิดอพซิส ทาเลียนา จำนวนทั้งสิ้น 6 โฮโมโลกัสโปรตีน ได้แก่ LOX1, LOX2, LOX3, LOX4, LOX5 และ LOX6 (Utame, 2011) ส่วนโปรตีน AOS (Allene oxide

synthase) เป็นเอนไซม์ที่เร่งปฏิกิริยาการดีไฮเดรชัน (Dehydration) ของไฮโดรเปอร์ออกไซด์ ไปเป็นอัลลีนออกไซด์ (Allene oxide) โดยไม่พบว่ามีโฮโมโลกัสโปรตีน หรือมีลักษณะเป็นยีนเดี่ยว (Single gene) ในจีโนมของอะราบิดอพซิส ทาเลียนา (Schaller *et al*., 2009 และ Turner *et al*., 2002)

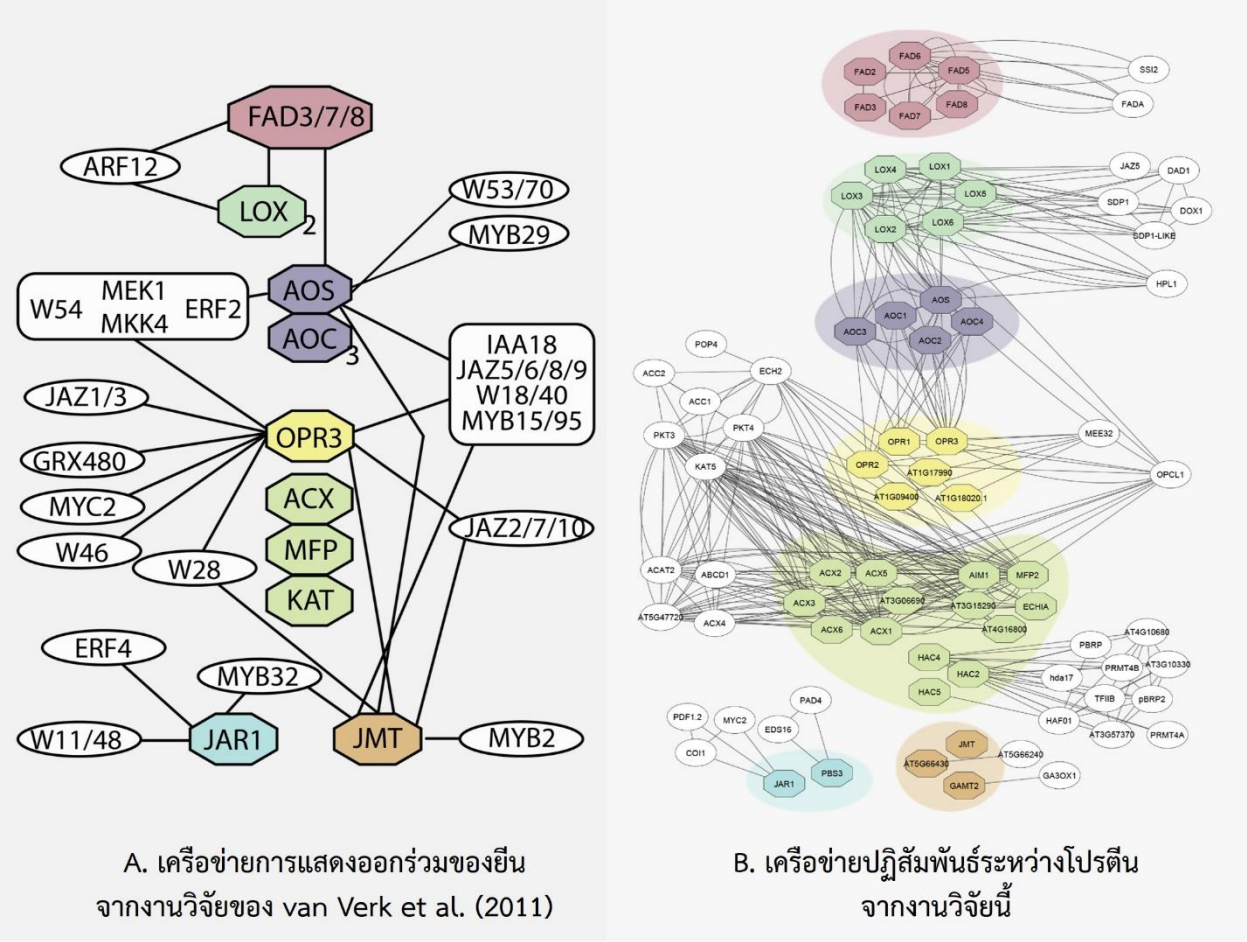
สำหรับ AOC (Allene oxide cyclase) จะเป็นเอนไซม์ที่เร่งปฏิกิริยาการสร้างวงแหวนที่มีความจำเพาะทางสเตอริโอเคมี (Stereospecific cyclization) ของการเปลี่ยนแปลงจากอัลลีนออกไซด์ ที่ไม่เสถียรไปเป็น cis-(+)-OPDA ((9S, 13S)-12 oxo-(10, 15Z)-phytodienoic acid) โดยที่จะพบโฮโมโลกัสโปรตีนของ AOC เป็นจำนวน 4 โฮโมโลกัสโปรตีนของจีโนมอะราบิดอพซิส คือ AOC1, AOC2, AOC3 ที่มีการแสดงออกในทุกส่วนของเนื้อเยื่อใบ (Leaf tissues) ในขณะที่ AOC4 จะมีการแสดงออกในส่วนของเส้นใบหลัก (Main veins) (Wasternack *et al*., 2013) และในส่วนของ OPR3 (OPDA reductases3) ที่จะตรวจสอบพบโฮโมโลกัสโปรตีนทั้งสิ้น 6 โฮโมโลกัสโปรตีน แต่จะมีเพียง OPR3 เท่านั้นที่เกี่ยวข้องกับการสังเคราะห์กรดจัสโมนิกในอะราบิดอพซิส ทาเลียนา (Wasternack *et al.,* 2013) สำหรับโปรตีนลำดับถัดมาทั้งสามชนิดได้แก่ ACX (Acyl-CoA oxidase), MFP (Multifunctional protein ที่ประกอบไปด้วยเอนไซม์ Enoyl-CoA hydratase และ β-hydroxy-acyl-CoA dehydrogenase) และ KAT (3-Ketoacyl-CoA thiolase) จะเกี่ยวข้องกับปฏิกิริยาเบต้า-ออกซิเดชัน (β-Oxidation) ของขั้นตอนการสังเคราะห์กรดจัสโมนิก (Schaller *et al*., 2009) ที่สามารถตรวจสอบพบโฮโมโลกัสโปรตีนเป็นจำนวน 6, 5 และ 3 โฮโมโลกัสโปรตีน ตามลำดับ ในส่วนของโปรตีน JMT (Jasmonic acid carboxyl methyltransferase) จะทำหน้าที่เร่งปฏิกิริยาเมธิลเลชัน (Methylation) ของกรดจัสโมนิกไปเป็น MeJA (Turner *et al*., 2002) และโปรตีน JAR1 (Jasmonic acid-amino acid synthetase) ที่เป็นเอนไซม์ที่เร่งปฏิกิริยาขั้นตอนสุดท้ายของการสังเคราะห์กรดจัสโมนิก (Wasternack *et al*., 2013) พบโฮโมโลกัสโปรตีนจำนวน 3 และ 2 โฮโมโลกัสโปรตีน ตามลำดับ จากการตรวจสอบข้อมูลของโปรตีนหลักจะพบว่าโปรตีนแต่ละชนิดที่เกี่ยวข้องกับการสังเคราะห์กรดจัสโมนิก จะเป็นเอนไซม์ซึ่งทำหน้าที่ในการเร่งปฏิกิริยาที่มีความจำเพาะต่อสารตัวกลางของวิถีการสังเคราะห์กรดจัสโมนิก บทบาทของโฮโมโลกัสโปรตีนจึงมีไม่มาก

ตารางที่ 1 ข้อมูลโปรตีนในวิถีการสังเคราะห์กรดจัสโมนิก และจำนวนโฮโมโลกัสโปรตีนที่ได้จากการสืบค้น

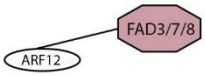
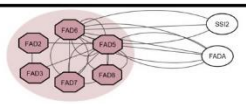
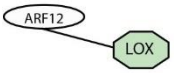
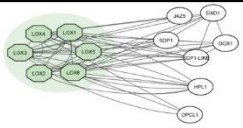
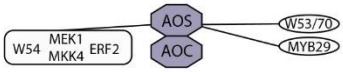
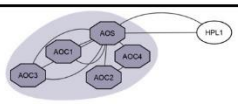
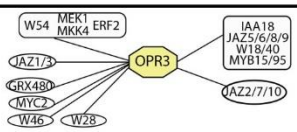
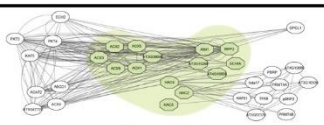
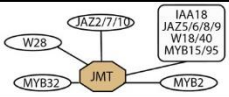
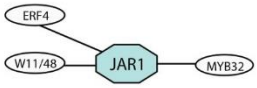
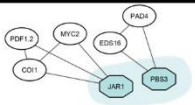
รายชื่อโปรตีน	โดเมน	ขนาด (กรดอะมิโน)	คำอธิบาย	จำนวนโฮโมโลกัสโปรตีน
FAD3/7/8	DUF3474—FA_desaturase	386/446/435	Fatty acid desaturase 3/7/8	6
LOX3	PLAT—Lipoxygenase	919	Lipoxygenase 3	6
AOS	P450	518	Allene oxide synthase	1
AOC1	Allene_ox_cyc	254	Allene oxide cyclase 1	4
OPR3	Oxidored_FMN	391	12-Oxophytodienoate reductase 3	6
ACX3	Acyl-CoA_dh_M—Acyl-CoA_dh_1—ACOX	675	Acyl-coenzyme A oxidase 3	6
MFP2	ECH_1—3HCDH_N—3HCDH	725	Peroxisomal fatty acid beta-oxidation multifunctional protein	5
KAT	zf-TAZ—HAT_KAT1—ZZ—zf-TAZ	1,706	Histone acetyltransferase	3
JMT	Methyltransf_7	389	Jasmonic acid carboxyl methyltransferase	3
JAR1	GH3	575	Jasmonic acid-amino acid synthetase	2

ภาพที่ 2 แสดงลักษณะของเครือข่ายที่เกี่ยวข้องกับการสังเคราะห์กรดจัสโมนิก ซึ่งเป็นเครือข่ายการแสดงร่วมของยีนที่ตอบสนองต่อสภาวะที่เป็นสิ่งเร้าต่อการตอบสนองของพืชอะราบิดอพซิส ทาเลียนา ที่ได้จากการวิเคราะห์ข้อมูลไมโครอาร์เรย์ของการแสดงออกของยีน 372 ชุดข้อมูล ที่มีจำนวนอาร์เรย์ชนิด 25K ทั้งหมด 1,437 อาร์เรย์ วิเคราะห์ร่วมกับข้อมูลยีนและทรานสคริปชันแฟคเตอร์ของกระบวนการสังเคราะห์กรดจัสโมนิกของพืชอะราบิดอพซิส ทาเลียนา ที่ดัดแปลงภาพมาจากงานวิจัยของ van Verk *et al.* (2011) (ภาพ 2A) และภาพแสดงเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนที่ได้จากงานวิจัยนี้ ซึ่งได้จากการสร้างและปรับแต่งด้วยโปรแกรม Cytoscape (Shannon *et al*., 2003) เพื่อจัดเรียงให้โปรตีนแต่ละกลุ่มแสดงผลสำหรับการเปรียบเทียบกับงานวิจัยของ van Verk *et al*. (2011) (ภาพที่ 2B) ซึ่งจัดแบ่งกลุ่มโฮโมโลกัสโปรตีนตามสีของโปรตีนหลักตามภาพ 2A ได้แก่ FAD3/7/8, LOX, AOS/AOC, OPR3, ACX/MFP/KAT, JAR1, และ JMT เมื่อนำข้อมูลเครือข่ายของกระบวนการสังเคราะห์กรดจัสโมนิกมาจัดแบ่งออกตามกลุ่มโปรตีนหลักข้างต้น ก็จะทำให้ได้เครือข่ายย่อยของโปรตีนหลักแต่ละชนิดกับโปรตีนที่เกี่ยวข้อง ดังแสดงในภาพที่ 3A ซึ่งแสดงภาพเครือข่ายย่อยของการแสดงออกร่วมของยีนที่แสดงออกได้เป็นโปรตีนหลักกับยีนและทรานสคลิปชันแฟคเตอร์ที่เกี่ยวข้อง กับภาพที่ 3B ที่แสดงเครือข่ายย่อยของกลุ่มโฮโมโลกัสโปรตีนของโปรตีนหลักที่มีปฏิสัมพันธ์กับโปรตีนที่เกี่ยวข้อง

เมื่อนำโฮโมโลกัสโปรตีนทั้งหมดของการสังเคราะห์กรดจัสโมนิกที่ได้มาสร้างเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนตามที่แสดงดังภาพที่ 2B และ 3B จะทำให้ได้เครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนของโปรตีนหลักแต่ละชนิดที่มีการเชื่อมโยงกันระหว่างโปรตีนหลัก LOX, AOS, AOC, OPR3, ACX, MFP และ KAT ซึ่งสอดคล้องกับลักษณะการทำงานในการเร่งปฏิกิริยาของโปรตีนหลักเหล่านี้ที่มีการเร่งปฏิกิริยาต่อเนื่องกัน และในส่วนของเครือข่ายที่ไม่มีการเชื่อมโยงระหว่างโปรตีนหลักหรือมีลักษณะเป็นเครือข่ายย่อยของโปรตีนหลักแต่ละชนิด ได้แก่ FAD3/7/8, JMT และ JAR1 ที่อาจจะเป็นโปรตีนหลักที่มีกลไกในการเร่งปฏิกิริยาที่แยกย่อยออกไปทำให้ไม่มีการปฏิสัมพันธ์ระหว่างโปรตีนหลักเหล่านี้กับโปรตีนหลักชนิดอื่น (Wasternack *et al*., 2013) สำหรับโปรตีนอื่น (วงรีสีขาว รูปที่ 2B และ 3B) จะเป็นกลุ่มเครือข่ายของโปรตีนที่ได้จากข้อมูลปฏิสัมพันธ์ระหว่างโปรตีนจากฐานข้อมูล STRING (Szklarczyk *et al*., 2015) ซึ่งเป็นข้อมูลที่ได้จากการวิเคราะห์ข้อมูลโปรตีนระดับจีโนม ข้อมูลการแสดงออกของยีน และข้อมูลจากผลการทดลองในห้องปฏิบัติการ ที่มีลักษณะของข้อมูลที่บ่งชี้ถึงปฏิสัมพันธ์ระหว่างโปรตีนในสองลักษณะทั้งทางพันธุกรรมและทางกายภาพ (Genetic and Physical interactions) ทำให้ลักษณะของเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนที่ได้ของโปรตีนหลักแต่ละชนิด ไม่สอดคล้องกับเครือข่ายการแสดงออกร่วมของยีนจากงานวิจัยของ van Verk *et al*. (2011) ซึ่งเป็นการวิเคราะห์ข้อมูลระดับการแสดงออกของยีนของชุดข้อมูลไมโครอาร์เรย์เพื่อค้นหาทรานส์คริปชันแฟคเตอร์ (Transcription factors) และโปรตีนที่มีบทบาทในการควบคุมกระบวนการสังเคราะห์กรดจัสโมนิกของอะราบิดอพซิส ทาเลียนา (ภาพที่ 2A และ 3A)



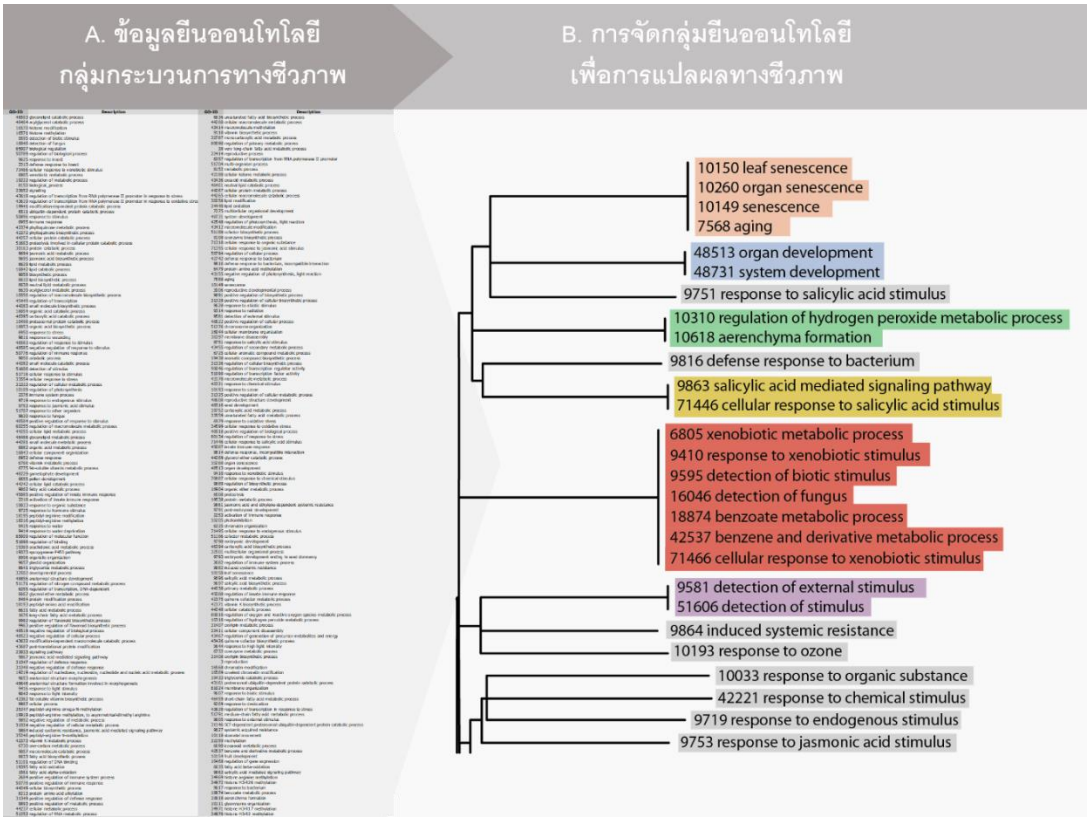
ภาพที่ 2 ภาพการแสดงผลเครือข่ายที่เกี่ยวข้องกับวิถีการสังเคราะห์กรดจัสโมนิก (A) เครือข่ายการแสดงออกร่วมของยีนที่ดัดแปลงมาจากงานวิจัยของ van Verk *et al.* (2011) และ (B) เครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนจากงานวิจัยนี้

โปรตีนหลัก	A. เครือข่ายการแสดงออกร่วมของยีน จากงานวิจัยของ van Verk et al. (2011)	B. เครือข่ายปฏิสัมพันธ์ระหว่างโปรตีน จากงานวิจัยนี้
FAD 3/7/8		
LOX		
AOS AOC		
OPR3		
ACX MFP KAT		
JMT		
JAR1		

ภาพที่ 3 ภาพการแสดงผลเครือข่ายที่เกี่ยวข้องกับวิถีการสังเคราะห์กรดจัสโมนิก โดยแยกตามโปรตีนหลักแต่ละชนิด (A) เครือข่ายการแสดงออกร่วมของยีนที่ได้มาจากงานวิจัยของ van Verk *et al*. (2011) และ (B) เครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนจากงานวิจัยนี้

จากข้อมูลเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนของกระบวนการสังเคราะห์กรดจัสโมนิกจะถูกนำมาแปลผลทางชีวภาพของโปรตีนทั้งหมดที่เป็นสมาชิกในเครือข่าย ด้วยการแปลผลยีนออนโทโลยีในกลุ่มกระบวนการทางชีวภาพ (GO_Biological_Process) ผ่านเครื่องมือ BiNGO (Maere *et al*., 2005) ซึ่งเป็นปลั๊กอินของโปรแกรม Cytoscape (Shannon *et al*., 2003) ที่ทำให้ได้ผลการวิเคราะห์ของข้อมูลยีนออนโทโลยีในกลุ่มกระบวนการทางชีวภาพ เพื่อใช้ในการแปลผลทางชีวภาพ เป็นจำนวนทั้งสิ้น 174 ยีนออนโทโลยี (ภาพที่ 4A) จากจำนวนยีนออนโทโลยีที่ได้ออกมามีเป็นจำนวนมาก มีความซับซ้อน และถ้าหากตั้งค่าพารามิเตอร์เพื่อสร้างข้อจำกัดของผลการวิเคราะห์ ก็จะทำให้ได้ผลการวิเคราะห์ที่ไม่ครบถ้วน จึงนำข้อมูลยีนออนโทโลยีที่ได้ทั้งหมดมาทำการจัดกลุ่มด้วยหลักการของการวิเคราะห์กลุ่มแบบขั้นตอน (Hierarchical clustering analysis) ของโปรตีนสมาชิกในแต่ละชุดของยีนออนโทโลยี ทำให้สามารถจัดกลุ่มและลดจำนวนข้อมูลยีนออนโทโลยีเพื่อใช้ในการแปลผลทางชีวภาพต่อไปได้ง่ายขึ้น (ภาพที่ 4B)

เมื่อทำการแปลผลทางชีวภาพของโฮโมโลกัสโปรตีนภายในเครือข่ายโดยอาศัยวิธีการจัดกลุ่มยีนออนโทโลยีด้วยหลักการวิเคราะห์กลุ่มแบบขั้นตอนของโปรตีนสมาชิกในแต่ละชุดของยีนออนโทโลยี (Duangnapa *et al*., 2015) จะทำให้ได้ผลแสดงดังภาพที่ 4 ซึ่งทำให้การแปลผลของยีนออนโทโลยีของกลุ่มกระบวนการทางชีวภาพ (Biological process) ที่มีจำนวนทั้งสิ้น 174 ยีนออนโทโลยี สามารถตรวจสอบและทำการแปลผลได้ง่ายขึ้น โดยการจัดกลุ่มยีนออนโทโลยีนี้จะทำให้ยีนออนโทโลยีที่มีความสอดคล้องกันในแง่ของยีน/โปรตีนสมาชิกมาอยู่ในกลุ่มคลัสเตอร์เดียวกัน ที่ทำให้ง่ายต่อการแปลผลมากกว่าการแปลผลโดยตรงจากรหัสหรือรายชื่อของยีนออนโทโลยีที่มีเป็นจำนวนมากและการแสดงผลจากโปรแกรมวิเคราะห์มักจะเรียงลำดับจากนัยสำคัญของยีนออนโทโลยีที่น่าจะมีความเกี่ยวข้องกับกลุ่มยีน/โปรตีนเป็นสำคัญ (Maere *et al*., 2005) ทำให้มีโอกาสที่จะพลาดข้อมูลบางอย่างจากการแปลผลยีนออนโทโลยีที่เกี่ยวข้องได้ จากตัวอย่างในภาพที่ 4B จะเห็นว่าบทบาทของโปรตีนจากเครือข่ายการสังเคราะห์กรดจัสโมนิกมีความเกี่ยวข้องการตอบสนองต่อสิ่งเร้าหลากหลายชนิด วิถีพาธเวย์ของกรดซาลิไซลิก พัฒนาการของอวัยวะ เมตาบอลิซึมของสารแปลกปลอมและเบนโซเอท เป็นต้น ซึ่งข้อมูลยีนออนโทโลยีที่ได้ไม่มีความแตกต่างจากการวิเคราะห์ข้อมูลเครือข่ายการแสดงออกร่วมของยีนที่ดัดแปลงมาจากงานวิจัยของ van Verk *et al*. (2011) ดังนั้นการสร้างเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนจึงไม่สามารถนำมาใช้ได้กับวิถีการสังเคราะห์กรดจัสโมนิก ซึ่งแตกต่างจากงานวิจัยของ Duangnapa *et al*. (2015) ที่แสดงให้เห็นถึงบทบาทและหน้าที่ที่สำคัญของกลุ่มโฮโมโลกัสโปรตีนจากวิถีการสังเคราะห์ซาลิไซลิกในพืชอะราบิดอพซิส ทาเลียนา ได้



ภาพที่ 4 ภาพการแสดงผล (A) ข้อมูลยีนออนโทโลยีของกลุ่มกระบวนการทางชีวภาพที่ได้จากการวิเคราะห์เครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนจากงานวิจัยนี้ และ (B) ผลจากการจัดกลุ่มยีนออนโทโลยีเพื่อการแปลผลทางชีวภาพ

สรุปผลการวิจัย

โดยสรุปการสร้างเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนของวีถีการสังเคราะห์กรดจัสโมนิกสามารถทำได้โดยการใช้ฐานข้อมูลสาธารณะและเครื่องมือทางชีวสารสนเทศที่เปิดให้ใช้ฟรี แต่เครือข่ายที่ได้จะมีความถูกต้องสมบูรณ์ก็ขึ้นอยู่กับการพัฒนาและปรับปรุงฐานข้อมูลที่นำมาใช้ และโฮโมโลกัสโปรตีนจะมีบทบาทที่สำคัญกับบางสภาวการณ์ภายในเซลล์ ซึ่งไม่สามารถใช้ร่วมกันได้กับการแสดงออกของโปรตีนที่มีความจำเพาะกับชนิดของเซลล์และอวัยวะ ดังนั้นการสร้างเครือข่ายปฏิสัมพันธ์ระหว่างโปรตีนจึงไม่สามารถนำมาใช้ได้กับวิถีการสังเคราะห์กรดจัสโมนิก การเลือกใช้การวิเคราะห์ข้อมูลลำดับเบสระดับจีโนมเพื่อตรวจสอบตำแหน่งควบคุมการแสดงออกของโปรตีนหลักในกระบวนการสังเคราะห์กรดจัสโมนิก รวมถึงข้อมูลปฏิสัมพันธ์ระหว่างทรานส์คริปชันแฟคเตอร์และยีนเป้าหมาย น่าจะเป็นทางเลือกที่เหมาะสมเพื่อทำความเข้าใจกลไกระดับโมเลกุลที่เกี่ยวข้องได้ดียิ่งขึ้น และการจัดกลุ่มยีนออนโทโลยีมีส่วนช่วยในการแปลผลทางชีวภาพ แต่ถ้าหากมีการพัฒนาเครื่องมือการจัดกลุ่มและปรับปรุงฐานข้อมูลยีนออนโทโลยีให้สอดคล้องกับวัตถุประสงค์ของการใช้งานได้ ก็จะทำให้การแปลผลทางชีวภาพมีความถูกต้องและตรงตามความต้องการมากยิ่งขึ้น

กิตติกรรมประกาศ

ขอขอบคุณผู้ช่วยศาสตราจารย์พงษ์รัตน์ ดำรงโรจน์วัฒนา กับข้อเสนอแนะสำหรับการเขียนบทความวิจัยนี้ ขอขอบคุณนางสาวจารินี คงทรัพย์ และนางสาวจิดาภา สอนศิริ ที่ช่วยเหลือในการปรับแต่งภาพประกอบและภาพเครือข่ายให้มีความสวยงาม

เอกสารอ้างอิง

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M, Rubin, G. M. and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. *Nat. Genet*, *25*(1), 25 – 29.

Duangnapa, K. and Sootanan, P. (2015). Homologous protein interaction network and functional analysis of salicylic acid biosynthesis pathway. *Proceeding of the 6th international conference on Computational Systems-Biology and Bioinformatics (CSBio2015)*, 59 – 63.

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L. L., Tate, J. and Punta, M., (2014). The pfam protein families. *Nucleic Acids Res*, *40*, D222 – D230.

Fitche, W. M. (1970). Distinguishing homologous from analogous proteins. *Syst. Zool*, *19*(2), 99 – 113.

GO Consortium. (2017). Nucleic Acids Res.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, *41*, 95-98.

Jensen, S. A. (2001). Orthologs and paralogs - we need to get it right. *Genome Biol*, *2*(8), 1002.1 – 1002.3.

Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D. L., Garcia, H. M., Karthikeyan, A. S., Lee, C. H., Nelson, W. D., Ploetz, L., Singh, S., Wensel, A. and Huala, E. (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res*, *40*, D1202 – D1210.

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.

Pellegrini, M., Haynor, D. and Johnson, J. M. (2004). Protein interaction networks, Expert Rev. *Proteomics*, *1*(2), 239 – 249.

Pieterse, C.M., Leon-Reyes, A., Van der Ent, S., Van Wees, S. C. (2009). Networking by small-molecule hormones in plant immunity. *Nat. Chem. Biol*, *5*(5), 308 – 316.

Salwinski, L., Miller, C. S., Smith, A. J., Pettit, F. K., Bowie, J. U. and Eisenberg, D. (2004). The database of interacting proteins: 2004 update. *Nucleic Acids Res*, *32* (Database issue), D449 – D451.

Sayers, E. W., Barrett, T., Benson, D. A., Bolton, E., Bryant, S. H., Caneses, K., Chetvernin, V., Church, D. M., Dicuccio, M., Federgen, S., Feolo, M., Fingerman, I. M., Geer, L. Y., Helmberg, W., Kupustin, Y., Krasnov, S., Landsman, D., Lipman, D. J., Lu, Z., Madden, T. L., Magej, T., Maglott, D. R., Marchler, B. A., Miller, V., Karsch, M. I., Ostell, J., Panchenko, A., Phan, L., Pruitt, K. D., Schuler, G. D., Sequeira, E., Sherry, S. T., Shumway, M., Sirotkin, K., Slotta, D., Souvorov, A., Starchenko, G., Tatusova, T. A., Wagner, L., Wang, Y., Wilbur, W. J., Yaschenko, E. and Ye, J. (2012). Database resources of the national center for biotechnology information. *Nucleic Acids Res*, *40*, D13 – D25.

Schaller, A. and Stintzi, A. (2009). Enzymes in jasmonate biosynthesis – Structure, function, regulation. *Phytochemistry*, *70*, 1532 – 1538.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, *13*(11), 2498 – 2504.

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K. P., Kuhn, M., Bork, P., Jensen, L. J. and von Mering, C. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life, *Nucleic Acids Res*, 43 (Database issue), D447 – D452.

Turner, J. G., Ellis, C. and Devoto, A. (2002). The jasmonate signal pathway. *The Plant Cell*, S153 – S164.

Umate, P. (2011). Genome-wide analysis of lipoxygenase gene family in Arabidopsis and rice. *Plant Signaling & Behavior*, *6*(3), 335 – 338.

van Verk, M. C., Bol, J. F., Linthorst, H. J. M. (2011). Prospecting for genes involved in transcriptional regulation of plant defenses, a bioinformatics approach, *BMC Plant Biology*, *11*, 88.

Wasternack, C. and Hause, B. (2013). Jasmonates: biosynthesis, perception, signal transduction and action in plant stress response, growth and development. An update to the 2007 review in Annals of Botany. *Annal. of Botany*, *111*, 1021 – 1058.