บทที่ 1
บทนำ

1.1 ความสำคัญและที่มาของปัญหาที่ทำการวิจัย

สิ่งมีชีวิตจำเป็นต้องมีการควบคุมการสังเคราะห์โปรตีนหรือการแสดงออกของยีนเพื่อให้โปรตีนที่ได้มีปริมาณเพียงพอสำหรับเมตาบอลิซึมต่างๆนั้น การแสดงออกของยีน คือ การสร้างผลิตผลจากยีนที่เป็นข้อมูลทางพันธุกรรมบนสายดีเอนเอโดยจะต้องมีการถอดรหัสและการแปลรหัสพันธุกรรม เพื่อสังเคราะห์โปรตีนที่จำเป็นต้องใช้ภายในเซลล์หรือระหว่างเซลล์ การแสดงออกของยีนเพื่อสร้างโปรตีนแต่ละชนิดเกิดขึ้นในเวลาและปริมาณที่แตกต่างกัน ทั้งนี้ขึ้นอยู่กับชนิดและบทบาทหน้าที่ของโปรตีนนั้น โดยการควบคุมการสังเคราะห์โปรตีนในร่างกายนั้นเกี่ยวข้องกับความสมดุลของหลายกระบวนการที่สำคัญได้แก่ กระบวนการถอดรหัส (transcription) และการตกแต่งหลังการถอดรหัส (post transcriptional modification) กระบวนการแปลรหัส (translation) และการตกแต่งหลังการแปลรหัส (post translational modification) รวมทั้งกระบวนการสลาย mRNA (mRNA degradation) และการสลายโปรตีน (protein degradation) [2] การควบคุมการสังเคราะห์โปรตีน คือ การควบคุมอัตราการผลิตและการสลายของโปรตีนเพื่อให้โปรตีนที่ได้มีปริมาณเพียงพอต่อขบวนการเมตาบอลิซึมต่างๆ หากมีการกระตุ้นหรือยับยั้งขั้นตอนใดๆ ในกระบวนการเหล่านี้จะมีผลกระทบต่อปริมาณโปรตีนที่เป็นผลิตผลทั้งสิ้น ทั้งนี้ขึ้นอยู่กับสภาพแวดล้อม หน้าที่ และความต้องการของเซลล์ชนิดนั้น การควบคุมเมตาบอลิซึมในเซลล์โปรคาริโอต (prokaryote) มีกลไกที่แตกต่างจากเซลล์ยูคาริโอต เนื่องจากโปรคาริโอตมีพันธุกรรมที่ไม่ซับซ้อนอีกทั้งกระบวนการสังเคราะห์อาร์เอนเอ (RNA) และโปรตีนเกิดขึ้นในไซโตพลาสซึม (cytoplasm) ซึ่งใช้เวลาที่ใกล้เคียงกัน ในขณะที่การควบคุมการแสดงออกของยีนในยูคาริโอต เช่น พืช และสัตว์ จะซับซ้อนกว่าในโปรคาริโอต เช่น แบคทีเรีย เพราะในยูคาริโอต (eukaryote) มียีนอยู่มากมายหลายชนิดบนโครโมโซมหลายอันทั้งนี้ขึ้นอยู่กับชนิดของสิ่งมีชีวิตนั้น การควบคุมการสังเคราะห์โปรตีนในโปรคาริโอตส่วนใหญ่เกิดขึ้นที่ระดับการถอดรหัสพันธุกรรม การควบคุมนี้เป็นแบบโอเพอรอน (operon) สำหรับในยูคาริโอตการควบคุมการสังเคราะห์โปรตีนเกิดขึ้นได้ทั้งระดับยีน (gene) ระดับการถอดรหัส (transcriptional control) และการควบคุมการแปลรหัส (translational control) ซึ่งการควบคุมในระดับการถอดรหัสนั้นเกิดจากโปรตีนควบคุม (transcription factors) เข้ามาทำงานเพื่อกระตุ้นหรือยับยั้งกระบวนการถอดรหัส จะเห็นว่า การควบคุมการแสดงออกของยีนในระดับการถอดรหัสนั้นสำคัญมาก อีกทั้งยังเป็นสาเหตุสำคัญที่ก่อให้เกิดโรคทางพันธุกรรมได้ [3, 4]

จะเห็นได้ว่า กลไกการควบคุมการแสดงออกของยีนในระดับการถอดรหัสนั้นสำคัญมากต่อการดำรงชีวิตของสิ่งมีชีวิตต่างๆ นักวิทยาศาสตร์จึงพยายามที่จะเข้าใจกลไกเหล่านี้จากข้อมูลด้านจีโนมที่มีมากขึ้น กอปรกับเทคโนโลยีด้านไมโครอะเรย์ (microarray) ได้ส่งผลให้นักวิทยาศาสตร์ได้พยายามคิดค้นวิธีการต่างๆ เพื่อมาศึกษาการแสดงออกของยีนของโปรคาริโอตและยูคาริโอต [5, 6] โดยการอนุมานลักษณะการควบคุมการแสดงออกของยีนจากข้อมูลดีเอ็นเอจากเทคนิคไมโครอะเรย์ด้วยวิธีทางคอมพิวเตอร์ [7-9] เช่น ordinary differential equations (ODE), partial differential equations (PDE) Bayesian networks [10] และ Boolean networks [1, 9, 11, 12] เป็นต้น ซึ่งวิธีที่ง่ายที่สุดที่ใช้ศึกษาเครือข่ายควบคุมการแสดงออกของยีน (gene regulatory network) ด้วยเทคนิคบูลีน (Boolean network) แต่อย่างไรก็ตาม การศึกษาเครือข่ายควบคุมการแสดงออกของยีนในระดับการถอดรหัสโดยใช้บูลีนนั้นมีข้อจำกัดคือ สามารถวิเคราะห์ข้อมูล (input) ที่มีค่าเพียง 2 ระดับ นั่นคือ 0 หรือ 1 ดังนั้น

สำหรับการวิเคราะห์เครือข่ายควบคุมการแสงดออกของยีนนั้น จำเป็นต้องมีการแปลงค่าการแสดงออกของยีน (raw data) ให้อยู่เพียง 2 ระดับ (binary values) คือ 0 หรือ 1 ซึ่ง 0 อาจหมายถึงไม่มีการแสดงออกของยีนนั้นๆ และ 1 คือ มีการแสดงออกของยีนนั้นๆ เรียกขั้นตอนนี้ว่า "discretization" โดยใช้วิธีการทางสถิติ (discretization method) จะเห็นได้ว่า การแปลงค่าการแสดงออกของยีนซึ่งเป็นข้อมูลจริง มาเป็นข้อมูลเพียง 2 ระดับนั้น อาจไม่มีเพียงพอ และทำให้เกิดการทำนายที่ผิดพลาด (false prediction) เครือข่ายควบคุมการแสงดออกของยีนมีความซับซ้อน ยากต่อการวิเคราะห์ต่อไป

วิธีการที่ใช้ในการแปลงข้อมูลการแสดงออกของยีนในขั้นตอนนี้มีหลากหลายแบบ อย่างไรก็ตาม ในการศึกษาเครือข่ายควบคุมการทำงานของยีนด้วยเทคนิคบูลีนที่มีมาก่อนหน้านี้นั้น ส่วนมากเน้นการปรับปรุงวิธีการเทคนิคบูลีนมากกว่าการพิจารณาถึงการเลือกวิธีการที่ใช้ในการแปลงข้อมูล ซึ่งอาจเป็นปัจจัยหลักหนึ่งที่ทำเครือข่ายควบคุมการแสดงออกของยีนนั้นไม่สอดคล้องกับข้อมูลที่มีอยู่ นอกจากนี้ การปรับปรุงเทคนิคบูลีนที่สามารถสร้างเครือข่ายควบคุมการแสดงออกของยีนที่มีความซับซ้อนน้อยลง ให้สอดคล้องกับข้อมูลทางชีววิทยาน่าจะทำให้เครือข่ายควบคุมการแสงดออกของยีนที่สร้างขึ้นนั้นมีความซับซ้อนน้อยลง การจะศึกษาเครือข่ายควบคุมการแสดงออกของยีนโดยใช้เทคนิคบูลีนที่มีมาก่อนหน้านี้นั้น ยังขาดการพิจารณาถึงเงื่อนไขทางด้านชีววิทยา (biological constraint) เช่น โปรตีนควบคุม (transcription factors) สามารถควบคุมการแสดงออกของยีนเป้าหมายในระดับเมตาบอลิก (เอนไซม์) ได้ แต่เอนไซม์ไม่ควรควบคุมการแสดงออกของโปรตีนควบคุม เป็นต้น การใส่ biological constraint ให้กับการวิเคราะห์ด้วยเทคนิคบูลีนน่าจะลดความซับซ้อนของเครือข่ายควบคุมการแสดงออกของยีน โดยการลดความผิดพลาดของการทำนายที่มักจะเกิดขึ้นจากการวิเคราะห์ด้วยวิธีนี้ ทำให้เครือข่ายที่สร้างขึ้นนั้นง่ายต่อการศึกษาและวิเคราะห์ต่อไป

ดังนั้น การวิจัยในครั้งนี้ ผู้วิจัยมีความมุ่งหวังที่พัฒนาและปรับปรุงเทคนิควิธีการสร้างเครือข่ายการแสดงออกของยีนแบบมีเงื่อนไขโดยใช้เทคนิคบูลีน เรียกว่า Constraint-based Boolean network พร้อมทั้งคำนึงถึงการประเมินหาวิธีการที่เหมาะสมต่อการแปลงข้อมูลการแสดงออกของยีนให้เป็นข้อมูล 2 ระดับ เพื่อช่วยให้สามารถสร้างเครือข่ายควบคุมการแสดงออกของยีนให้สอดคล้องกับระบบของสิ่งมีชีวิตมากยิ่งขึ้น ซึ่งสามารถลดความซับซ้อนของเครือข่ายควบคุมการแสดงออกของยีน นอกจากนี้สามารถนำเอาวิธีที่พัฒนาขึ้นไปประยุกต์ใช้วิเคราะห์เครือข่ายควบคุมการแสดงออกของยีนของกระบวนการในพืช เช่น กระบวนการสังเคราะห์แป้ง ซึ่งเป็นแหล่งพลังงานหลักของมนุษย์

1.2 วัตถุประสงค์ของโครงการ

เพื่อปรับปรุงและพัฒนาวิธีการในการสร้างเครือข่ายควบคุมการแสดงออกของยีนแบบมีเงื่อนไขด้วยเทคนิคบูลีน เพื่อนำไปใช้ศึกษาเครือข่ายควบคุมการแสดงออกของยีนในกระบวนการสังเคราะห์แป้ง

1.3 ขอบเขตของโครงการวิจัย

1.3.1 ปรับปรุงและพัฒนาโปรแกรมคอมพิวเตอร์ (Constraint-based Boolean network) เพื่อสร้างเครือข่ายควบคุมการแสดงออกของยีนโดยใช้ข้อมูลจากไมโครอะเรย์ของยีสต์ภายใต้สภาวะการเลี้ยงแบบ diauxic shift

1.3.2 ประเมินหาวิธีการที่เหมาะสมที่สุดที่ใช้แปลงค่าการแสดงออกของยีนเป็นข้อมูล 2 ระดับ (discretization method)

1.3.3 ประยุกต์ใช้โปรแกรมคอมพิวเตอร์ที่พัฒนาขึ้น เพื่อสร้างและศึกษาเครือข่ายการแสดงออกของยีนในกระบวนการสังเคราะห์แป้ง

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1.4.1 โปรแกรมคอมพิวเตอร์ที่ใช้ในการสร้างและศึกษาเครือข่ายควบคุมการแสดงออกของยีนโดยใช้ข้อมูลไมโครอะเรย์

1.4.2 เครือข่ายควบคุมการแสดงออกของยีนในกระบวนการสังเคราะห์แป้ง เพื่อนำไปสู่แนวทางในการปรับปรุงสายพันธุ์พืชที่มีความสามารถในการสังเคราะห์แป้งเช่น ข้าวและมันสำปะหลัง ที่สามารถสังเคราะห์แป้งให้มีคุณสมบัติตามต้องการ เหมาะต่อการนำไปใช้ในอุตสาหกรรมต่างๆ และนำไปสู่การเพิ่มมูลค่าของผลิตภัณฑ์ทางการเกษตร

1.4.3 มีการผลิตบุคลากรที่มีความเชี่ยวชาญเฉพาะทางด้านเทคโนโลยีชีวสารสนเทศ (Bioinformatics)

1.5 แผนการถ่ายทอดเทคโนโลยีหรือผลการวิจัยสู่กลุ่มเป้าหมาย

1.5.1 เผยแพร่ผลงานวิจัยในรูปสิ่งตีพิมพ์และงานประชุมวิชาการในระดับชาติและระดับนานาชาติ

1.5.2 ร่วมมือกับกลุ่มวิจัย

1.6 แผนการดำเนินงานตลอดโครงการ

ตารางที่ 1.1 แผนการดำเนินงาน

ขั้นตอนการวิจัย	เดือนที่ 1-3	เดือนที่ 4-6	เดือนที่ 7-9	เดือนที่ 10-12
1. การเก็บและรวบรวมข้อมูล				
2. การเตรียมข้อมูล				
3.ประเมินวิธีการที่ใช้ในการแปลงข้อมูลการแสดงออกของยีน				
4. พัฒนาโปรแกรมที่ใช้ในการสร้างเครือข่ายการแสดงออกของยีนแบบมีเงื่อนไข				
5. การทดสอบและประเมินโปรแกรม				
6. การประยุกต์ใช้โปรแกรมเพื่อสร้างและศึกษาเครือข่ายควบคุมการแสดงออกของยีนในกระบวนการสังเคราะห์แป้ง				
7. สรุปผลและเขียนรายงาน				