

## บทที่ 5

### บทสรุป

ในงานวิจัยนี้ได้ทำการวิเคราะห์ข้อมูลการแสดงผลของยีนที่เรียกว่า ข้อมูลไมโครอาร์เรย์ของยีสต์สายพันธุ์แซคคาโรไมซิส เซรีวิซิเย โดยได้แบ่งการทดลองออกเป็น 2 ส่วนหลัก คือในการทดลองส่วนแรก เป็นการศึกษาและวิเคราะห์ข้อมูลไมโครอาร์เรย์ของยีสต์ในระหว่างกระบวนการได้ออกซิซิฟท์ เพื่อทำการจำลองการส่งผลกระทบของปริมาณความเข้มข้นของกลูโคสและความหนาแน่นเซลล์ที่มีต่อระดับการแสดงออกของยีน ซึ่งทำให้เราเข้าใจพฤติกรรมการแสดงออกของยีนยีสต์มากขึ้น และการทดลองในส่วนที่สอง เป็นการสร้างเครือข่ายการควบคุมกันระหว่างยีนของกลุ่มยีนที่มีการรายงานในทางชีววิทยาแล้วว่ามีความสัมพันธ์ระหว่างกันอย่างไร ซึ่งได้สรุปภาพรวมของงานวิจัยทั้งหมดไว้ในหัวข้อที่ 5.1 และข้อเสนอแนะและแนวทางในการวิจัยต่อไปในอนาคตไว้ในหัวข้อที่ 5.2 ดังต่อไปนี้

#### 5.1 สรุปผลภาพรวมของงานวิจัย

จากการทดลองทั้งสองส่วนในบทที่ 3 และ 4 ที่ได้กล่าวมาแล้ว แสดงให้เห็นว่าวิธีการมาร์คอฟเชนมอนติคาร์โลซึ่งเป็นวิธีการทางสถิติสามารถนำมาประยุกต์ใช้ในการวิเคราะห์ข้อมูลทางชีววิทยา เช่น ข้อมูลไมโครอาร์เรย์ ได้ ถึงแม้ว่าการทดลองทั้งสองส่วนนั้นจะใช้วิธีการของมาร์คอฟเชนมอนติคาร์โลที่แตกต่างกันในการวิเคราะห์ข้อมูล นั่นคือ การทดลองในบทที่ 3 ได้ใช้วิธีการกิบป์ แซมพลิง และการทดลองในบทที่ 4 ได้ใช้อัลกอริทึมเมโทรโพลิส – แฮสติงส์ ในการวิเคราะห์ข้อมูลโดยมีวัตถุประสงค์ของการทำการทดลองเพื่อทำการสร้างแบบจำลองของข้อมูลที่มีลักษณะแตกต่างกัน แต่อย่างไรก็ตามหลักการทำงานหลัก ๆ ของทั้งสองวิธีการนี้ยังคงมีลักษณะที่คล้ายคลึงกันคือ จะทำการสร้างข้อมูลจำลองขึ้นมาแบบสุ่ม แล้วพิจารณาว่าข้อมูลจำลองที่ได้สร้างขึ้นมานั้นมีความเหมือนจริงและสมควรจะได้รับการยอมรับไว้หรือไม่ และอีกสิ่งหนึ่งที่ได้เห็นได้อย่างชัดเจน คือ การทำงานของวิธีการมาร์คอฟเชนมอนติคาร์โลมักจะควบคู่ไปกับวิธีการของเบย์เซียนเสมอ เนื่องจากว่า ด้วยวิธีการของเบย์เซียนนั้น ไม่อาจสามารถทำการสร้างข้อมูลจำลองขึ้นมาได้ ซึ่งในการวิเคราะห์ข้อมูลเพื่อเพิ่มความน่าเชื่อถือเพิ่มมากขึ้น ก็ควรเป็นการวิเคราะห์จากข้อมูลที่มีอยู่อย่างเพียงพอไม่น้อยเกินไปจนไม่สามารถวิเคราะห์อะไรได้ ดังนั้นในงานวิจัยนี้จึงได้ใช้วิธีการมาร์คอฟเชนมอนติคาร์โลมาเป็นวิธีการหลักในการวิเคราะห์ข้อมูลในด้านต่าง ๆ ดังที่ได้เสนอไปแล้ว

ในส่วนของการทดลองในบทที่ 3 ซึ่งเป็นการจำลองการศึกษาการส่งผลกระทบของปริมาณความเข้มข้นของกลูโคสและปริมาณความหนาแน่นเซลล์ที่มีต่อระดับการแสดงออกของยีน

ซึ่งสามารถสร้างแบบจำลองแสดงพฤติกรรมของการแสดงออกของยีนยีสต์ในระหว่างกระบวนการได้อ็อกซิซิฟท์ ที่ทำการออกแบบมาจากการศึกษาและวิเคราะห์พฤติกรรมการแสดงออกของยีนยีสต์ที่มีการรายงานไว้ในทางชีววิทยา และได้ทำการสร้างข้อมูลจำลองของพารามิเตอร์ต่างๆ ที่ส่งผลกระทบต่อการแสดงออกของยีนด้วยวิธีการกิบป์ แซมพลิง จนสามารถทำการอนุมานการแจกแจงโพลทีเรียของข้อมูลที่เราสงใจออกมาได้ ซึ่งจะเห็นได้ว่าการแจกแจงโพลทีเรียของการแสดงออกโดยเฉลี่ยของยีนนั้นเป็นแบบเอ็กซ์โพเนนเชียล และด้วยแบบจำลองนี้สามารถทำให้เราแยกแยะความแตกต่างของรูปแบบลักษณะการแสดงออกของยีนที่มีลักษณะที่แตกต่างกันได้ โดยอาศัยการพิจารณาจากค่าจำลองของซูเปอร์พารามิเตอร์ในแบบจำลอง แต่ตามธรรมชาติของสิ่งมีชีวิตโดยทั่วไปยังมีปัจจัยอื่น ๆ ที่ส่งผลกระทบต่อแสดงออกของยีนอีกหลายปัจจัย ดังนั้นควรมีการศึกษาเพิ่มเติมเพื่อให้ได้แบบจำลองที่มีความถูกต้องเพิ่มมากขึ้น

นอกจากนี้จากการวิเคราะห์ข้อมูลไมโครอาร์เรย์ที่เราเห็นกันอยู่ทั่วไป มักจะเป็นการวิเคราะห์ความสัมพันธ์ของระดับการแสดงออกของยีนด้วยการหาค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation) ของยีนในกลุ่มเดียวกันหรือระหว่างยีนที่เราสนใจ ดังนั้นจากการทดลองในบทที่ 3 นี้ จึงถือได้ว่าเป็นการเสนอวิธีการในการวิเคราะห์ข้อมูลไมโครอาร์เรย์ในอีกรูปแบบหนึ่ง ที่ทำการวิเคราะห์และสร้างเป็นแบบจำลองเชิงกราฟ โดยการศึกษาลักษณะความสัมพันธ์ระหว่างการแสดงออกของยีนที่มีรูปแบบลักษณะการแสดงออกที่แตกต่างกัน กับปริมาณความเข้มข้นของกลูโคสและปริมาณความหนาแน่นเซลล์ว่าปัจจัยทั้งสองนี้ส่งผลกระทบต่อแสดงออกของยีนด้วยปริมาณที่แตกต่างกัน

สำหรับการทดลองในบทที่ 4 ซึ่งอาจกล่าวได้ว่าเป็นหัวใจของงานวิจัยในครั้งนี้ โดยเป็นการสร้างแบบจำลองการควบคุมกันระหว่างยีน HAP2 HAP3 HAP4 และ CYC1 ซึ่งได้มีการรายงานไว้แล้วในทางชีววิทยาว่ายีนทั้งสี่นี้มีความสัมพันธ์กันโดยกลุ่มโปรตีน HAP2 HAP3 HAP4 ต่างไปควบคุมการแสดงออกของยีน CYC1 โดยเป็นการทดลองเพื่อสร้างแบบจำลองที่เป็นไปได้ของยีนทั้งสี่นี้ด้วยวิธีการของอัลกอริทึมเมโทรโพลิส - แฮสติงส์ ซึ่งเปรียบการทดลองในครั้งนี้ได้ว่าเป็นการสร้างแบบจำลองของสิ่งที่รู้อยู่แล้วขึ้นมาใหม่ เพื่อเป็นการยืนยันว่า ด้วยการทำงานของวิธีการดังกล่าวสามารถทำการสร้างและพิจารณายอมรับแบบจำลองที่ได้สร้างขึ้นมาได้อย่างถูกต้องอย่างน้อยเพียงใด ซึ่งจากผลการทดลองในบทที่ 4 ก็แสดงให้เห็นแล้วว่าด้วยวิธีการของอัลกอริทึมเมโทรโพลิส - แฮสติงส์ สามารถทำการสร้างแบบจำลองที่มีความถูกต้องและน่าเชื่อถือในระดับหนึ่ง แต่ก็ยังพบว่าบางแบบจำลองที่สร้างและถูกยอมรับมานั้นยังคงมีลักษณะที่แตกต่างจากแบบจำลองจริงที่มีการรายงานไว้ และเมื่อนำผลการทดลองที่ได้ไปเปรียบเทียบกับงานวิจัยที่ได้

ทำการทดลองไว้ก่อนหน้า (Manorat *et al.*, 2005) ก็พบว่าได้ผลการทดลองเป็นแบบจำลองที่มีลักษณะใกล้เคียงกัน

## 5.2 ข้อเสนอแนะและแนวทางการวิจัยในอนาคต

จากที่ได้กล่าวมาแล้วว่าการเข้าใจในกระบวนการควบคุมกันระหว่างยีนของเซลล์นั้นมีความสำคัญอย่างมาก ซึ่งอาจนำไปสู่การเข้าใจกระบวนการทำงานของเซลล์ที่มีความซับซ้อนได้ ซึ่งก็ยังเป็นแนวทางที่นักชีววิทยาให้ความสนใจ และมีความพยายามที่จะทำความเข้าใจกลไกเหล่านี้เพื่อให้สามารถก้าวไปสู่การจำลองพฤติกรรมของเซลล์ที่มีความซับซ้อนที่ประกอบด้วยปัจจัยต่าง ๆ มากมายเพื่อให้เซลล์สามารถดำรงชีวิตอยู่ได้ ดังนั้นแนวทางในการทำงานวิจัยในด้านนี้จึงยังได้รับความสนใจอยู่ นอกจากนี้ก้าวต่อไปของการศึกษาทางด้านนี้ มักจะต้องการให้สามารถทำความเข้าใจสิ่งมีชีวิตได้ทั้งระบบ (System Biology) ซึ่งการศึกษาจากส่วนที่เล็ก ๆ ส่วนหนึ่งก็อาจเป็นจุดเริ่มต้นนำไปสู่การศึกษาระบบที่มีขนาดใหญ่ขึ้นได้

จากการจำลองการศึกษาการส่งผลกระทบของปริมาณความเข้มข้นของกลูโคสและความหนาแน่นเซลล์ที่มีต่อระดับการแสดงออกของยีน อาจสามารถนำไปพัฒนาต่อเพื่อประยุกต์ใช้กับการสร้างแบบจำลองของเซลล์ ที่ต้องมีปัจจัยต่าง ๆ ของเซลล์เข้ามาเกี่ยวข้องด้วยสัดส่วนที่แตกต่างกันไป ซึ่งปัจจัยต่าง ๆ เหล่านี้ก็เปรียบเสมือนเป็นพารามิเตอร์ต่าง ๆ ที่อยู่ในแบบจำลอง หรือพัฒนาต่อเพื่อทำการสร้างแบบจำลองการควบคุมกันระหว่างยีน

สำหรับงานวิจัยในส่วนถัดมา ซึ่งเป็นการสร้าง (Reconstruction) เครือข่ายการควบคุมกันระหว่างกลุ่มยีนที่ได้มีการรายงานไว้แล้วทางชีววิทยา แต่เนื่องจากการทดลองครั้งนี้ได้ทำการทดลองกับข้อมูลเพียงชุดเดียว ฉะนั้นเพื่อให้เกิดความน่าเชื่อถือมากยิ่งขึ้นควรได้มีการนำขั้นตอนวิธีการทดลองในครั้งนี้ไปใช้ในการสร้างแบบจำลองเครือข่ายการควบคุมกันระหว่างกลุ่มยีนอื่น ๆ อีก ก่อนที่จะนำไปใช้สำหรับสร้างเครือข่ายระหว่างยีนที่ยังไม่ทราบความสัมพันธ์ระหว่างกัน แต่เนื่องจากว่าในงานวิจัยส่วนนี้เป็นการศึกษาจากข้อมูลการแสดงออกของยีนที่ผ่านขั้นตอนการเปลี่ยนแปลงจากข้อมูลชนิดต่อเนื่องให้เป็นข้อมูลชนิดไม่ต่อเนื่องก่อน โดยเงื่อนไขที่ใช้ในการแปลงนี้เป็นการสังเกตจากค่าการแสดงออกของยีน โดยตรงซึ่งกำหนดเป็นค่าคงที่ในช่วงข้อมูลหนึ่งที่สามารถจำแนกความแตกต่างของการแสดงออกของยีนได้เป็นการแสดงออกที่เพิ่มขึ้น ลดลง และคงที่ ซึ่งหากมีการเปลี่ยนแปลงเงื่อนไขไปใช้ค่าคงที่อื่น ๆ แทน หรือมีการนำวิธีการในการแปลงข้อมูลจากข้อมูลชนิดต่อเนื่องให้เป็นข้อมูลชนิดไม่ต่อเนื่องเข้ามาใช้แทนวิธีการสังเกตจากข้อมูลโดยตรง อาจจะมีผลทำให้ได้ผลลัพธ์ที่มีความแตกต่างออกไปจากผลการทดลองในการวิจัยครั้งนี้ นอกจากนี้อาจต้องมีการนำวิธีการอื่นของมาร์คอฟเชนมอนติคาร์โล อาทิเช่น วิธี Reversible jump

MCMC เข้ามาช่วยในการสร้างแบบจำลองการควบคุมกันระหว่างยีน และอาจนำข้อมูลลำดับของดีเอ็นเอเข้ามาวิเคราะห์ร่วมด้วย เพื่อให้ได้ผลลัพธ์ที่มีความน่าเชื่อถือมากยิ่งขึ้น ซึ่งจะเป็นการพัฒนาต่อไปในอนาคต



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่  
Copyright© by Chiang Mai University  
All rights reserved