บทที่ 3

การสร้างแบบจำลองเพื่อศึกษาผลกระทบของปริมาณความเข้มข้นกลูโคสและความหนาแน่นเซลล์ต่อระดับการแสดงออกของยีนในยีสต์สายพันธุ์ *แซคคาโรไมซิส เซริวิซิเย*

จากงานวิจัย Kirimasthong *et al.* (2006) ซึ่งได้ทำการจำลองความสัมพันธ์ระหว่างปริมาณความเข้มข้นของกลูโคสและระดับการแสดงออกยีนในยีสต์สายพันธุ์แซคคาโรไมซิส เซริวิซิเย ในระหว่างกระบวนการไดอ๊อคซิคชิฟท์ (Diauxic Shift) ออกมาเป็นแบบจำลองเชิงกราฟ โดยมีสมมติฐานว่าความสัมพันธ์ระหว่างกลูโคสและการแสดงออกของยีนมีลักษณะเป็นโมเดลถดถอยแบบเอ็กโพเนนเชียล (exponential regression model) และได้ทำการทดลองโดยใช้วิธีการมาร์คอฟเชนมอนติคาร์โล (Markov Chain Monte Carlo) ในการสร้างข้อมูลจำลอง (simulated data) ขึ้นมา และสามารถอนุมานจากข้อมูลจำลองนั้นเพื่อหาการแจกแจงโพสทีเรียและทำการประมาณค่าของพารามิเตอร์ต่าง ๆ ที่ส่งผลกระทบต่อการแสดงออกของยีนได้ ดังนั้นในงานวิจัยนี้จึงได้ทำการศึกษาและพัฒนาต่อยอดจากงานวิจัยข้างต้น แต่ยังคงใช้ข้อมูลที่ทำการศึกษาเป็นข้อมูลชุดเดียวกัน คือ กลุ่มข้อมูลการแสดงออกของยีนในยีสต์แซคคาโรไมซิส เซริวิซิเย ในระหว่างกระบวนการไดอ๊อคซิคชิฟท์ทั้งหมด 5 กลุ่ม ที่ถูกแบ่งตามลักษณะการแสดงออกของยีนเพื่อทำการศึกษาความสัมพันธ์ระหว่างปัจจัยต่าง ๆ ที่ส่งผลกระทบต่อการแสดงออกของยีนในเชิงลึกมากขึ้น

ในการศึกษาครั้งนี้ ได้ทำการศึกษาข้อมูลการแสดงออกของยีนของยีสต์แซคคาโรไมซิส เซริวิซิเย ในระหว่างกระบวนการไดอ๊อคซิคชิฟท์ ถึงแม้ว่ายีสต์จะเป็นสิ่งมีชีวิตเล็ก ๆ แต่ก็มีความสำคัญอย่างมาก นั่นคือ ยีสต์ได้ถูกนำไปใช้ในกระบวนการผลิตขนมปังหรือเครื่องดื่มประเภทแอลกอฮอล์ (Niederberger, 1989) นอกจากนี้ยังถูกนำไปใช้ในกระบวนการผลิตทางเทคโนโลยีชีวภาพ (Biotechnology) และอุตสาหกรรมเกี่ยวกับเภสัชกรรม (Pharmaceutical industry) (Aristidou and Penttila, 2000) สำหรับการเพาะเชื้อยีสต์นั้น ยีสต์จะถูกเพาะเชื้อไว้ในกลูโคส โดยเมื่อเวลาผ่านไปความเข้มข้นของกลูโคสจะค่อย ๆ ลดลงจนหมดไปแล้วยีสต์ก็จะเปลี่ยนไปใช้เอทานอลแทนเพื่อให้เซลล์สามารถมีชีวิตอยู่ได้ ซึ่งกระบวนการเปลี่ยนแปลงนี้จะเรียกว่า กระบวนการไดอ๊อคซิคชิฟท์

Derisi *et al.* (1997) ได้ทำการเก็บข้อมูลการแสดงออกของยีนในลักษณะดีเอ็นเอไมโครอาร์เรย์ของยีสต์ในระหว่างกระบวนการไดอ๊อคซิคชิฟท์ ทั้งหมด 7 การทดลองที่ระยะเวลาต่างกัน และได้แสดงระดับการแสดงออกของยีนต่าง ๆ ที่มีการแสดงออกที่เพิ่มขึ้นหรือลดลง เมื่อปริมาณความเข้มข้นของกลูโคสลดลงและความหนาแน่นเซลล์เพิ่มขึ้น

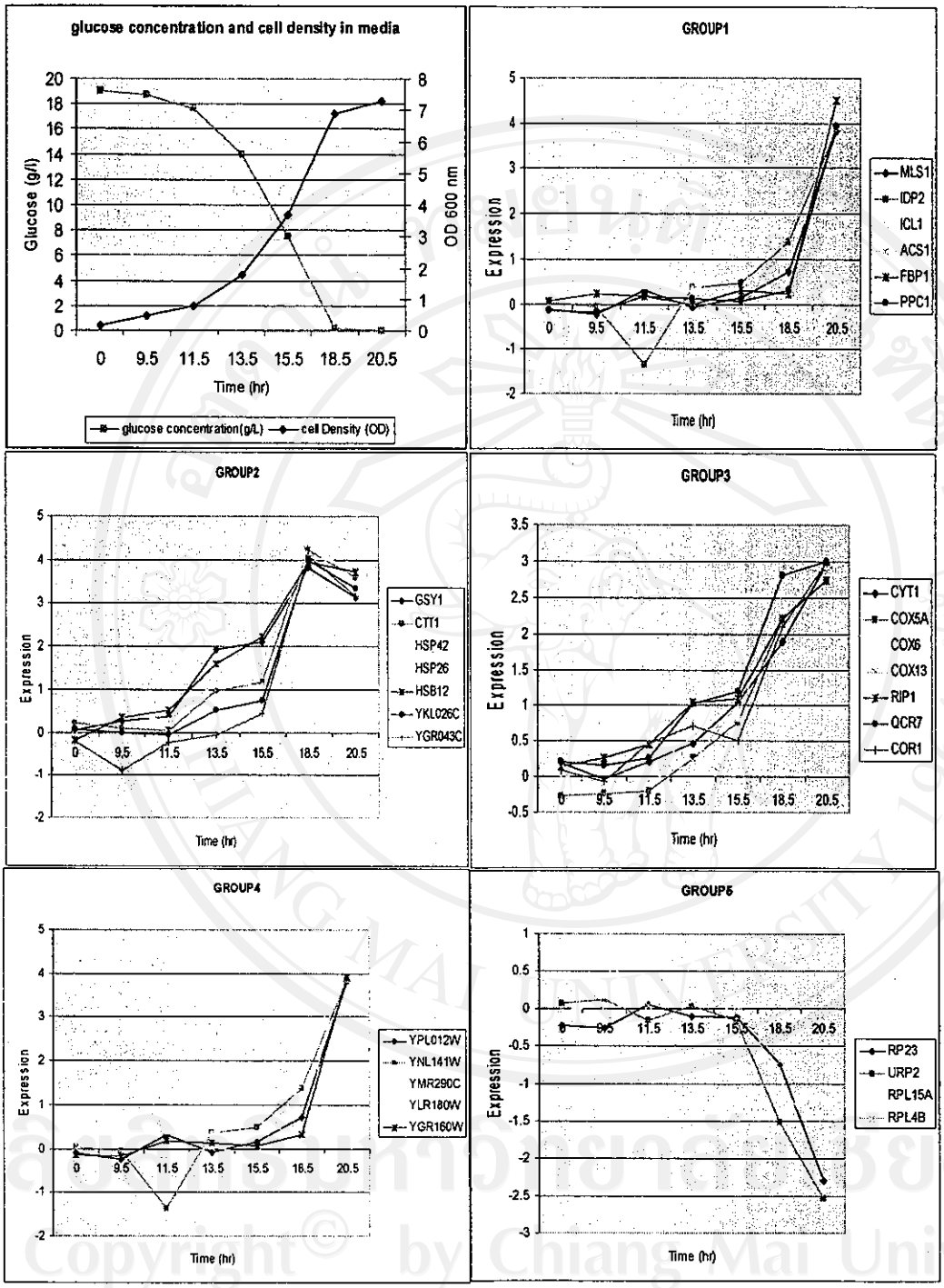
ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

การศึกษาครั้งนี้เป็นผลจากการศึกษาพฤติกรรมการแสดงออกของยีนของยีสต์สายพันธุ์แซคคาโรไมซิส เซริวิซิเย ในระหว่างกระบวนการไดอ๊อคซิคชิฟท์ โดยมีสมมติฐานว่าการแสดงออกของยีนน่าจะมีปัจจัยต่างๆ เข้ามาเกี่ยวข้อง ซึ่งปัจจัยที่เห็นได้ชัดว่ามีผลต่อการแสดงออกของยีนในระหว่างกระบวนการไดอ๊อคซิคชิฟท์นี้ คือ กลูโคสและความหนาแน่นเซลล์ ดังนั้นการศึกษานี้จึงมีวัตถุประสงค์เพื่อสร้างแบบจำลองเชิงกราฟที่เป็นตัวแทนแสดงการมีอิทธิพลของปริมาณความเข้มข้นของกลูโคสและปริมาณความหนาแน่นของเซลล์ ที่ส่งผลกระทบต่อระดับการแสดงออกของยีนที่ระยะเวลาต่างกันภายใต้กระบวนการไดอ๊อคซิคชิฟท์ โดยปริมาณความเข้มข้นของกลูโคส ความหนาแน่นของเซลล์ และระดับการแสดงออกของยีนมีความสัมพันธ์กันในลักษณะของโมเดลถดถอยแบบเอ็กโพเนนเชียลชนิดหลายตัวแปร (Multiple exponential regression model) ซึ่งในงานวิจัยนี้จะใช้วิธีการอนุมานด้วยเบย์เซียนโดยอาศัยมาร์คอฟเชนมอนติคาร์โลในการอนุมานหาการแจกแจงโพสทีเรียของโมเดลถดถอยที่ได้ทำการเสนอไว้ ซึ่งเป็นโมเดลที่ทำให้สามารถอนุมานระดับการแสดงออกของยีนออกมาได้ และค่าของพารามิเตอร์ต่าง ๆ ที่มีส่วนเกี่ยวข้องกับการแสดงออกของยีนในแบบจำลองเชิงกราฟจะถูกประมาณค่าด้วยโปรแกรม WinBUGS (Bayesian inference Using Gibbs Sampling for Windows) ซึ่งเป็นโปรแกรมที่อาศัยวิธีการกิบบ์ แซมพลิง ในการคำนวณ

3.1 ชุดข้อมูลการทดลอง

งานวิจัยนี้ได้ทำการศึกษาข้อมูลไมโครอาร์เรย์ของยีสแซคคาโรไมซิส เซริวิซิเย ในระหว่างกระบวนการไดอ๊อคซิคชิฟท์จาก DeRisi *et al.* (1997) ซึ่งสามารถดาวน์โหลดชุดข้อมูลไมโครอาร์เรย์นี้ได้จาก http://rana.lbl.gov/EisenData.htm สำหรับในการเพาะเลี้ยงเชื้อยีสต์นั้น ยีสต์จะถูกเพาะเชื้อไว้ในกลูโคสและใช้กลูโคสในการเจริญเติบโต ทำให้ปริมาณความเข้มข้นของกลูโคสจะค่อย ๆ ลดลงเมื่อเวลาผ่านไป โดยอัตราการเจริญเติบโตของเซลล์ยีสต์พบว่าเป็นแบบเอ็กโพเนนเชียล (Exponential growth) ซึ่งยีสต์จะทำการเปลี่ยนกลูโคสไปเป็นเอทานอล ซึ่งเป็นกระบวนการที่เรารู้จักกันดีที่เรียกว่า กระบวนการหมัก (Fermentation process) หลังจากที่ปริมาณความเข้มข้นของกลูโคสหมดลงเซลล์ยีสต์จะเปลี่ยนไปใช้เอทานอลแทนการใช้กลูโคส โดยการชิฟท์ (shift) จากกระบวนการหมักโดยใช้ออกซิเจนไปสู่การหายใจโดยใช้ออกซิเจน ซึ่งกระบวนการนี้จะถูกเรียกว่า ไดอ๊อคซิคชิฟท์ (Wikipedia, 2006)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved



รูป 3.1 รูปแบบลักษณะการแสดงออกของทั้งห้ากลุ่มยีนที่เกี่ยวข้องกับกระบวนการไดอ๊อคซิคชิฟท์

ข้อมูลชุดนี้เป็นข้อมูลไมโครอาร์เรย์ของยีนประมาณ 6,400 ยีน ที่ทำการวัดระดับการแสดงออกของยีนที่เวลาต่าง ๆ กันทั้งหมด 7 ช่วงเวลาในระหว่างกระบวนการไดอ๊อคซิคชิฟท์ ซึ่งจะถูกแบ่งกลุ่มออกเป็น 5 กลุ่มตามลักษณะการแสดงออกที่แตกต่างกัน (Derisi *et al.*, 1997)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

แสดงไว้ดังรูป 3.1 ยีนที่อยู่ในกลุ่มเดียวกันถือว่ามีหน้าที่คล้ายคลึงกัน ซึ่งยีนทั้งห้ากลุ่มนี้จะเป็นข้อมูลนำเข้า (input data) ของแบบจำลองเชิงกราฟเพื่อทำการทดลองต่อไป โดยยีนทั้งห้ากลุ่มนี้เป็นยีนที่เกี่ยวข้องกับกระบวนการทางชีววิทยาดังต่อไปนี้ 1) TCA และ glyoxalate cycles; 2) glycogen และ trehalose synthesis; 3) cytochrome – c; 4) SAM1 และ unknown function และ 5) ribosomal protein

(Derisi *et al.*, 1997)

3.2 วิธีการกิบบ์ แซมพลิง

วิธีการกิบบ์ แซมพลิง (Gilks *et al.*, 1996) เป็นวิธีการหนึ่งของมาร์คอฟเชนมอนติคาร์โล ซึ่งจะถูกใช้เป็นวิธีการในการสร้างข้อมูลจำลองขึ้นมา วิธีการนี้มักจะถูกใช้งานร่วมกับวิธีการอนุมานด้วยเบย์เซียนเพื่อทำการอนุมานหาการแจกแจงโพสทีเรียของกลุ่มข้อมูลที่เราสนใจ โดยวิธีการทำงานของกิบบ์ แซมพลิงสามารถอธิบายเป็นขั้นตอนหลัก ๆ ได้ดังต่อไปนี้

1. การออกแบบ (Designing) โครงสร้างของแบบจำลองเชิงกราฟเพื่อเป็นตัวแทนการแสดงออกของยีนในแต่ละกลุ่มยีน
2. การกำหนดค่าเริ่มต้น (Initializing) ให้กับทุกพารามิเตอร์ที่ไม่ทราบค่าและกำหนดการแจกแจงเบื้องต้น (Prior distribution) ให้กับพารามิเตอร์เหล่านั้น
3. การสุ่มข้อมูล (Sampling) ของพารามิเตอร์ต่าง ๆ แบบสุ่มจากการแจกแจงที่มีเงื่อนไข (Conditional distribution) โดยวิธีการของกิบบ์ แซมพลิงนี้จะทำงานแบบทำซ้ำเพื่อทำการสร้างข้อมูลจากการแจกแจงที่มีเงื่อนไขสำหรับโหนดหรือพารามิเตอร์ที่ไม่ทราบค่าทุกโหนดที่อยู่ในแบบจำลองเชิงกราฟ แสดงการแจกแจงที่มีเงื่อนไขได้ดังสมการที่ (13)

$$P(v \mid V_{-v}) \propto P(v, V_{-v})$$

$$= P(v \mid parents[v]) * \prod_{w \in children[w]} P(w \mid parents[w]) \qquad (12)$$

โดยที่ v เป็นโหนดหรือพารามิเตอร์ที่อยู่ในแบบจำลองเชิงกราฟ ส่วน V_{-v} เป็นโหนดหรือพารามิเตอร์อื่น ๆ ที่เหลือที่ไม่ใช่ v และ w เป็นโหนดหรือพารามิเตอร์ลูกของ v ซึ่งจากการแจกแจงที่มีเงื่อนไขของโหนดใด ๆ ในสมการที่ (12) ที่ประกอบด้วยเทอมไพรเออร์ ที่พิจารณาเฉพาะโหนด v และโหนดพ่อแม่ของโหนด v : $P(v \mid parents[v])$ และเทอมไลค์ลิฮูดที่พิจารณาจากโหนดลูกของโหนด v และโหนดพ่อแม่ร่วมของโหนดลูกเหล่านั้น ดังนั้นจึงเห็นได้ว่าการแจกแจงที่มีเงื่อนไขสำหรับแต่ละโหนดจะ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ขึ้นอยู่กับค่าของโหนดพ่อแม่ โหนดลูก และโหนดพ่อแม่ร่วมของโหนดลูกของมันเท่านั้น

4. การติดตามผลลัพธ์ (Monitoring) ที่ได้จากการจำลองข้อมูลว่ามีลักษณะการเปลี่ยนแปลงเป็นเช่นไร และข้อมูลจำลองในช่วงใดที่มีค่ายังไม่คงที่ (burn - in) ซึ่งจะไม่ถูกนำมาใช้ในการประมาณค่าของพารามิเตอร์ต่าง ๆ
5. การอนุมานจากผลลัพธ์ที่ได้ (Inferring) เพื่อประมาณค่าของพารามิเตอร์ที่ไม่ทราบค่า

3.3 แบบจำลองเชิงกราฟเพื่อเป็นตัวแทนระดับการแสดงออกของยีน

เรามีสมมติฐานว่า ความสัมพันธ์ระหว่างระดับการแสดงออกของยีนในแต่ละกลุ่มกับปริมาณความเข้มข้นของกลูโคสและความหนาแน่นเซลล์ เป็นโมเดลถดถอยแบบเอ็กโพเนนเชียลชนิดหลายตัวแปร (Multiple exponential regression model) ซึ่งเป็นสมมติฐานที่มาจากการสังเกตรูปแบบลักษณะการแสดงออกของยีนต่าง ๆ เปรียบเทียบกับปริมาณความเข้มข้นของกลูโคสและปริมาณความหนาแน่นของเซลล์ ซึ่งแสดงไว้ดังรูป 3.1 โดยพบว่าระดับการแสดงออกของยีนจะเพิ่มขึ้นหรือลดลงเป็นแบบเอ็กโพเนนเชียล เมื่อความเข้มข้นของกลูโคสลดลงในขณะที่ความหนาแน่นเซลล์เพิ่มขึ้น ซึ่งสมมติฐานนี้ได้นำไปสู่การออกแบบจำลองเชิงกราฟแสดงดังรูป 3.2 ที่เป็นโครงสร้างที่เป็นอิสระแบบมีเงื่อนไข (Conditional independence structure) เพื่อเป็นตัวแทนของระดับการแสดงออกของยีนในแต่ละกลุ่มยีน ที่สัมพันธ์กับปริมาณความเข้มข้นของกลูโคส และความหนาแน่นเซลล์แบบเอ็กโพเนนเชียลนี้ แบบจำลองนี้ถือได้ว่าเป็นเครื่องมือช่วยในการศึกษาผลกระทบของปัจจัยต่าง ๆ เช่น ปริมาณความเข้มข้นของกลูโคสและความหนาแน่นเซลล์ที่มีผลต่อระดับการแสดงออกของยีน นอกจากนี้แบบจำลองนี้ยังสามารถใช้เป็นเทคนิคในการประมาณค่าของพารามิเตอร์ต่าง ๆ สำหรับแต่ละกลุ่มยีนได้อีกด้วย

หลังจากทำการออกแบบแบบจำลองเชิงกราฟแล้ว จึงทำการกำหนดค่าเริ่มต้นและการแจกแจงไพรเออร์ให้กับทุกโหนดที่ไม่ทราบค่าที่อยู่ในแบบจำลองเชิงกราฟ โดยที่โหนดแต่ละโหนดจะเป็นตัวแทนของแต่ละพารามิเตอร์

แบบจำลองเชิงกราฟในรูป 3.2 ซึ่งถือว่าเป็นตัวแทนแสดงพฤติกรรมการแสดงออกของยีนและสามารถอธิบายแบบจำลองได้ดังต่อไปนี้

ขั้นแรกเราได้ทำการกำหนดการแจกแจงไพรเออร์ให้กับระดับการแสดงออกของยีนในแต่ละกลุ่มยีน ซึ่งแทนด้วยโหนด Y ด้วยการแจกแจงแบบปกติ (Normal distribution) เนื่องจากในทางชีววิทยามีสมมติฐานว่า ระดับการแสดงออกของยีนนั้นมีการแจกแจงเป็นแบบปกติ จากการแจกแจงไพรเออร์ของโหนด Y จะเห็นได้ว่าการแสดงออกของยีนจะขึ้นอยู่กับสองพารามิเตอร์ คือ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

พารามิเตอร์ค่าเฉลี่ย (mean parameter) *mu* และพารามิเตอร์ พรีซิชัน (precision parameter) *tau* โดยที่ tau = 1/variance หมายความว่า พารามิเตอร์ค่าเฉลี่ย *mu* จะส่งผลกระทบต่อระดับการแสดงออกของแต่ละยีนที่ระยะเวลาต่างๆ กัน และพารามิเตอร์พรีซิชัน *tau* จะส่งผลกระทบต่อความแปรปรวนของระดับการแสดงออกของยีนในแต่ละกลุ่มยีน ดังนั้นเมื่อค่าของพารามิเตอร์ *tau* ถูกประมาณค่าด้วยโปรแกรม WinbUGS แล้ว ค่าของพารามิเตอร์ *sigma* ซึ่งแทนส่วนเบี่ยงเบนมาตรฐานของยีนในแต่ละกลุ่มก็จะถูกประมาณค่าตามมาด้วยเช่นกัน

ในขั้นต่อไป จากสมมติฐานของเราที่ว่า ปริมาณความเข้มข้นของกลูโคสและความหนาแน่นเซลล์จะส่งผลกระทบแบบเอ็กโพเนนเชียลต่อระดับการแสดงออกของยีน ดังนั้นโมเดลถดถอยแบบเอ็กโพเนนเชียลชนิดหลายตัวแปรสำหรับพารามิเตอร์ *mu* ของระดับการแสดงออกของยีน ซึ่งแปรผกผันกับปริมาณความเข้มข้นของกลูโคส และแปรผันตรงกับปริมาณความหนาแน่นเซลล์ สามารถแสดงได้ดังสมการ

$$mu[i,j]=lamda_g[i]lamda_{od}[i]*\exp(-(lamda_g[i]glucose[j])+(lamda_{od}[i]OD[j])) \quad (13)$$

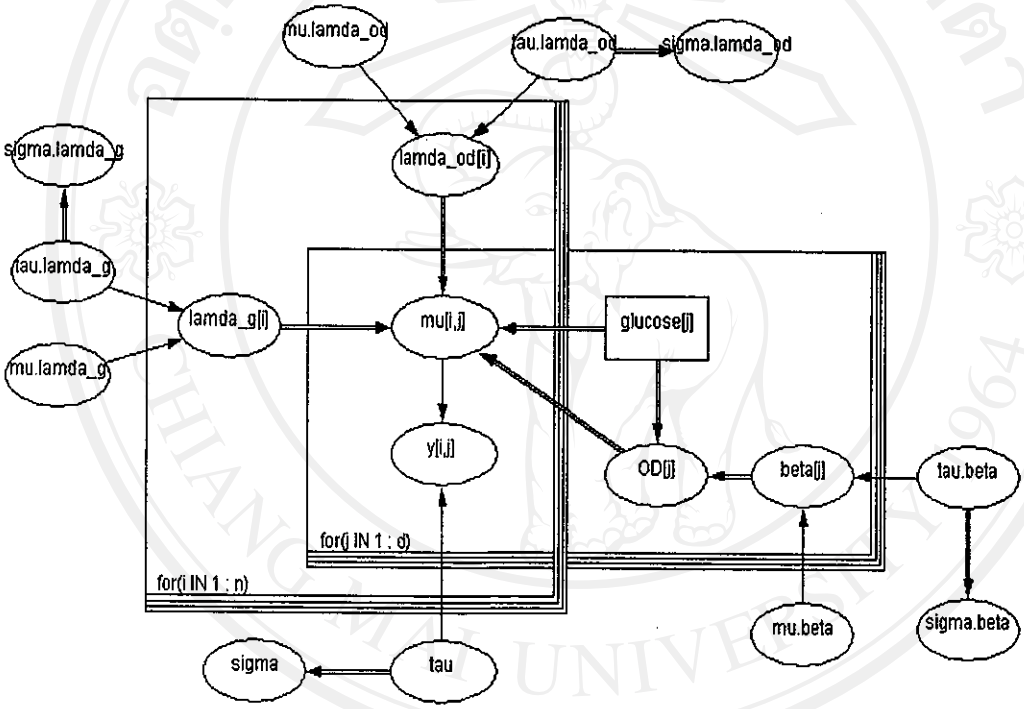
จากโมเดลถดถอยของ *mu*[i,j] ในสมการที่ (13) ซึ่งเป็นระดับการแสดงออกของยีนโดยเฉลี่ยของยีนที่ i^{th} ในช่วงระยะเวลาที่ j^{th} ซึ่งประกอบด้วยพารามิเตอร์ที่เราสนใจ คือพารามิเตอร์ *glucose* แทนปริมาณความเข้มข้นของกลูโคส ซึ่งจะถูกกำหนดด้วยค่าคงที่ที่ช่วงระยะเวลาต่าง ๆ กันในระหว่างกระบวนการไดอ๊อคซิคชิฟท์ และพารามิเตอร์ *OD* แทนปริมาณความหนาแน่นเซลล์ ที่ช่วงระยะเวลาต่าง ๆ กันในระหว่างกระบวนการไดอ๊อคซิคชิฟท์ โดยอัตราการเจริญเติบโตของเซลล์ยีสต์เป็นแบบเอ็กโพเนนเชียลซึ่งได้รับผลกระทบมาจากกลูโคส เนื่องจากเซลล์ยีสต์จะใช้กลูโคสในการเจริญเติบโตและดำรงชีวิตอยู่ ซึ่งพบว่าความหนาแน่นเซลล์จะแปรผกผันกับปริมาณความเข้มข้นของกลูโคส ดังนั้นค่าของพารามิเตอร์ *OD* ในสมการที่ (13) จึงคำนวณได้มาจากโมเดลถดถอยของปริมาณความเข้มข้นของกลูโคสที่ระยะเวลาต่างกัน แสดงได้ดังสมการ

$$OD[j]=beta[j]*\exp(-(beta[j]glucose[j])) \quad (14)$$

จากสมการที่ (13) จะเห็นได้ว่าพารามิเตอร์ *mu* ยังขึ้นอยู่กับพารามิเตอร์ของสมการเอ็กโพเนนเชียล นั่นคือ พารามิเตอร์ $lamda_g$ (พารามิเตอร์ของกลูโคส) และพารามิเตอร์ $lamda_{od}$ (พารามิเตอร์ของความหนาแน่นเซลล์) ในขณะที่ค่าของพารามิเตอร์ *OD* ในสมการที่ (14) จะขึ้นอยู่กับพารามิเตอร์ *beta* ซึ่งเราได้กำหนดการแจกแจงไพรเออร์ให้กับพารามิเตอร์ทั้งสามนี้ด้วยการแจกแจงแบบปกติที่ขึ้นอยู่กับซุปเปอร์พารามิเตอร์ (super - parameters) ของตัวเอง นั่นคือ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

พารามิเตอร์ *mu, tau* และ *sigma* สำหรับแต่ละพารามิเตอร์ โดยการแจกแจงไพรเออร์ของซุปเปอร์พารามิเตอร์ $mu.lamda_g$, $mu.lamda_{od}$ และ $mu.beta$ จะถูกกำหนดด้วยการแจกแจงแบบปกติ ที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 นั่นคือ dnorm(0,1) และกำหนดการแจกแจงไพร-เออร์ของซุปเปอร์พารามิเตอร์ $tau.lamda_g$, $tau.lamda_{od}$ และ $tau.beta$ ด้วยการแจกแจงแกมมา (Gamma distribution) คือ dgamma(0.001,0.001) และเมื่อค่าของพารามิเตอร์ $tau.lamda_g$, $tau.lamda_{od}$ and $tau.beta$ ถูกประมาณค่าด้วยโปรแกรม WinBUGS แล้ว ค่าของพารามิเตอร์ส่วนเบี่ยงเบนมาตรฐาน $sigma.lamda_g$, $sigma.lamda_{od}$ *and* $sigma.beta$ ก็จะถูกประมาณค่าด้วยเช่นเดียวกัน



รูป 3.2 แบบจำลองเชิงกราฟแสดงพฤติกรรมการแสดงออกของยีน โดยแสดงโครงสร้างที่เป็นอิสระแบบมีเงื่อนไขของระดับการแสดงออกของยีนในแต่ละกลุ่ม (โหนด Y) ที่ขึ้นอยู่กับกลูโคสและความหนาแน่นเซลล์ผ่านทางพารามิเตอร์ *mu* (ค่าเฉลี่ยของระดับการแสดงออกของยีน) โดยแต่ละโหนดจะเป็นตัวแทนของพารามิเตอร์ต่าง ๆ ในแบบจำลอง ซึ่งขึ้นอยู่กับโหนดพ่อแม่ โหนดลูก และโหนดพ่อแม่ร่วม ของมันเท่านั้น

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

3.4 อัลกอริทึมของการทดลอง

จากขั้นตอนและวิธีการในการทำงานของวิธีการกิบบ์ แซมพลิง โดยอาศัยการแจกแจงที่มีเงื่อนไขในการคำนวณที่ได้กล่าวมาแล้วข้างต้น สามารถสรุปเป็นอัลกอริทึมในการทำงานได้ดังต่อไปนี้

Input: y: Gene expression data of each group

Output: Simulated data of each parameter

$t = 0$

Initialize $mu^{(0)}, tau^{(0)}, sigma^{(0)}, OD^{(0)}, lamda_g^{(0)}, lamda_{od}^{(0)}, beta^{(0)},$
$mu.lamda_g^{(0)}, tau.lamda_g^{(0)}, sigma.lamda_g^{(0)},$
$mu.lamda_{od}^{(0)}, tau.lamda_{od}^{(0)}, sigma.lamda_{od}^{(0)},$
$mu.beta^{(0)}, tau.bata^{(0)}, sigma.beta^{(0)}$

do{

$$mu^{(t+1)} = P(mu^{(t)} \mid V_{-mu^{(t)}}) \sim P(mu^{(t)} \mid lamda_g^{(t)}, lamda_{od}^{(t)}, OD^{(t)}, glucose) P(y \mid mu^{(t)}, tau^{(t)})$$

$$OD^{(t+1)} = P(OD^{(t)} \mid V_{-OD^{(t)}}) \sim P(OD^{(t)} \mid beta^{(t)}, glucose) * P(mu^{(t)} \mid lamda_g^{(t)}, lamda_{od}^{(t)}, OD^{(t)}, glucose)$$

$$tau^{(t+1)} = P(tau^{(t)} \mid V_{-tau^{(t)}}) \sim P(tau^{(t)}) P(y \mid mu^{(t)}, tau^{(t)}) P(sigma^{(t)} \mid tau^{(t)})$$

$$sigma^{(t+1)} = P(sigma^{(t)} \mid V_{-sigma^{(t)}}) \sim P(sigma^{(t)} \mid tau^{(t)})$$

$$lamda_g^{(t+1)} = P(lamda_g^{(t)} \mid V_{-lamda_g^{(t)}}) \sim P(lamda_g^{(t)} \mid mu.lamda_g^{(t)}, tau.lamda_g^{(t)}) * P(mu^{(t)} \mid lamda_g^{(t)}, lamda_{od}^{(t)}, OD^{(t)}, glucose)$$

$$lamda_{od}^{(t+1)} = P(lamda_{od}^{(t)} \mid V_{-lamda_{od}^{(t)}}) \sim P(lamda_{od}^{(t)} \mid mu.lamda_{od}^{(t)}, tau.lamda_{od}^{(t)}) * P(mu^{(t)} \mid lamda_g^{(t)}, lamda_{od}^{(t)}, OD^{(t)}, glucose)$$

$$beta^{(t)} = P(beta^{(t)} \mid V_{-beta^{(t)}}) \sim P(beta^{(t)} \mid mu.beta^{(t)}, tau.beta^{(t)}) * P(OD^{(t)} \mid beta^{(t)}, glocose)$$

$$mu.lamda_g^{(t+1)} = P(mu.lamda_g^{(t)} \mid V_{-mu.lamda_g^{(t)}}) \sim P(mu.lamda_g^{(t)}) * P(lamda_g^{(t)} \mid mu.lamda_g^{(t)}, tau.lamda_g^{(t)})$$

$$tau.lamda_g^{(t+1)} = P(tau.lamda_g^{(t)} \mid V_{-tau.lamda_g^{(t)}}) \sim P(tau.lamda_g^{(t)}) * P(lamda_g^{(t)} \mid mu.lamda_g^{(t)}, tau.lamda_g^{(t)})$$

$$sigma.lamda_g^{(t+1)} = P(sigma.lamda_g^{(t)} \mid V_{-sigma.lamda_g^{(t)}}) \sim P(sigma.lamda_g^{(t)} \mid tau.lamda_g^{(t)})$$

$$mu.lamda_{od}^{(t+1)} = P(mu.lamda_{od}^{(t)} \mid V_{-mu.lamda_{od}^{(t)}}) \sim P(mu.lamda_{od}^{(t)}) * P(lamda_{od}^{(t)} \mid mu.lamda_{od}^{(t)}, tau.lamda_{od}^{(t)})$$

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

$$tau.lamda_{od}^{(t+1)} = P(tau.lamda_{od}^{(t)} \mid V_{-tau.lamda_{od}^{(t)}}) \sim P(tau.lamda_{od}^{(t)}) * P(lamda_{od}^{(t)} \mid mu.lamda_{od}^{(t)}, tau.lamda_{od}^{(t)})$$

$$sigma.lamda_{od}^{(t+1)} = P(sigma.lamda_{od}^{(t)} \mid V_{-sigma.lamda_{od}^{(t)}}) \sim P(sigma.lamda_{od}^{(t)} \mid tau.lamda_{od}^{(t)})$$

$$mu.beta^{(t+1)} = P(mu.beta^{(t)} \mid V_{-mu.beta^{(t)}}) \sim P(mu.beta^{(t)}) * P(beta^{(t)} \mid mu.bata^{(t)}, tau.bata^{(t)})$$

$$tau.beta^{(t+1)} = P(tau.beta^{(t)} \mid V_{-tau.beta^{(t)}}) \sim P(tau.beta^{(t)}) * P(beta^{(t)} \mid mu.bata^{(t)}, tau.bata^{(t)})$$

$$sigma.beta^{(t+1)} = P(sigma.beta^{(t)} \mid V_{-sigma.beta^{(t)}}) \sim P(sigma.beta^{(t)} \mid tau.beta^{(t)})$$

$$t = t + 1$$

}until (Parameters converge)

3.5 การทดลอง

สำหรับการวิเคราะห์ข้อมูลในงานครั้งนี้ เริ่มต้นจากทำการศึกษาพฤติกรรมการแสดงออกของยีนยีสต์สายพันธุ์แซคคาโรไมซิส เซริวิซิเย ในระหว่างกระบวนการไดอ๊อคซิดชิฟท์ จนพบว่ามีปัจจัยที่เข้ามาเกี่ยวข้องกับการแสดงออกของยีนยีสต์นี้สองปัจจัยด้วยกัน คือ ปริมาณความเข้มข้นกลูโคสและความหนาแน่นเซลล์ จากการศึกษาพฤติกรรมของยีนดังกล่าวจึงทำให้เกิดสมมติฐานที่ว่าปริมาณความเข้มข้นกลูโคสและความหนาแน่นเซลล์น่าจะส่งผลกระทบต่อการแสดงออกของยีนยีสต์ในระหว่างกระบวนการไดอ๊อคซิดชิฟท์เป็นโมเดลถดถอยแบบเอ็กโพเนนเชียลชนิดหลายตัวแปร ซึ่งสมมุติฐานนี้จะถูกนำไปใช้ในการออกแบบจำลองเชิงกราฟเพื่อเป็นตัวแทนแสดงการแสดงออกของยีนต่างๆ ทั้งห้ากลุ่มยีนต่อไป หลังจากทำการออกแบบแบบจำลองเชิงกราฟ และกำหนดค่าเริ่มต้นและการแจกแจงไพรเออร์ให้กับทุกโหนดที่ไม่ทราบค่าแล้ว เราจะใช้โปรแกรม WinBUGS ซึ่งสามารถดาวน์โหลดได้จาก http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml ในการสร้างข้อมูลจำลองของแต่ละพารามิเตอร์ที่อยู่ในแบบจำลองขึ้นมา เพื่อใช้ในการอนุมานการแจกแจงโพสทีเรียและประมาณค่าของพารามิเตอร์ต่างๆ ซึ่งโปรแกรม WinBUGS นี้ จะใช้วิธีการอนุมานด้วยเบย์เซียน (Bayesian inference) และวิธีการกิบบ์แซมพลิง (Gibb sampling) เป็นวิธีการในการคำนวณและประมาณค่าของพารามิเตอร์ต่าง ๆ

แบบจำลองความสัมพันธ์แบบเอ็กโพเนนเชียลนี้จะถูกใช้ในการสร้างข้อมูลจำลองใหม่ของแต่ละพารามิเตอร์ขึ้นมา โดยทำการรันโปรแกรมแบบทำซ้ำทั้งหมด 5,000 รอบ ซึ่งข้อมูลที่ถูกสร้าง

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ขึ้นในช่วงแรก ๆ จะมีการเปลี่ยนแปลงมาก แล้วการเปลี่ยนแปลงจะค่อย ๆ ลดลงจนกระทั่งข้อมูลที่ได้มีการกระจายในรูปแบบคงที่ (Convergence) จากนั้นจึงทำการเลือกข้อมูลที่ถูกสร้างขึ้นมาที่อยู่ในช่วงที่มีการกระจายในรูปแบบที่คงที่แล้วออกมาทั้งหมด 1,000 ตัว โดยตัดข้อมูลที่อยู่ในช่วงที่ยังมีการกระจายที่เปลี่ยนแปลงอยู่มากหรือช่วงเบิร์น – อินออก ซึ่งข้อมูลที่มีการกระจายในรูปแบบที่คงที่ทั้ง 1,000 ตัวจะถูกนำไปใช้ในการอนุมานการแจกแจงโพสทีเรียของพารามิเตอร์ *mu* สำหรับแต่ละกลุ่มยีน และนำไปใช้ในการประมาณค่าของพารามิเตอร์ต่าง ๆ ที่มีผลกระทบต่อระดับการแสดงออกของยีน

จากนั้นเพื่อเป็นการตรวจสอบสมมติฐานที่ตั้งไว้ว่า ความสัมพันธ์ระหว่างระดับการแสดงออกของยีนในแต่ละกลุ่มกับปริมาณความเข้มข้นของกลูโคสและความหนาแน่นเซลล์ เป็นโมเดลถดถอยแบบเอ็กโพเนนเชียลชนิดหลายตัวแปร จึงได้ทำการออกแบบแบบจำลองเชิงกราฟซึ่งมีลักษณะเช่นเดียวกับแบบจำลองที่ได้ทำการออกแบบไว้ในรูป 3.2 แต่กำหนดสมการความสัมพันธ์ให้กับพารามิเตอร์ *mu* และพารามิเตอร์ *OD* ด้วยโมเดลถดถอยเชิงเส้น (Linear regression model) แสดงดังสมการที่ (15) และสมการที่ (16) ตามลำดับ

$$mu[i,j]=a_g[i]glucose[j]+a_{od}[i]OD[j]+a_0 \quad (15)$$

$$OD[j]=b_g[j]glucose[j]+b_0 \quad (16)$$

จากแบบจำลองซึ่งกำหนดสมการความสัมพันธ์แบบเชิงเส้นดังกล่าว จึงใช้โปรแกรม WinBUGS ทำการสร้างข้อมูลของพารามิเตอร์ต่างๆ ขึ้นมาเช่นเดียวกับการทดลองของแบบจำลองเชิงกราฟที่กำหนดสมการความสัมพันธ์แบบเอ็กโพเนนเชียล เพื่อทำการเปรียบเทียบผลลัพธ์ที่ได้จากการทดลองจากแบบจำลองทั้งสองกรณีนี้

3.6 ผลการทดลอง

จากการพิจารณาการแจกแจงของพารามิเตอร์ *mu* แสดงดังรูป 3.3 ซึ่งเป็นค่าจำลองของพารามิเตอร์ *mu* ที่เป็นผลลัพธ์จากโปรแกรม WinBUGS แสดงด้วยกราฟ caterpillar plot โดยเส้นสีดำแต่ละเส้นที่อยู่ในกราฟจะแสดงถึงช่วงค่าข้อมูลที่จำลองได้ของพารามิเตอร์ *mu* ของแต่ละยีนในระยะเวลาที่แตกต่างกัน และจุดทึบสีดำที่อยู่บนเส้นตรงจะแสดงถึงค่าเฉลี่ยของค่าจำลองของพารามิเตอร์ *mu* ตัวอย่างเช่น เส้นสีดำที่ตำแหน่ง [1, 7] ในยีนกลุ่มที่ 1 มีความหมายว่า ค่าจำลองที่สร้างขึ้นมาได้ของพารามิเตอร์ *mu* ของยีนตัวที่ 1 คือยีน MLS1 ที่ระยะเวลาในช่วงที่ 7 ของกระบวนการไดอีอคซิคชิฟท์ พบว่ามีค่าอยู่ในช่วงสามถึงห้า และมีค่าเฉลี่ยของพารามิเตอร์ *mu* ในช่วงข้อมูลที่ถูกจำลองขึ้นมานี้ประมาณเท่ากับสี่ เป็นต้น

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

จากรูป 3.3 ซึ่งเป็นผลที่ได้จากการทดลองพบว่า ค่าที่จำลองได้ของพารามิเตอร์ *mu* ซึ่งเป็นพารามิเตอร์แสดงค่าเฉลี่ยของการแสดงออกของยีนในแต่ละกลุ่มยีนมีค่าใกล้เคียงกับค่าจริงของระดับการแสดงออกของยีนที่แสดงไว้ในรูป 3.1 และเมื่อพิจารณาเฉพาะค่าเฉลี่ยของพารามิเตอร์ *mu* ของข้อมูลจำลองในแต่ละช่วงของทั้งห้ากลุ่มยีนแสดงดังรูป 3.4 พบว่าลักษณะรูปแบบการแสดงออกของยีนในแต่ละกลุ่มยีนที่ได้จากการจำลองในครั้งนี้ ยังคงมีลักษณะคล้ายคลึงกับรูปแบบการแสดงออกของยีนที่เป็นชุดข้อมูลตั้งต้นในรูป 3.1 โดยกราฟของยีนในกลุ่มที่ 1 ถึงกลุ่มที่ 4 พบว่าลักษณะรูปแบบของค่าเฉลี่ยของพารามิเตอร์จะค่อยข้างคงที่ในช่วงระยะเวลาที่หนึ่งถึงห้า และเพิ่มขึ้นอย่างรวดเร็วในช่วงระยะเวลาที่หกและเจ็ด ซึ่งมีลักษณะตรงข้ามกับยีนในกลุ่มที่ 5 คือ ในช่วงระยะเวลาที่หกและเจ็ด รูปแบบลักษณะของค่าเฉลี่ยของพารามิเตอร์ *mu* จะมีค่าลดลงอย่างรวดเร็ว นั้นแสดงให้เห็นได้ว่าการแจกแจงโพสทีเรียของพารามิเตอร์ *mu* นี้เป็นการแจกแจงแบบเอ็กโพเนนเชียล

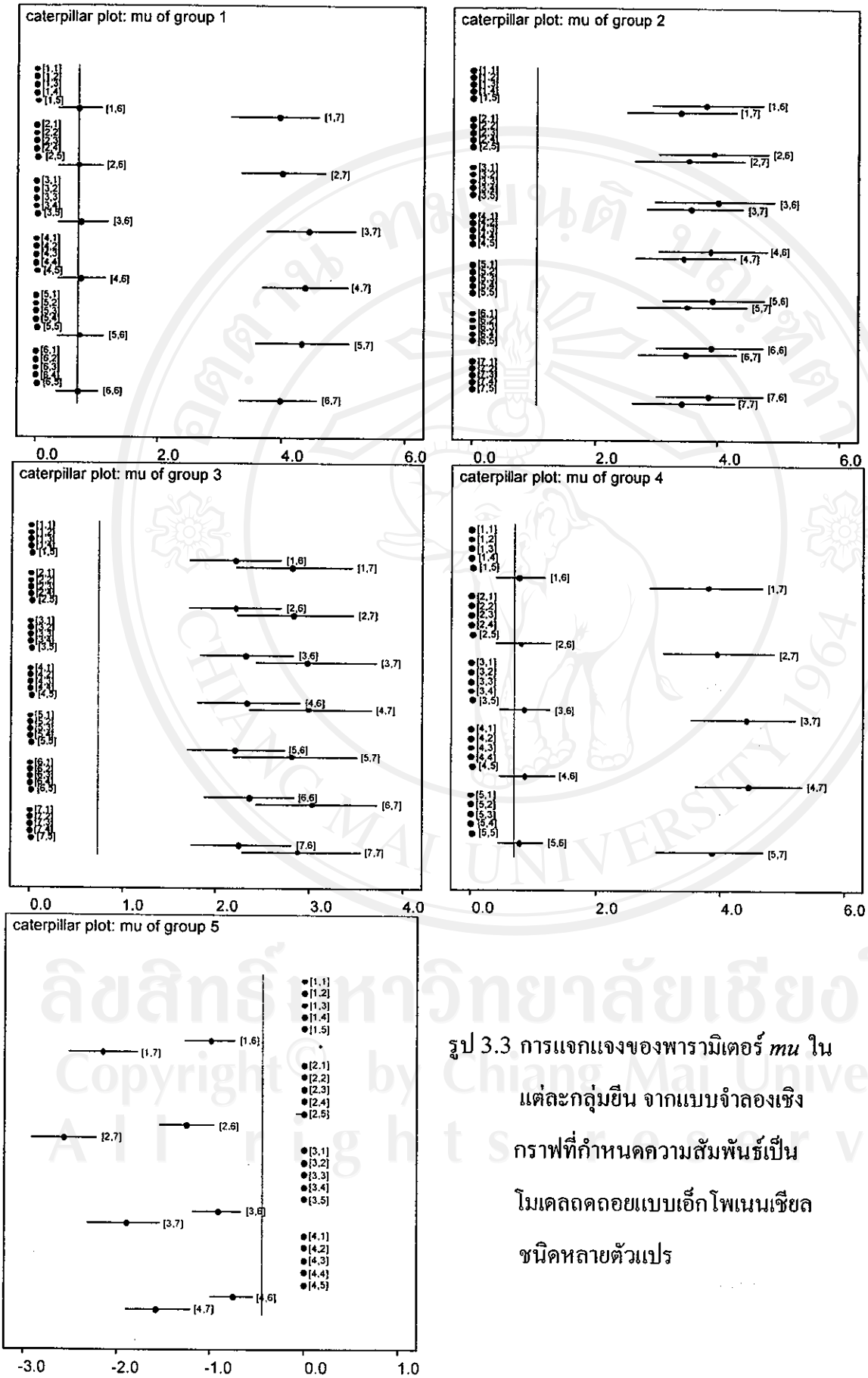
จากแบบจำลองเชิงกราฟที่ได้ทำการออกแบบไว้ พบว่ายังมีพารามิเตอร์ที่มีความสำคัญที่มีส่วนส่งผลกระทบต่อระดับการแสดงออกของยีน นั่นคือ พารามิเตอร์ $lamda_g$ (พารามิเตอร์ของกลูโคส) พารามิเตอร์ $lamda_{od}$ (พารามิเตอร์ของความหนาแน่นเซลล์) ซึ่งทั้งสองพารามิเตอร์นี้เป็นพารามิเตอร์ของโมเดลถดถอยแบบเอ็กโพเนนเชียลชนิดหลายตัวแปรของพารามิเตอร์ *mu* และพารามิเตอร์ที่มีความสำคัญอีกพารามิเตอร์หนึ่ง คือ พารามิเตอร์ *sigma* (พารามิเตอร์ส่วนเบี่ยงเบนมาตรฐานของแต่ละกลุ่มยีน) สามารถแสดงค่าเฉลี่ยของพารามิเตอร์ทั้งสามนี้ ซึ่งเป็นค่าโดยเฉลี่ยของค่าจำลองที่สร้างขึ้นมาได้ทั้งหมดในแต่ละกลุ่มยีน ดังตาราง 3.1

ตาราง 3.1 ค่าเฉลี่ยของพารามิเตอร์ $lamda_g$, $lamda_{od}$ และ *sigma* ที่มีส่วนกระทบต่อระดับการแสดงออกของยีนในแต่ละกลุ่มยีนโดยผ่านทางพารามิเตอร์ *mu*

GROUP	PARAMETER				
	$mu.lamda_g$	$Mu.lamda_{od}$	$Sigma.lamda_g$	$Sigma.lamda_{od}$	sigma
group1	0.5016	1.873	0.05096	0.07212	0.4314
group2	1.022	2.23	0.08211	0.09852	0.808
group3	0.6608	1.3	0.06279	0.05874	0.5094
group4	0.307	0.8934	0.05553	0.05062	0.4628
group5	-0.5278	2.099	0.1479	0.2425	0.193

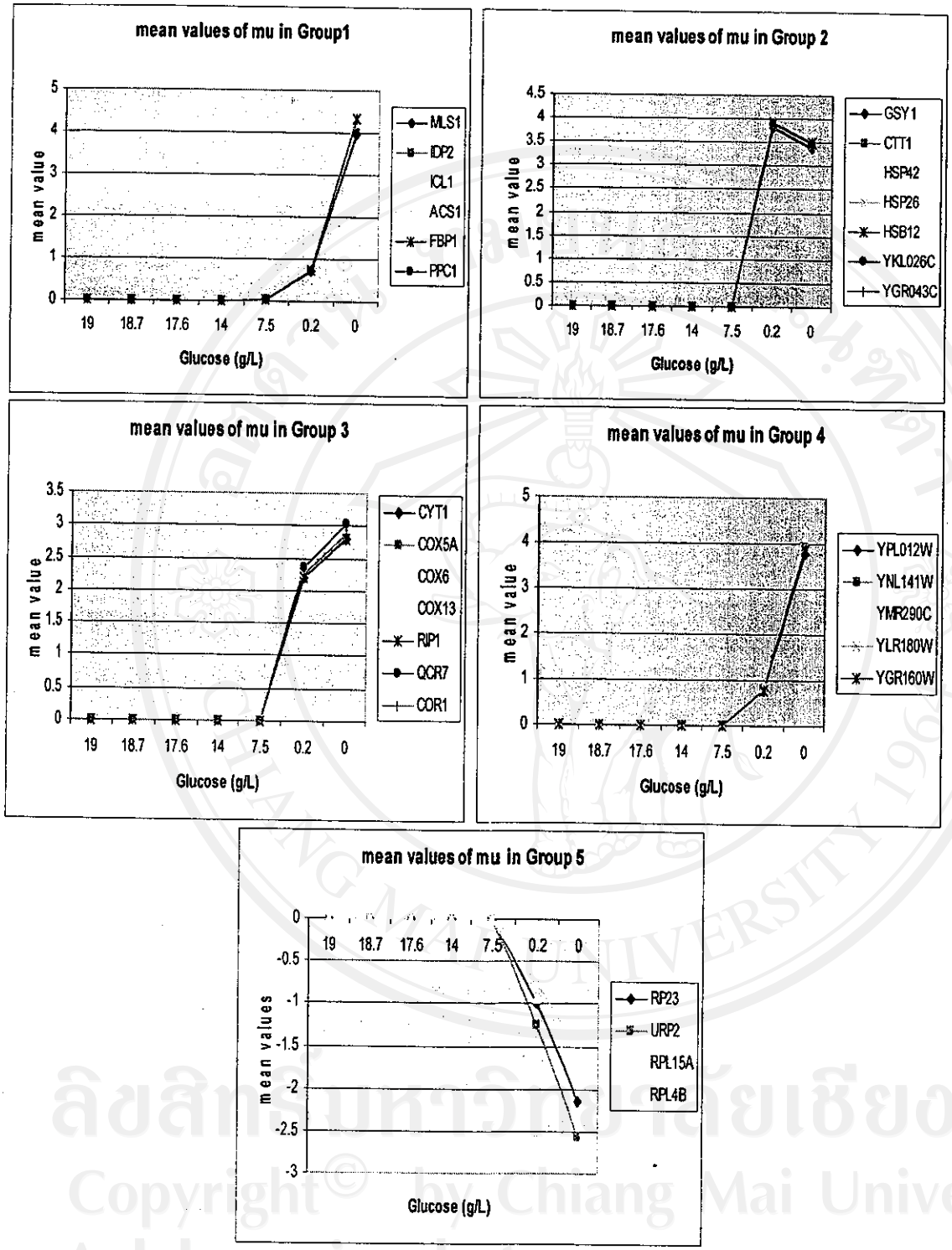
จากตาราง 3.1 พบว่าค่าของพารามิเตอร์ *sigma* ในแต่ละกลุ่มยีนมีค่าน้อย ที่เป็นเช่นนี้อาจเนื่องมาจากข้อมูลที่ได้จากการทดลองนี้มีความคล้ายคลึงกับระดับการแสดงออกของยีนมาก และชุดข้อมูลยีนที่เป็นข้อมูลตั้งต้นในการทดลองนี้ได้ถูกแบ่งกลุ่มมาแล้ว อาจทำให้ข้อมูลยีนในแต่ละกลุ่มยีนมีความแตกต่างกันไม่มาก สำหรับค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของพารามิเตอร์

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved



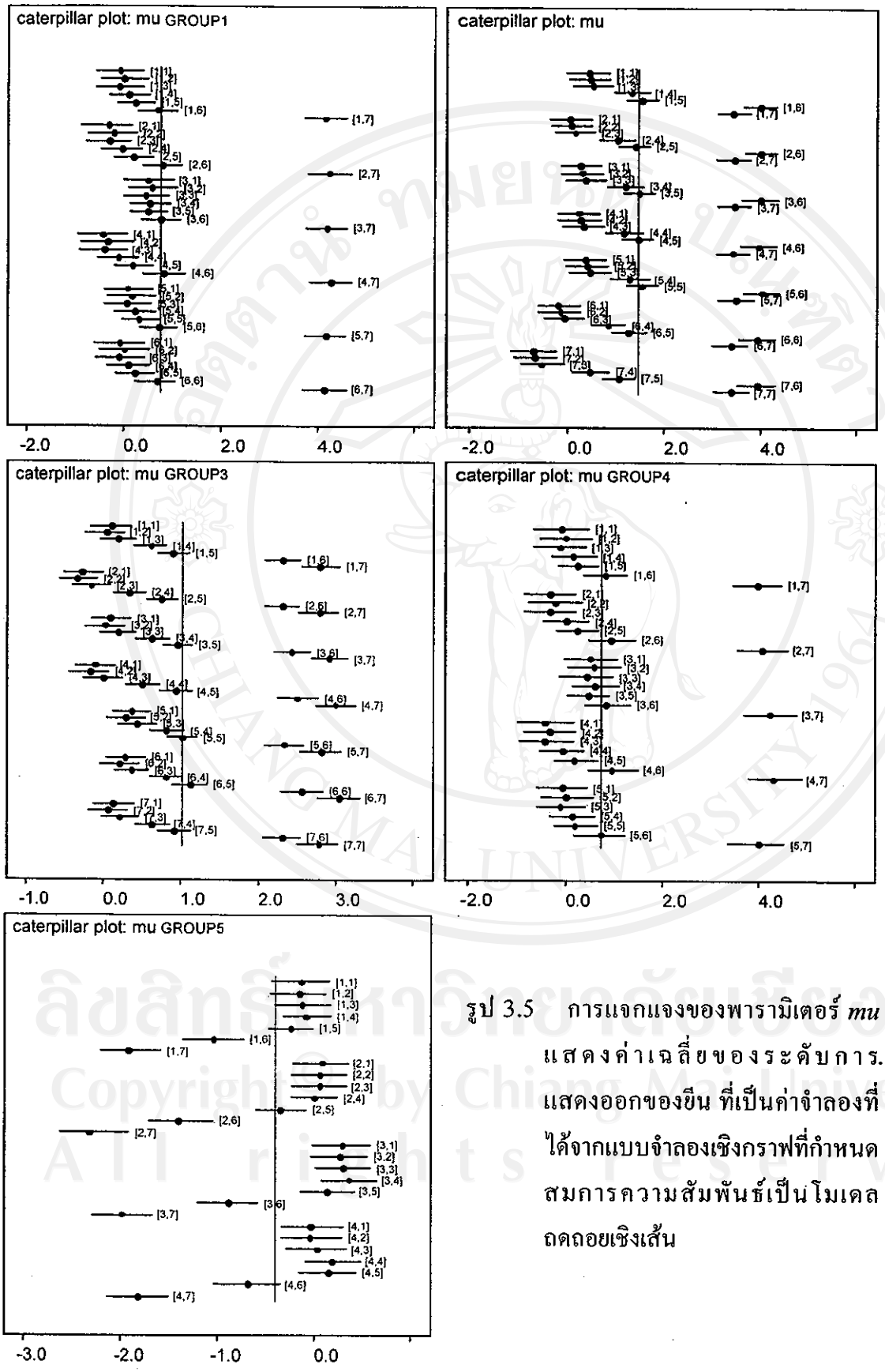
รูป 3.3 การแจกแจงของพารามิเตอร์ *mu* ใน แต่ละกลุ่มยีน จากแบบจำลองเชิงกราฟที่กำหนดความสัมพันธ์เป็นโมเดลถดถอยแบบเอ็กโพเนนเชียล ชนิดหลายตัวแปร

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved



รูป 3.4 ค่าเฉลี่ยของค่าจำลองของพารามิเตอร์ *mu* เพื่อแสดงรูปแบบการแจกแจงของพารามิเตอร์ *mu* สำหรับแต่ละกลุ่มยีนที่ได้มาจากการจำลองที่กำหนดความสัมพันธ์เป็นโมเดลถดถอยแบบเอ็กโพเนนเชียลชนิดหลายตัวแปร ซึ่งพบว่ารูปแบบลักษณะของแต่ละยีนยังคงมีลักษณะคล้ายคลึงกับค่าการแสดงออกจริงของยีนต่าง ๆ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved



รูป 3.5 การแจกแจงของพารามิเตอร์ *mu* แสดงค่าเฉลี่ยของระดับการแสดงออกของยีน ที่เป็นค่าจำลองที่ได้จากแบบจำลองเชิงกราฟที่กำหนดสมการความสัมพันธ์เป็นโมเดลถดถอยเชิงเส้น

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

$lamda_g$ และ $lamda_{od}$ ซึ่งก็คือค่าของพารามิเตอร์ $mu.lamda_g$ $mu.lamda_{od}$ และ $sigma.lamda_g$ $sigma.lamda_{od}$ ตามลำดับ โดยค่าเฉลี่ยของค่าจำลองของพารามิเตอร์เหล่านี้แสดงไว้ในตาราง 3.1 ได้บ่งบอกถึงปริมาณการส่งผลกระทบของกลูโคสและความหนาแน่นเซลล์ที่มีต่อระดับการแสดงออกของยีนในแต่ละกลุ่มที่ได้จากการจำลองในครั้งนี้ พบว่าค่าของพารามิเตอร์ $mu.lamda_g$ $sigma.lamda_g$ และ $sigma.lamda_{od}$ (ยกเว้นค่าของพารามิเตอร์ $mu.lamda_{od}$ ที่มีค่าแตกต่างกันไปในแต่ละกลุ่มยีน) ในกลุ่มยีนที่ 1 ถึงกลุ่มยีนที่ 4 มีค่าใกล้เคียงกัน ซึ่งมีค่าแตกต่างจากกลุ่มที่ 5 อย่างเห็นได้ชัด ที่เป็นเช่นนี้อาจเนื่องมาจากรูปแบบลักษณะการแสดออกของยีนในกลุ่มที่ 1 ถึงกลุ่มที่ 4 มีลักษณะคล้ายคลึงกัน คือมีการแสดงออกที่เพิ่มขึ้นอย่างรวดเร็วในช่วงเวลาท้ายๆ ซึ่งต่างจากรูปแบบลักษณะการแสดงออกของยีนในกลุ่มที่ 5 ที่ในช่วงเวลาท้ายของกระบวนการไดอ็อคซิดชิฟท์มีการแสดงออกที่ลดลงอย่างรวดเร็ว ดังนั้นค่าของพารามิเตอร์ที่ส่งผลกระทบต่อยีนที่มีลักษณะรูปแบบการแสดงออกที่ต่างกันจึงมีค่าที่แตกต่างกันตามไปด้วย

เมื่อพิจารณาผลลัพธ์ที่ได้จากแบบจำลองที่กำหนดสมการความสัมพันธ์ให้กับพารามิเตอร์ *mu* และ *OD* เป็นโมเดลถดถอยเชิงเส้นโดยพิจารณาที่ค่าจำลองของพารามิเตอร์ *mu* แสดงดังรูป 3.5 พบว่าค่าที่จำลองได้ของทั้งห้ากลุ่มยีนมีค่าใกล้เคียงกับระดับการแสดงออกของยีนจริงๆ ในรูป 3.1 และรูปแบบลักษณะการแสดงออกของแต่ละกลุ่มยีนก็มีลักษณะใกล้เคียงกับลักษณะการแสดงออกของยีนจริงที่เป็นข้อมูลตั้งต้นในรูป 3.1

จากการเปรียบเทียบผลลัพธ์ที่ได้จากแบบจำลองที่กำหนดด้วยสมการความสัมพันธ์ที่แตกต่างกันพบว่า ผลลัพธ์ของการสร้างข้อมูลจำลองของพารามิเตอร์ *mu* ที่ได้มีลักษณะคล้ายคลึงกันและใกล้เคียงกับการแสดงออกของยีนที่เป็นข้อมูลตั้งต้นที่เราสนใจนำมาศึกษา ดังนั้นจึงไม่สามารถสรุปตามสมมติฐานที่ได้ตั้งไว้ว่าปริมาณความเข้มข้นของกลูโคสและความหนาแน่นเซลล์ส่งผลกระทบต่อระดับการแสดงออกของยีนในระหว่างกระบวนการไดอ็อคซิดชิฟท์เป็นแบบเอ็กโพเนนเชียลได้

3.7 บทสรุป

จากงานวิจัยนี้ แสดงให้เห็นว่าการจำลองการศึกษาในครั้งนี้สามารถแสดงปริมาณการส่งผลกระทบของความเข้มข้นกลูโคสและความหนาแน่นเซลล์ที่มีต่อระดับการแสดงออกของยีนในระหว่างกระบวนการไดอ็อคซิดชิฟท์ได้โดยผ่านทางพารามิเตอร์ของกลูโคสและความหนาแน่นเซลล์ โดยอาศัยแบบจำลองที่เป็นโครงสร้างแบบมีเงื่อนไขเป็นเครื่องมือในการอนุมานหาการแจกแจงโพสทีเรียของระดับการแสดงออกโดยเฉลี่ยของยีน (พารามิเตอร์ *mu*) ด้วยการใช้วิธีการกิบบ์แซมพลิงและการอนุมานด้วยเบย์เซียน โดยไม่สามารถสรุปได้ว่าปริมาณความเข้มข้นของกลูโคส

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

และความหนาแน่นเซลล์ส่งผลกระทบต่อระดับการแสดงออกของยีนเป็นแบบเอ็กโพเนนเชียลตามสมมติฐานที่ตั้งไว้ได้ และจากการศึกษาในครั้งนี้ทำให้เราเข้าใจพฤติกรรมการแสดงออกของยีนของเซลล์ยีสต์ในระหว่างกระบวนการไดอ็อคซิคชิฟท์เพิ่มมากขึ้น นอกจากนี้เมื่อทำการเปรียบเทียบผลลัพธ์ที่ได้กับงานวิจัยก่อนหน้า (Kirimasthong *et al.*, 2006) ซึ่งเป็นการศึกษาการส่งผลกระทบของปริมาณความเข้มข้นของกลูโคสที่มีต่อการแสดงออกของยีนเพียงอย่างเดียวเท่านั้น พบว่าเมื่อนำปริมาณความหนาแน่นเซลล์เข้ามาพิจารณาร่วมด้วย ทำให้ได้ผลลัพธ์ของการแจกแจงของพารามิเตอร์ *mu* ที่มีลักษณะใกล้เคียงกับการแสดงออกจริงของยีนมากกว่าในงานวิจัยแรก

นอกจากนี้การศึกษาครั้งนี้ยังได้แสดงให้เห็นถึงวิธีการวิเคราะห์ข้อมูลดีเอ็นเอไมโครอาร์เรย์อีกรูปแบบหนึ่ง ที่ทำการวิเคราะห์และสร้างเป็นแบบจำลองเชิงกราฟ โดยการศึกษาลักษณะความสัมพันธ์ระหว่างการแสดงออกของยีนที่มีรูปแบบลักษณะการแสดงออกที่แตกต่างกัน กับปริมาณความเข้มข้นของกลูโคสและปริมาณความหนาแน่นเซลล์ว่าปัจจัยทั้งสองนี้ส่งผลกระทบต่อการแสดงออกของยีนด้วยปริมาณที่แตกต่างกันโดยพิจารณาที่ค่าจำลองของพารามิเตอร์ของกลูโคสและความหนาแน่นเซลล์

ดังนั้นควรมีการพัฒนาแบบจำลองนี้ต่อไปเพื่อให้แบบจำลองที่ได้ออกแบบไว้นี้ให้มีความถูกต้องและน่าเชื่อถือมากยิ่งขึ้นจนสามารถนำไปใช้กับข้อมูลทางชีววิทยากลุ่มอื่นๆ ได้ ซึ่งอาจต้องใช้ความรู้ทางด้านชีววิทยาเข้ามาช่วยในการพิจารณา และตามธรรมชาติของสิ่งมีชีวิตโดยทั่วไปยังอาจมีปัจจัยอื่น ๆ ที่ส่งผลต่อการแสดงออกของยีนอีกหลายปัจจัย

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved