

บทที่ 5

อภิปรายผลการทดลอง

แรงจูงใจและประโยชน์ในการศึกษา

จากการขยายตัวของฐานข้อมูลลำดับของกรดอะมิโนที่ได้จากการอ่านรหัสพันธุกรรม ทำให้การศึกษาโปรตีนจากลำดับของกรดอะมิโนได้ขยายตัว อย่างรวดเร็ว

วิธีการทดสอบโปรตีน 2 ชนิดว่าชอบโมโลกัสนกันหรือไม่ โดยพิจารณาจากลำดับของกรดอะมิโนมีหลายวิธี เช่น maximum matching (Needleman, 1969), Basic Local Alignment Search Tool (BLAST) (Samuel Karlin and Steven Altschul, 1990-1993) hidden Markov models (Kevin et al, 1999) เป็นต้น วิธีที่นิยมและใช้กันมากคือการทดสอบความคล้ายกันของลำดับของกรดอะมิโน แต่ปัญหาที่พบคือ ไม่สามารถทดสอบกับโปรตีนที่มี sequence identity ต่ำๆ หรือ โปรตีนที่มีช่องว่าง (gap) ขนาดใหญ่ หรือ จำนวนมากได้

งานวิจัยนี้ได้เปลี่ยนลำดับของกรดอะมิโนให้อยู่ในรูปของแผนภูมิ 2 มิติเพื่อทดสอบชอบโมโลกัสนกันของโปรตีน โดยมีสมมติฐานว่า โปรตีนที่ชอบโมโลกัสนกันจะให้แผนภูมิ 2 มิติของอะมิโนที่คล้ายกัน และวิธีการนี้จะลดปัญหาในเรื่องของการที่ลำดับของกรดอะมิโนมีที่นำมาทดสอบมี sequence identity ต่ำ หรือ ช่องว่างได้ และได้สร้างโครงข่ายประสาทเทียมขึ้นช่วยในการทดสอบความเหมือนกันของรูปแบบของแผนภูมิ 2 มิติของกรดอะมิโนสั้นๆ

อภิปรายผลการทดลอง

งานวิจัยชิ้นนี้ประกอบด้วย 2 ส่วนคือ

1. แผนภูมิโครงสร้าง 2 มิติ

จากลำดับของกรดอะมิโนที่มาจากโปรตีนจำนวน 30 คู่ซึ่งชอบโมโลกัสนกัน และมี sequence identity ต่างๆ กัน พบว่าโปรตีนที่ชอบโมโลกัสนกันจะให้แผนภูมิ 2 มิติที่มีลักษณะเหมือนกัน หรือคล้ายกัน โดยเฉพาะอย่างยิ่งบริเวณที่กรดอะมิโนชนิด hydrophobic มีการรวมกลุ่มเป็น cluster พบว่าลักษณะของ cluster จะคล้ายๆ กัน

วิธีการนี้แก้ปัญหาในเรื่องลำดับของกรดอะมิโนที่นำมาทดสอบมี sequence identity ต่ำ หรือ ช่องว่าง ได้เนื่องจากการพิจารณาความเหมือนหรือคล้ายกันของแผนภูมิ 2 มิติ นั้นจะดูจากลักษณะ cluster ของกรดอะมิโนที่มีสมบัติเป็น hydrophobic เป็นอันดับแรก แล้วจึงพิจารณากรดอะมิโนรอบข้าง cluster ว่ามีสมบัติเป็นอย่างไร ดังนั้นการที่โปรตีนมี sequence identity ต่ำก็ไม่มีผลในการพิจารณาความคล้ายหรือเหมือนของแผนภูมิ 2 มิติ เนื่องจากไม่ได้พิจารณาความเหมือนกันของกรดอะมิโนในลำดับโดยตรง และถ้าโปรตีนที่นำมาทดสอบมี ช่องว่าง ก็จะพบว่าลักษณะของ cluster ของกรดอะมิโนที่มีสมบัติเป็น hydrophobic ก็ยังคงมีลักษณะที่เหมือนกันหรือคล้ายกัน แต่อาจจะมีการเลื่อนตำแหน่งของ cluster หรือเปลี่ยนรูปร่างไปบ้างตามขนาดของช่องว่าง และบริเวณเกิดช่องว่างนั่นเอง

ในการพิจารณาฮอโมโลยีของโปรตีนโดยใช้แผนภูมิ 2 มิติของกรดอะมิโนเพียงอย่างเดียว นั้นพบว่าบางครั้งไม่เพียงพอที่จะบอกว่า โปรตีนทั้ง 2 ชนิดนั้นฮอโมโลจิสกันหรือไม่ เนื่องจากพบว่า มี โปรตีน บางชนิดไม่เป็นฮอโมโลจิสโปรตีนแต่แผนภูมิ 2 มิติของกรดอะมิโนที่ได้มีลักษณะเหมือนกันเช่นโปรตีนที่พัฒนามาเพื่อจะทำหน้าที่เดียวกัน แต่มาจากต้นกำเนิดที่ต่างกัน เป็นต้น ดังนั้นจึงต้องมีการพิจารณาลำดับทางวิวัฒนาการควบคู่กันไปด้วย เพื่อเพิ่มความถูกต้องยิ่งขึ้น

และข้อจำกัดที่สำคัญอย่างหนึ่งของการใช้แผนภูมิ 2 มิติของกรดอะมิโนคือ บางครั้งการพิจารณาความเหมือนหรือคล้ายกันของแผนภูมิ 2 มิติกรดอะมิโนนั้นเป็นเรื่องยาก ต้องใช้ความชำนาญในการพิจารณาเนื่องจากการพิจารณาความเหมือนหรือคล้ายกันนั้นไม่ได้พิจารณาเพียงแต่ลักษณะเท่านั้น บางครั้งต้องพิจารณาสมบัติของกรดอะมิโนด้วย

2. โครงข่ายประสาทเทียม

จากโครงข่ายประสาทเทียมที่สร้างขึ้นพบว่า เมื่อเพิ่มจำนวน hidden layer มากขึ้นผลการจำแนกรูปแบบของแผนภูมิ 2 มิติที่เกิดจากลำดับของกรดอะมิโนจำนวน 18 ตัวให้เปอร์เซ็นต์ความถูกต้องมากขึ้น โดยจำนวน hidden layers ที่เหมาะสมกับ test

set ชุดนี้คือ 60 และ 70 สาเหตุที่ทำให้เปอร์เซ็นต์ความถูกต้องมีค่าสูงเนื่องมาจาก test set มีขนาดเล็ก และไม่ต่างจาก Training set มากนัก

แต่โครงข่ายประสาทเทียมนี้มีประสิทธิภาพไม่มากนักในการนำไปใช้งานจริง เนื่องจากโครงข่ายประสาทเทียมนี้ผ่านการสอน(Train)จากข้อมูลจำนวนไม่มากและหลากหลายรูปแบบ ดังนั้นจึงควรมีการเพิ่ม training set เพื่อเป็นการเพิ่มประสิทธิภาพให้กับ โครงข่ายประสาทเทียม ให้สามารถนำไปใช้งานได้จริง

ปัญหาในการทดลอง

1. ปัญหาที่เกิดจากการแบ่งกลุ่มของกรดอะมิโนเพื่อใช้ในการสร้าง แผนภูมิ 2 มิติ เนื่องจากมีมাত্রามาก ถึง 40 มাত্রาที่ใช้ในการบอกค่า hydrophobicity ของกรดอะมิโนแต่ละตัว
2. ปัญหาในการเลือก training set และ test set เนื่องจากจำนวนของลำดับของกรดอะมิโนมีมากมาย จึงมีความยากลำบากในการเลือก ลำดับของกรดอะมิโนมาเป็น training set และ test set

แนวทางการวิจัยในอนาคต

1. ทำการเพิ่มจำนวน training set เพื่อเพิ่มประสิทธิภาพในการทดสอบความคล้ายกันของรูปแบบสั้นๆ
2. มีการทดสอบหาจำนวนของกรดอะมิโนในแต่ละรูปแบบสั้น ที่เหมาะสม
3. นำโครงข่ายประสาทเทียมที่ได้ ไปช่วยในการทดสอบการมีต้นกำเนิดเดียวกันของโปรตีนที่มีขนาดใหญ่ (ประกอบด้วยกรดอะมิโนจำนวนมาก)โดยใช้วิธี Gocal Alignment
4. ขยายการทดสอบแบบคู่ (pair) ให้ไปสู่ การทดสอบแบบ multiple alignment)
5. ทำการแบ่งกลุ่มของกรดอะมิโนก่อนที่จะนำไป train โครงข่ายประสาทเทียม จากนั้นให้ใช้กลุ่มแทนการเปลี่ยนกรดอะมิโนเป็นข้อมูลเชิงทวิภาคของกรดอะมิโนทั้ง 20 ตัว ก็จะลดความซับซ้อนของ input ได้