

สิริวรรณ แต้วิจิตร 2552: การจำแนกซัพแฟมิลีและวิเคราะห์โครงสร้างของเอนไซม์
ด้วยวิธีการพิจารณาการติดกันของกลุ่มกรดอะมิโนและโพรไฟล์ฮิดเดินมาร์คอฟโมเดล
ปริญาวิทยาศาสตรมหาบัณฑิต (วิศวกรรมคอมพิวเตอร์) สาขาวิศวกรรม
คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก:
รองศาสตราจารย์กฤษณะ ไวยมัย, Ph.D. 111 หน้า

การค้นพบเอนไซม์ใหม่มีอัตราการค้นพบที่สูงขึ้นอย่างรวดเร็ว เพื่อลดเวลาและค่าใช้จ่าย
ในการจำแนกประเภทเอนไซม์ จึงมีการพัฒนางานวิจัยเพื่อสร้างระบบอัตโนมัติด้วยวิธีการต่างๆ
ได้แก่ (1) วิธีการเปรียบเทียบเอนไซม์ใหม่และเอนไซม์ต้นแบบที่ทราบฟังก์ชันแล้ว วิธีดังกล่าวให้
ความถูกต้องสูงเมื่อเอนไซม์ใหม่มีค่าความเหมือนกับเอนไซม์ต้นแบบมากกว่า 60% (2) วิธีการ
พิจารณาส่วนย่อยของกลุ่มกรดอะมิโน ซึ่งสามารถบ่งบอกฟังก์ชันได้ (3) วิธีการสร้างฟีเจอร์เพื่อ
เป็นตัวแทนเอนไซม์ เช่น พิจารณาความถี่ในการเกิดกรดอะมิโน วิธีดังกล่าวให้ความถูกต้องได้ดี
แต่กระนั้นผลลัพธ์ที่ได้ไม่สามารถแสดงความสัมพันธ์เกี่ยวกับโครงสร้างของเอนไซม์ใหม่ได้

ในงานวิจัยนี้ทำการพัฒนาระบบจำแนกประเภทเอนไซม์ซัพแฟมิลีที่ให้ความถูกต้องสูง
และให้ผลลัพธ์ที่สามารถนำไปวิเคราะห์โครงสร้างเอนไซม์ใหม่ ในรูปแบบการติดกันของ
โครงสร้างทุติยภูมิ โดยมีสมมติฐานที่ว่าส่วนย่อยของโครงสร้างเอนไซม์สามารถบ่งบอกหน้าที่
การทำงานของเอนไซม์ได้ สำหรับส่วนย่อยของโครงสร้างนั้นสืบค้นจากตารางเมตริกซ์คอนแทค
แมป ซึ่งเป็นการแปลงข้อมูลโครงสร้าง 3 มิติให้อยู่ในรูปตารางเมตริกซ์ที่เข้าใจง่าย และนำแต่ละ
ส่วนย่อยมาประกอบกันเป็นตัวแทนสายโปรตีนด้วยโพรไฟล์ฮิดเดินมาร์คอฟโมเดล ดังนั้นข้อมูล
ที่ใช้ในการสกัดฟีเจอร์จำเป็นต้องทราบโครงสร้าง 3 มิติแล้วเท่านั้น คิดเป็นสัดส่วนเพียง 11.21%
ของข้อมูลทั้งหมดในงานวิจัยนี้ และสร้างแบบจำลองด้วยซัพพอร์ตเวกเตอร์แมชชีน จากผลการ
ทดลองพบว่าฟีเจอร์จากตารางเมตริกซ์คอนแทคแมป มีความสัมพันธ์และมีนัยสำคัญเพียงพอใน
การจำแนกประเภทเอนไซม์ โดยแบบจำลองให้ค่าความถูกต้องถึง 73.71% เมื่อทดสอบด้วยวิธี
jackknife และเมื่อเพิ่มองค์ความรู้เกี่ยวกับการแทนที่ระหว่างกรดอะมิโน สามารถเพิ่มความถูกต้อง
ได้ถึง 75.87% เมื่อทดสอบด้วยวิธี 10-fold cross validation