

บทที่ 1

บทนำ

ชีวสารสนเทศศาสตร์ (Bioinformatics) เป็นศาสตร์แขนงใหม่ซึ่งมีความเกี่ยวข้องกับการนำความรู้ด้านเทคโนโลยีสารสนเทศมาใช้ในการศึกษา และใช้เพื่อการจัดการกับข้อมูลจำนวนมากที่ได้จากการวิจัยด้านวิทยาศาสตร์ชีวภาพ (Biological Science) ทำให้ก่อให้เกิดองค์ความรู้ใหม่ๆ โดยเฉพาะอย่างยิ่งต้องงานวิจัยทางด้านวิทยาศาสตร์ชีวภาพ และงานทางด้านวิทยาศาสตร์การแพทย์ ซึ่งในปัจจุบันข้อมูลของยีนและโปรตีนมีจำนวนมากขึ้นอย่างรวดเร็ว ดังนั้นชีวสารสนเทศศาสตร์จึงแขนงวิชาที่กำลังมีบทบาทต่อวงการวิทยาศาสตร์ และเป็นเครื่องมือสำคัญในงานวิจัยยุคหลังจีโนม (Post-genomic era) อาจกล่าวได้ว่าแขนงวิชาชีวสารสนเทศศาสตร์ ถือกำเนิดขึ้นจากความจำเป็นที่นักวิทยาศาสตร์ต้องจัดให้มีระบบบริหารจัดการข้อมูลจำนวนมากมหาศาล ที่เป็นผลจากโครงการจีโนมของสิ่งมีชีวิตต่าง ๆ ความรู้ความเข้าใจและการประยุกต์ใช้ชีวสารสนเทศในงานวิจัยจะช่วยให้นักวิจัยสามารถสืบค้นข้อมูลยีนและโปรตีน ศึกษาความสัมพันธ์ของการทำงานของยีนทั้งระบบ (Genome wide) เข้าใจกลไกการควบคุมการแสดงออกของยีน รวมทั้งความสัมพันธ์ระหว่างโครงสร้างและหน้าที่ของโปรตีน

1.1 หลักการ ทฤษฎี เหตุผล และ สมมุติฐาน

การพัฒนาเทคโนโลยีด้านจีโนม เป็นเทคโนโลยีที่เกี่ยวข้องโดยตรงกับสารพันธุกรรม ซึ่งมีหน้าที่ควบคุมลักษณะต่างๆ ของสิ่งมีชีวิตทุกชนิด ดังนั้น การที่สามารถระบุตำแหน่งและหน้าที่ของยีนได้จะช่วยให้เราเข้าใจสิ่งมีชีวิตนั้นๆ มากยิ่งขึ้นซึ่งก่อให้เกิดประโยชน์มากมายมหาศาลเช่น การนำเทคโนโลยีชีวสารสนเทศ มาร่วมกับงานวิจัยวิทยาศาสตร์ชีวภาพและวิทยาศาสตร์การแพทย์ จะช่วยให้สามารถวิเคราะห์ข้อมูลจีโนมที่มีจำนวนมากขึ้น ชีวสารสนเทศจะช่วยให้สามารถศึกษาความสัมพันธ์ระหว่างโครงสร้าง และหน้าที่ของโปรตีนในภาพรวมระดับเซลล์ ตลอดจนเปิดประตูสู่การค้นหายาใหม่ๆ ด้วยกระบวนการค้นหามีประสิทธิภาพในยุคหลังจีโนม หรือมีฉะนั้นอย่างน้อยก็จะสามารถใช้ยีนนั้นผลิตโปรตีนที่ใช้ในการรักษาโรคได้หลายชนิด เช่น โรคมะเร็ง บาดแผลเรื้อรังที่รักษาไม่หาย เป็นต้น

ในประเทศไทย เทคโนโลยีชีวภาพได้เข้ามามีบทบาทในการพัฒนาเศรษฐกิจและสังคมของประเทศในทุกด้านรวมทั้งด้านการเกษตร โดยเฉพาะอย่างยิ่งข้าว ซึ่งเป็นธัญญาหารของคนไทยและ

ชาวโลกกว่า 4,000 ล้านคน ข้าวเป็นสินค้าที่ประเทศไทยส่งออกมากเป็นอันดับหนึ่งของโลก สร้างรายได้ให้กับประเทศปีละประมาณ 75,000 ล้านบาท ประเทศไทยยังเป็นแหล่งกำเนิดหลากหลายพันธุ์ข้าวที่สำคัญอีกแหล่งหนึ่ง ดังนั้นการวิจัยและพัฒนาด้านข้าว จึงเป็นพื้นฐานที่สำคัญ ที่จะช่วยยกระดับความสามารถและประสิทธิภาพในการผลิต และปรับปรุงพันธุ์ข้าวได้แก่ การค้นหายีนที่กำหนดลักษณะเด่นของพันธุ์ข้าวที่ดีไว้ เช่น เครื่องหมายดีเอ็นเอของยีนต้านโรคและแมลง ยีนทนต่อน้ำท่วม และยีนควบคุมคุณภาพของข้าว เป็นต้น

จากโครงการจีโนมต่าง ๆ (Genome sequencing projects) ทำให้มีข้อมูลลำดับเบสจำนวนมหาศาล ซึ่งได้เปิดทางให้นักวิทยาศาสตร์ได้ค้นหายีนทั้งหมดในสิ่งมีชีวิตต่าง ๆ อย่างกว้างขวาง ซึ่งระบบชีวสารสนเทศ โดยเฉพาะการทำนายยีน (Gene Prediction) มีความจำเป็นต่อการสืบค้นหายีนได้อย่างรวดเร็ว นอกจากนี้ยังสามารถนำมาประยุกต์ใช้ในการวิเคราะห์บทบาทหน้าที่ อธิบายโครงสร้าง ตำแหน่งที่สำคัญต่าง ๆ และการแสดงออกของยีนต่อไป

ในปัจจุบัน นักวิจัยทางด้านชีวสารสนเทศศาสตร์ให้ความสนใจในเรื่องของการทำนายยีนเพื่อระบุตำแหน่งและหน้าที่ของยีน รวมทั้งจะมีความพยายามในการหาโครงสร้างและหน้าที่ของโปรตีน ตลอดจนศึกษาการควบคุมการทำงานของโปรตีน จากจีโนมของสิ่งมีชีวิตต่าง ๆ รวมทั้งมนุษย์อย่างแพร่หลาย ทำให้มีโปรแกรมซึ่งเป็นผลจากงานวิจัย ที่มีความสามารถในการทำนายยีนออกมาเป็นจำนวนมาก ยกตัวอย่างเช่น โปรแกรมชื่อ Genscan (Chris Burge และคณะ, 1997), Genemark (Mark Borodovsky และคณะ, 1993), GeneParser (Snyder และคณะ, 1993), Genie (D. Kulp และคณะ, 1996) และ Glimmer (S. Salzberg และคณะ, 1998) เป็นต้น ซึ่งเมื่อมองในภาพรวมแล้ว ผลการทำนายที่ได้จากโปรแกรมต่าง ๆ ออกมาไม่เหมือนกัน อาจมีความคล้ายคลึงกันแต่ในขณะเดียวกันจะมีความแตกต่างอยู่บ้าง จึงไม่สามารถหาข้อสรุปที่ชัดเจนได้ ทำให้ผู้วิจัยสนใจที่จะนำผลการทำนายที่ได้มาศึกษาเปรียบเทียบเพื่อนำไปพิจารณาหาความสัมพันธ์

วิทยานิพนธ์ฉบับนี้จะศึกษาเทคนิคการทำนายยีนด้วยวิธีต่าง ๆ ทั้งในเชิงทฤษฎี และการนำไปใช้งาน โดยเลือกนำผลการทำนายยีนของโปรแกรม Glimmer version 3.0 และผลการทำนายของโปรแกรมที่พัฒนาขึ้นเองโดยใช้ทฤษฎีของ Fickett J.W. (1982) มาศึกษาทดลอง และทำการเปรียบเทียบผลการทำนายยีนเพื่อนำไปพิจารณาหาความสัมพันธ์ของผลการทำนายยีนด้วยวิธีทั้งสอง

1.2 วัตถุประสงค์ของการวิจัย

- 1) เพื่อเปรียบเทียบผลการทำนายยีนของสิ่งมีชีวิตด้วยวิธีฟิคเกตและกลิมเมอร์

2) เพื่อพิจารณาหาความสัมพันธ์ระหว่างผลการทำนายของโปรแกรมทำนายของสิ่งมีชีวิตด้วยวิธีฟิคเกตและกลิมเมอร์

1.3 ประโยชน์ที่จะได้รับจากการศึกษาเชิงทฤษฎี และ เชิงประยุกต์

- 1) ในเชิงทฤษฎี จะทำให้เข้าใจถึงวิธีการทำนาย ทั้งแนวคิด และขอบเขตของเทคนิคการทำนาย
- 2) ในเชิงประยุกต์ จะได้นำผลการพิจารณาความสัมพันธ์ของการทำนายไปเป็นตัวช่วยในการตัดสินใจเลือกวิธีการทำนายที่เหมาะสมต่อความต้องการ

1.4 ขอบเขตการวิจัย

ศึกษาเทคนิคการทำนายด้วยวิธีต่างๆ พัฒนาโปรแกรมทำนายจากทฤษฎีของ Fickett J.W. (1982) โดยข้อมูลที่ใช้ในการศึกษาคือชุดลำดับดีเอ็นเอของแบคทีเรีย ซึ่งแบ่งเป็นลำดับเบสย่อย ๆ จำนวน 385 ไฟล์ เพื่อนำผลการทำนายที่ได้ไปศึกษาทดลองหาความสัมพันธ์กับผลการทำนายที่ได้จากโปรแกรม Grimmer version 3.0

1.5 วิธีการวิจัย

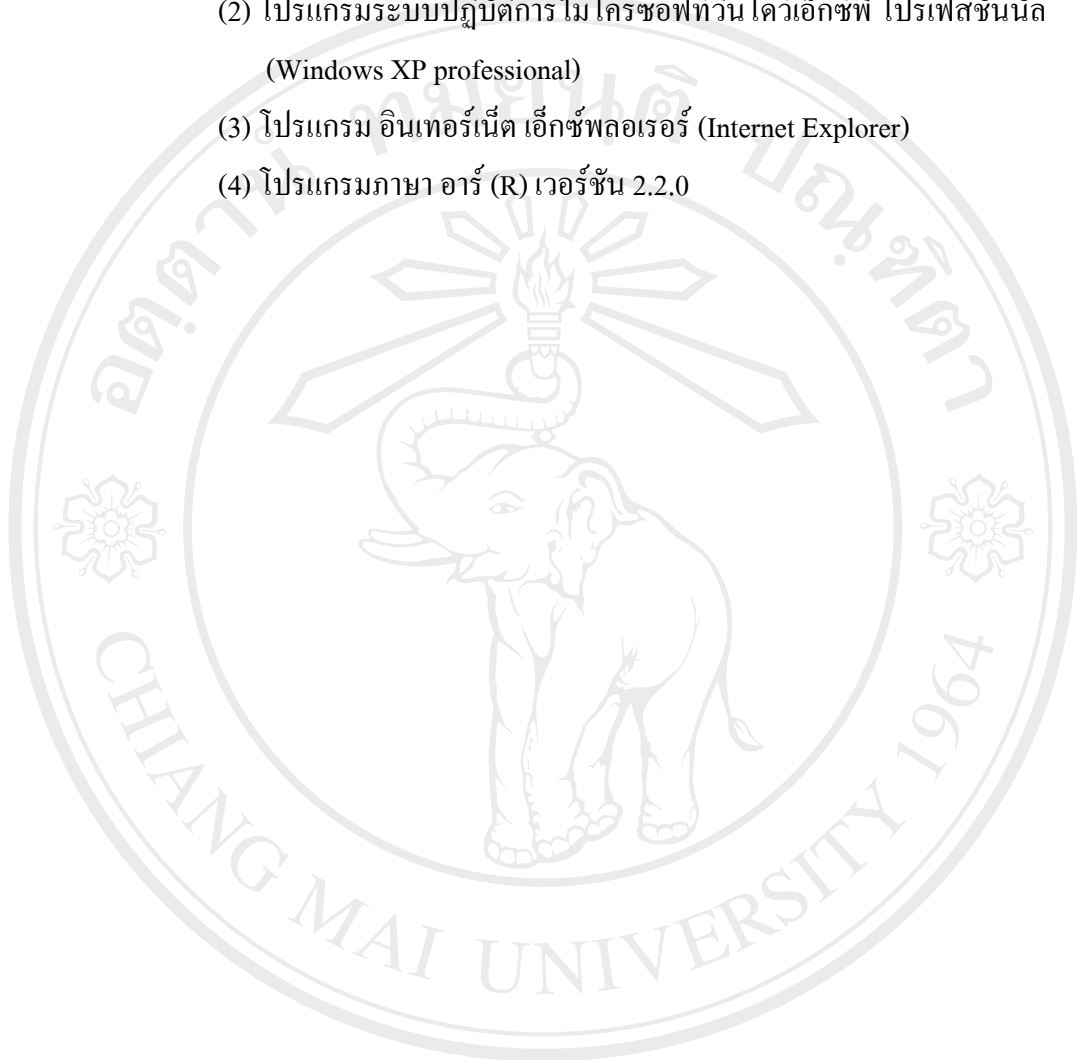
- 1) ศึกษาเทคนิคการทำนายด้วยวิธีต่างๆ ในเชิงทฤษฎี และการนำไปใช้งาน
- 2) พัฒนาโปรแกรมทำนายจากทฤษฎีของ Fickett J.W. (1982) และนำโปรแกรมดังกล่าวไปใช้ในการทำนายของข้อมูลลำดับดีเอ็นเอแบคทีเรีย
- 3) นำผลการทำนายที่ได้ ไปศึกษาทดลองหาความสัมพันธ์กับผลการทำนายที่ได้จากโปรแกรม Grimmer version 3.0
- 4) สรุปผลการทดลอง และจัดทำเอกสารสรุปผลงานวิจัยทั้งหมด

1.6 สถานที่ที่ใช้ในการดำเนินการวิจัยและรวบรวมข้อมูล

- 1) สถานที่
 - (1) ภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเชียงใหม่
 - (2) สำนักหอสมุด มหาวิทยาลัยเชียงใหม่
 - (3) ห้องปฏิบัติการวิจัย ชีวสารสนเทศศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่

2) อุปกรณ์ที่ใช้ในการวิจัย

- (1) เครื่องคอมพิวเตอร์ส่วนบุคคล (Personal Computer)
- (2) โปรแกรมระบบปฏิบัติการไมโครซอฟท์วินโดวส์เอ็กซ์พี โปรเฟสชันนัล
(Windows XP professional)
- (3) โปรแกรม อินเทอร์เน็ต เอ็กซ์พลอเรอร์ (Internet Explorer)
- (4) โปรแกรมภาษา อาร์ (R) เวอร์ชัน 2.2.0



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved