

คมสัน จันมา 2550: ขั้นตอนวิธีการเทียบเรียงกลุ่มลำดับข้อมูลชีวภาพโดยพิจารณา
ความถี่ส่วนย่อยของลำดับ ปรินซ์มหาวิทยาลัยมหาสารคาม (วิศวกรรม
คอมพิวเตอร์) สาขาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ ปรินซ์
กรรมการที่ปรึกษา: ผู้ช่วยศาสตราจารย์พันธุ์พิติ เปี่ยมสง่า, D.Sc. 100 หน้า

การเทียบเรียงกลุ่มลำดับข้อมูลทางชีวภาพเป็นงานที่จำเป็นสำหรับนักชีววิทยา ซึ่งใน
ปัจจุบันจำนวนของข้อมูลทางด้านชีวภาพได้เพิ่มขึ้นอย่างรวดเร็ว ดังนั้นเครื่องมือสำหรับการเทียบ
เรียงกลุ่มลำดับข้อมูลทางชีวภาพจึงจำเป็นต้องมีความถูกต้องในระดับที่นักชีววิทยาพอใจ และใช้
เวลาในการประมวลผลน้อย งานวิจัยนี้จึงนำเสนอวิธีการเทียบเรียงกลุ่มลำดับข้อมูลทางชีวภาพที่
ใช้เวลาในการทำงานน้อยมาก โดยที่ความถูกต้องเทียบเคียงได้กับเครื่องมือที่ใช้ในปัจจุบัน โดยใช้
วิธี n-gram ตัดข้อมูลเป็นส่วนย่อย และพิจารณาจากความถี่ของส่วนย่อยของแต่ละคู่ลำดับข้อมูล
เพื่อใช้เป็นแนวทางในการเทียบเรียง วัดความถูกต้อง และเวลาการประมวลผลโดยใช้ชุดข้อมูล
ทดสอบ BALiBASE และ PREFAB ผลลัพธ์ที่ได้จากการเทียบเรียงชุดทดสอบ PREFAB สามารถ
เทียบเรียงได้เร็วกว่า ClustalW 6.24 เท่า โดยที่ความถูกต้องอยู่ในระดับเดียวกัน