

ชื่อเรื่องการค้นคว้าแบบอิสระ ระบบการจัดกลุ่มข้อมูลการถอดรหัสพันธุกรรม

ของดีเอ็นเอ

ผู้เขียน

นางสาวชโลธร เหลี่ยมวิรัช

ปริญญา

วิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์)

อาจารย์ที่ปรึกษาการค้นคว้าแบบอิสระ

อาจารย์ ดร. จิรยุทธ ไชยจารุณิช

บทคัดย่อ

ข้อมูลการถอดรหัสพันธุกรรมของดีเอ็นเอได้มาจากการทดลองไมโครอาร์เรย์ดีเอ็นเอในห้องทดลองจุลชีววิทยา ข้อมูลการถอดรหัสพันธุกรรมของดีเอ็นเอถูกจัดเก็บอยู่ในรูปแบบของตารางที่เก็บข้อมูลเชิงตัวเลขซึ่งสามารถนำไปวิเคราะห์ด้วยวิธีการคำนวณเพื่อให้เข้าใจถึงหลักเกณฑ์พื้นฐานของการเจริญเติบโตและการพัฒนาการของสิ่งมีชีวิต ตลอดจนตรวจหาสาเหตุทางพันธุกรรมที่แฝงอยู่ของโรคที่เกิดขึ้นกับมนุษย์ ระบบการจัดกลุ่มข้อมูลการถอดรหัสพันธุกรรมของดีเอ็นเอที่พัฒนาขึ้นนี้มีวัตถุประสงค์เพื่อเป็นเครื่องมือช่วยวิเคราะห์ข้อมูลการถอดรหัสพันธุกรรมของดีเอ็นเอด้วยเทคนิคการจัดกลุ่มข้อมูลโดยใช้ขั้นตอนวิธีต่าง ๆ ได้แก่ การจัดกลุ่มโดยใช้ขั้นตอนวิธีเค-มีนส์ การจัดกลุ่มแบบลำดับชั้นด้วยขั้นตอนแบบซิงเกิลลิงก์เกจ การจัดกลุ่มแบบลำดับชั้นด้วยขั้นตอนแบบคอมพลิติกเกจ และ การจัดกลุ่มแบบลำดับชั้นด้วยขั้นตอนแบบเซ็นทรอยด์ลิงก์เกจ นอกจากนี้ยังได้นำเอาขั้นตอนวิธีของการจัดกลุ่มที่ผู้เขียนได้มีส่วนในการวิจัยและพัฒนาร่วมกับอาจารย์ที่ปรึกษาการค้นคว้าแบบอิสระมาใช้ในระบบนี้ด้วย ซึ่งได้แก่ การจัดกลุ่มโดยใช้ขั้นตอนวิธีเค-มีนส์ที่พิจารณาขอบเขตรัศมีของกลุ่ม และ การจัดกลุ่มโดยใช้ขั้นตอนวิธีเค-มีนส์หลายระดับ ระบบจะแสดงผลพัทธ์การจัดกลุ่มที่ได้อยู่ในรูปแบบที่เป็นกราฟฟิค และ จัดเก็บผลลัพธ์การจัดกลุ่มเป็นไฟล์ข้อความเพื่อให้ระบบสามารถนำมาจัดแสดงในภายหลังได้ สิ่งเหล่านี้เพียงพอที่จะทำให้ระบบการจัดกลุ่มการถอดรหัสพันธุกรรมของดีเอ็นเอที่พัฒนาขึ้นนี้ เป็นเครื่องมือที่ให้ความสะดวกแก่นักชีววิทยาในการวิเคราะห์ข้อมูลการถอดรหัสพันธุกรรมของดีเอ็นเอ

Independent Study Title DNA Transcription Data Clustering System

Author Miss Chalothorn Liamwirat

Degree Master of Science (Computer Science)

Independent Study Advisor Lecturer Dr. Jeerayut Chaijaruwanich

ABSTRACT

DNA microarray experiments in biological laboratory are sources of DNA transcription data. DNA transcription data is formed into a numerical table that can be analysed by several computational methods to understand fundamental aspects of growth and development as well as to explore the underlying genetic causes of human diseases. In this study, DNA transcription data clustering system is developed. It is a tool that helps the analysis of DNA transcription data using clustering techniques. Clustering techniques developed in this system include the k-means algorithm, single linkage hierarchical clustering, complete linkage hierarchical clustering, centroid linkage hierarchical clustering, the cluster radius bounded k-means algorithm and the multi-level k-means algorithm. The last two methods are originally proposed by the author in cooperation with author's independent study adviser. The results of clustering are displayed in graphic format and kept in text file that can be subsequently displayed the system. This DNA transcription data clustering system is shown to be satisfied by the user requirements and eases biologist for DNA transcription data analysis.