

CLASSIFICATION OF DENGUE VIRUS SEROTYPES BY VISUALIZATION TOOL**BOONYARAT VIRIYASAKSATHIAN 5137363 EGBE/M****M.Eng. (BIOMEDICAL ENGINEERING)**

THESIS ADVISORY COMMITTEE: YODCHANAN WONGSAWAT, Ph.D. (Electrical Engineering), PRAPAT SURIYAPHOL, Ph.D. (Biochemistry), NORASED NASONGKLA, Ph.D. (Polymer Science)

ABSTRACT

Software for analysis and interpretation of nucleotide sequences can display the result in different ways: color-coded bar, letter to letter and graphical display, *etc.* However, there is still no efficient software that can simultaneously classify the serotypes of Dengue virus and visually illustrate the nucleotide sequences of Dengue virus. Therefore, this research aimed to study and develop software, called CADViST, for analyzing the nucleotide sequences of Dengue virus with a novel graphical representation, called UnitX. The tasks of the thesis were divided into two parts: (1) design a graphic to represent the nucleotide sequences of Dengue virus, then (2) develop software that provides the following functions: (2.1) developing a nucleotide sequence search engine of the Dengue virus, (2.2) assigning the serotype to the unknown Dengue virus sequence, and (2.3) displaying the Basic Local Alignment Search Tool (BLAST) search result.

This thesis designed UnitX to display the relationship between amino and keto functions in graphical display. For the purpose of verification, UnitX was employed to classify the nucleotide sequences of Dengue virus. All 2,184 sample sequences, from GenBank database of the National Center for Biotechnology Information (NCBI), USA, comprise four serotypes of Dengue virus sequence: DENV1 to 4. Each serotype contains 952, 737, 405 and 90, nucleotide sequences, respectively. The experimental result concluded that UnitX was able to efficiently classify the nucleotide sequences of Dengue virus serotype. Hence, UnitX was able to be used as the graphical representation of the CADViST software for providing a facility to users in nucleotide analysis. For assigning the serotype, one-tenth of each serotype was randomly selected for constructing the template sequence via multiple sequence alignment, called ClustalW. BLAST was employed to provide the serotype to the unknown Dengue virus sequence. The results showed that the template sequence representing each serotype could efficiently distinguish the different serotypes of Dengue virus.

KEY WORDS: BIOINFORMATICS / CLASSIFICATION / DENGUE VIRUS / SEROTYPE /
VISUALIZATION / BLAST

106 pages

การจำแนกซีโรทัยป์ของไวรัสเด็งกีด้วยโปรแกรมคอมพิวเตอร์แบบวิซวลไลเซชัน

CLASSIFICATION OF DENGUE VIRUS SEROTYPES BY VISUALIZATION TOOL

บุญรัตน์ วิริยะศักดิ์เสถียร 5137363 EGBE / M

วศ.ม. (วิศวกรรมชีวการแพทย์)

คณะกรรมการที่ปรึกษาวิทยานิพนธ์ : ยศชนัน วงศ์สวัสดิ์, Ph.D. (Electrical Engineering),
ประพัฒน์ สุริยผล Ph.D. (Biochemistry), นรเศรษฐ์ ณ สงขลา, Ph.D. (Polymer Science)

บทคัดย่อ

ซอฟต์แวร์สำหรับการวิเคราะห์และแปลผลลำดับเบสของนิวคลีโอไทด์สามารถแสดงผลได้หลายแบบ เช่น การแสดงผลเป็นแถบสี เป็นตัวอักษรแบบตำแหน่งต่อตำแหน่ง และการแสดงผลแบบกราฟิก เป็นต้น แต่ยังไม่มียซอฟต์แวร์ที่มีประสิทธิภาพในการจำแนกซีโรทัยป์ของเชื้อไวรัสเด็งกีและทำให้เห็นภาพรวมของสายนิวคลีโอไทด์ทั้งเส้นของเชื้อไวรัสเด็งกี ดังนั้น โครงการวิจัยในวิทยานิพนธ์นี้ จึงมุ่งเน้นการศึกษาและพัฒนาซอฟต์แวร์โดยตั้งชื่อว่า CADViST สำหรับใช้ในการวิเคราะห์ลำดับเบสของนิวคลีโอไทด์ของเชื้อไวรัสเด็งกีด้วยการนำเสนอข้อมูลในรูปแบบกราฟิกที่ถูกพัฒนาขึ้นมาใหม่ชื่อ UnitX เป้าหมายของงานวิจัยแบ่งออกเป็น 2 ส่วน ดังนี้คือ (1) ออกแบบกราฟิกที่นำมาใช้ในการแสดงลำดับเบสของนิวคลีโอไทด์ของเชื้อไวรัสเด็งกี (2) พัฒนาซอฟต์แวร์ให้มีฟังก์ชันการทำงานดังนี้คือ (2.1) เป็นซอฟต์แวร์ที่ใช้ในการสืบค้นหาข้อมูลลำดับเบสของนิวคลีโอไทด์ของเชื้อไวรัสเด็งกี (2.2) เป็นซอฟต์แวร์ที่ใช้ในการจำแนกซีโรทัยป์ของลำดับเบสของนิวคลีโอไทด์ของเชื้อไวรัสเด็งกี (2.3) เป็นซอฟต์แวร์ที่ใช้แสดงผลจากโปรแกรม Basic Local Alignment Search Tool (BLAST) งานวิจัยนี้นำเสนอ UnitX ไปใช้ในการแสดงความสัมพันธ์ของปริมาณระหว่างหมู่ฟังก์ชันอะมิโนและคีโตในรูปของกราฟิก เพื่อทดสอบความสามารถของ UnitX UnitX จึงถูกนำไปใช้ในการจำแนกซีโรทัยป์ของลำดับนิวคลีโอไทด์ของเชื้อไวรัสเด็งกี นิวคลีโอไทด์ตัวอย่างจำนวน 2,148 ตัวอย่าง ได้มาจากฐานข้อมูล GenBank ของ National Center for Biotechnology Information (NCBI) ประกอบด้วย 4 ซีโรทัยป์ คือ ซีโรทัยป์ 1 ถึง 4 แต่ละซีโรทัยป์ประกอบด้วย 952, 737, 405 และ 90 ตามลำดับ ผลจากการทดลองสรุปได้ว่า UnitX สามารถนำมาใช้ในการจำแนกซีโรทัยป์ทั้ง 4 ของเชื้อไวรัสเด็งกีได้ เมื่อได้ผลดังนี้ จึงนำ UnitX ไปใช้เป็นหน่วยแสดงผลของซอฟต์แวร์ CADViST ซึ่งเป็นซอฟต์แวร์ที่ถูกพัฒนาเพื่ออำนวยความสะดวกแก่ผู้ใช้ให้สามารถวิเคราะห์ลำดับนิวคลีโอไทด์ได้ง่ายขึ้น 10% ของแต่ละซีโรทัยป์ถูกนำมาทำ Multiple Sequence Alignment เพื่อสร้างสายนิวคลีโอไทด์ของแต่ละ ซีโรทัยป์ด้วยโปรแกรม ClustalW โปรแกรม BLAST ถูกนำมาใช้เพื่อจำแนกซีโรทัยป์ของเชื้อไวรัสเด็งกี ผลลัพธ์แสดงให้เห็นว่า 10% นิวคลีโอไทด์ใน 4 ซีโรทัยป์ที่นำมาใช้สามารถระบุซีโรทัยป์ได้ถูกต้องทั้งหมด