

รายงานวิจัยฉบับสมบูรณ์

การพัฒนาซอฟต์แวร์ที่มีประสิทธิภาพสำหรับการวิเคราะห์ปฏิสัมพันธ์
ระหว่าง SNPs (Single Nucleotide Polymorphisms)
ที่มีส่วนเกี่ยวข้องต่อการเกิดโรค

Development of efficient software for analyses of SNP-SNP
interactions associated a disease

นักวิจัย
ดร.อุณิตษา สังข์เกตุ

โครงการวิจัยนี้ได้รับทุนสนับสนุนจากเงินรายได้มหาวิทยาลัย
มหาวิทยาลัยสงขลานครินทร์
ประจำปีงบประมาณ 2555 รหัสโครงการ SCI550388S

ชื่อโครงการ

ภาษาไทย: การพัฒนาซอฟต์แวร์ที่มีประสิทธิภาพสำหรับการวิเคราะห์ปฏิสัมพันธ์ระหว่าง SNPs (Single Nucleotide Polymorphisms) ที่มีส่วนเกี่ยวข้องต่อการเกิดโรค

ภาษาอังกฤษ: Development of efficient software for analyses of SNP-SNP interactions associated a disease

คณະนักวิจัย และหน่วยงานต้นสังกัด (คณະ/ภาควิชา หรือหน่วยงาน)

ดร.อุณิษา สังข์เกตุ ภาควิชาเทคโนโลยีชีวภาพโมเลกุลและชีวสารสนเทศ คณะวิทยาศาสตร์

กิตติกรรมประกาศ

งานวิจัยนี้สำเร็จลุล่วงลงได้ด้วยดีจากการช่วยเหลือและให้คำแนะนำเป็นอย่างดีจาก รศ.ดร.นพ. สุรศักดิ์ สังข์ทนต์ ณ อยุธยา และ ดร.นพ.สุรคเมธ มหาศิริมงคล ผู้เขียนขอกราบขอบพระคุณทั้งสองท่านมา ณ ที่นี้

บทคัดย่อ

การปฏิสัมพันธ์ระหว่าง SNPs (Single Nucleotide Polymorphisms) มีส่วนเกี่ยวข้องต่อการเกิดโรค เช่น โรคไข้หวัดใหญ่ เป็นต้น โรคไข้หวัดใหญ่เกิดจากเชื้อไวรัสไข้หวัดใหญ่ ซึ่งเป็นโรคที่ร้ายแรงและคร่าชีวิตผู้คนไปมากมาย เนื่องจากเชื้อไวรัสไข้หวัดใหญ่มีการปรับเปลี่ยนพันธุกรรมของตัวเองเพื่อสามารถต้านวัคซีนได้ ดังนั้นจึงจำเป็นต้องเข้าใจการปฏิสัมพันธ์ระหว่าง SNPs ของเชื้อไวรัสไข้หวัดใหญ่ในระหว่างการเกิดวิวัฒนาการของพันธุกรรม เพื่อใช้สำหรับพัฒนาวัคซีนที่ป้องกันเชื้อดังกล่าวได้ในอนาคต ในการทำความเข้าใจการปฏิสัมพันธ์ของ SNPs ในช่วงวิวัฒนาการของพันธุกรรมของเชื้อไวรัสไข้หวัดใหญ่ จะต้องจัดกลุ่มพันธุกรรมที่เหมือนกันหรือต่างกันของเชื้อไวรัสที่ระบาดในแต่ละช่วง ซึ่งในการจัดกลุ่มดังกล่าวถ้าจัดกลุ่มด้วยตามมนุษย์จะใช้เวลาานาน อีกทั้งมีความผิดพลาดเกิดขึ้นตามมา ในงานวิจัยนี้จึงได้ทำการพัฒนาซอฟต์แวร์ที่มีประสิทธิภาพสำหรับการวิเคราะห์ปฏิสัมพันธ์ระหว่าง SNPs ที่มีชื่อว่า SNPer เพื่อช่วยในการลดระยะเวลาในการวิเคราะห์ และกำจัดความผิดพลาดดังกล่าว จากการทดลองจัดกลุ่ม SNPs ในกลุ่มประชากรตัวอย่างไวรัสไข้หวัดใหญ่ 3 ตัวอย่างได้ SNPs ออกมา 3 กลุ่มคือ 1) SNPs ที่เหมือนกันทั้ง 3 ตัวอย่าง 2) SNPs ที่เหมือนกันในบางตัวอย่าง 3) SNPs ที่ไม่ซ้ำกันเลยใน 3 ตัวอย่าง จากการทดลอง SNPer ใช้เวลาในการจัดกลุ่ม SNPs ชุดนี้เพียงแค่ 3 วินาที โดยไม่เกิดข้อผิดพลาดใดๆ ซึ่งจะเห็นได้ว่าลดระยะเวลาในการทำงานเป็นอย่างมาก เมื่อเทียบกับการจัดกลุ่มด้วยตามมนุษย์ที่ใช้ระยะเวลาประมาณ 40 ชั่วโมง พร้อมกับมีข้อผิดพลาดเกิดขึ้น ทั้งนี้ SNPer สามารถถูกนำมาใช้ในการจัดกลุ่มของสิ่งมีชีวิตอื่นได้อีกด้วย

Abstract

SNP-SNP interactions associate diseases such as influenza. Influenza is an infectious disease caused by the influenza virus (IFV) and outbreaks each year. The vaccine resistance can be appeared if the genetics shifts or drifts of IFV are occurred. Close monitoring of genetics evolution of IFV is essential information for emerging seasonal influenza vaccine design and composition to protect people from the IFV. The SNP-SNP interactions of IFV of each influenza epidemic are classified for understanding the genetics evolution of IFV. However, classifying SNPs by manual visualization is time consuming and produces some error. Therefore, SNPer program was developed to speed up the capacity and eliminate errors of grouping the SNPs. The SNPs of IFV populations were grouped the into (1) universal SNPs, (2) likely common SNPs, and (3) unique SNPs. SNPer executed only three seconds with no error of SNP classification events analyzed with 40 hours with errors using manual visualization. SNPer can perform not only on SNPs of IFV but also on SNPs of other organisms.

บทสรุปผู้บริหาร (Executive Summary)

1. บทนำ (เขียนสั้น ๆ ถึงปัญหาและที่มาของการทำวิจัยเรื่องนี้ เหมือนการเขียน Introduction ของ Manuscript)

ในปัจจุบันได้มีการคิดค้นวัคซีนที่เพื่อป้องกันมนุษย์จากเชื้อโรคมามากมาย เช่น เชื้อไวรัสไข้หวัดใหญ่ เพื่อป้องกันโรคไข้หวัดใหญ่ ซึ่งโรคไข้หวัดใหญ่เป็นโรคที่อันตราย คนที่เป็นโรคนี้อาจเสียชีวิตได้ ถ้าไม่ได้รับการรักษาอย่างทันท่วงที อย่างไรก็ตามถึงแม้ว่าจะมีวัคซีนป้องกันเชื้อไวรัสไข้หวัดใหญ่ แต่เชื้อไวรัสไข้หวัดใหญ่มีการเปลี่ยนแปลงพันธุกรรมของตัวเองตลอดเวลาเพื่อดำเนินชีวิต ดังนั้นจึงจำเป็นต้องทำความเข้าใจการเปลี่ยนแปลงพันธุกรรมในช่วงวิวัฒนาการของไวรัสไข้หวัดใหญ่ เพื่อนำมาเป็นข้อมูลในการพัฒนาวัคซีนสำหรับต้านเชื้อไวรัสที่พันธุกรรมได้เปลี่ยนแปลงต่อไป การเปลี่ยนแปลงพันธุกรรมของไวรัสทำให้เกิดการกลายพันธุ์ระหว่าง SNPs (Single Nucleotide Polymorphisms) เช่น บาง SNPs เกิดขึ้นเฉพาะในกลุ่มประชากรไวรัสที่ต้านวัคซีน แต่ไม่เกิดขึ้นในกลุ่มประชากรไวรัสที่ไม่ต้านวัคซีน เป็นต้น อย่างไรก็ตามในการวิเคราะห์หาการกลายพันธุ์ระหว่าง SNPs ด้วยตามนุษย์จะใช้ระยะเวลาในการเปรียบเทียบ SNPs นานมาก ทำให้สายตาสลัดขี้ตา ส่งผลให้เกิดข้อผิดพลาดตามมา ในงานวิจัยนี้จึงได้สังเกตเห็นถึงความสำคัญของการพัฒนาซอฟต์แวร์ที่ช่วยลดระยะเวลา และกำจัดความผิดพลาดในการวิเคราะห์การกลายพันธุ์ระหว่าง SNPs

2. วัตถุประสงค์

เพื่อพัฒนาซอฟต์แวร์ที่มีประสิทธิภาพในการหาการกลายพันธุ์ระหว่าง SNPs

3. สรุป (สรุปผลการทดลองทั้งหมดของงานวิจัยทั้งชุดโครงการ/โครงการ ทั้งตีพิมพ์แล้วและยังไม่ได้ตีพิมพ์)

ในงานวิจัยนี้ได้พัฒนาซอฟต์แวร์ที่มีชื่อว่า SNPer เพื่อลดระยะเวลาและขจัดความผิดพลาด ในการวิเคราะห์การกลายพันธุ์ระหว่าง SNPs ของกลุ่มประชากรไวรัสไข้หวัดใหญ่ ได้ทำการทดลองวิเคราะห์หาการกลายพันธุ์ระหว่าง SNPs ของกลุ่มประชากรไวรัสโรคไข้หวัดใหญ่ 3 กลุ่มตัวอย่าง ได้ลักษณะการกลายพันธุ์ของ SNPs 3 ลักษณะคือ 1) SNPs ที่เหมือนกันทั้ง 3 กลุ่มตัวอย่าง 2) SNPs ที่เหมือนกันในบางกลุ่มตัวอย่าง 3) SNPs ที่ไม่ซ้ำกันเลยใน 3 กลุ่มตัวอย่าง ผลการทดลองที่ได้ สามารถนำมาใช้ในการพัฒนาวัคซีนได้ ยกตัวอย่างเช่น ถ้ากำหนดให้กลุ่มตัวอย่างที่ 1 และกลุ่มตัวอย่างที่ 2 เป็นเชื้อที่ไม่ต้านวัคซีน ส่วนกลุ่มตัวอย่างที่ 3 เป็นเชื้อที่ต้านวัคซีน จะได้ว่า SNPs ที่อยู่ในเฉพาะกลุ่มตัวอย่างที่ 3 เป็น SNPs ที่อันตรายและนำมาเป็นข้อมูลสำคัญในการพัฒนาวัคซีน เพื่อป้องกันเชื้อไวรัสในกลุ่มตัวอย่างที่ 3 ต่อไป SNPer ทำงานได้อย่างมีประสิทธิภาพมาก เพราะจากการทดลอง SNPer ใช้เวลาในการจัดกลุ่ม SNPs ของกลุ่มประชากรไวรัสไข้หวัดใหญ่ชุดนี้เพียงแค่ 3 วินาที โดยไม่เกิดข้อผิดพลาดใดๆ ซึ่งเป็นการจัดกลุ่มข้อมูล SNPs ได้รวดเร็วมาก เมื่อเปรียบเทียบกับการจัดกลุ่มด้วยตาของผู้เชี่ยวชาญที่ใช้ระยะเวลาประมาณ 40 ชั่วโมง และมีข้อผิดพลาดเกิดขึ้น นอกจากนี้ SNPer จะวิเคราะห์ SNPs ได้อย่างรวดเร็วแล้ว SNPer สามารถถูกนำมาใช้วิเคราะห์การกลายพันธุ์ระหว่าง SNPs ของสิ่งมีชีวิตอื่นได้อีกด้วย

ภาคผนวก

1. แนบสำเนาบทความที่ได้รับการตีพิมพ์แล้ว (Reprint)

ตามเอกสารท้ายเล่ม

2. ผลการวิจัยส่วนที่ยังไม่ได้ตีพิมพ์

ได้เขียน manuscript สำหรับการตีพิมพ์ในหัวข้อเรื่อง “Analyses of SNP-SNP Interactions Applying Parallel Computing” ดังเอกสารท้ายเล่ม โดยได้เปลี่ยนชื่อซอฟต์แวร์จากเดิม ParallLogicReg เป็น SNPInt และกำลังอยู่ในขั้นตอนการทดลองใช้ซอฟต์แวร์ SNPInt มาวิเคราะห์ข้อมูลใหม่ (โรคมะเร็งลำไส้) ซึ่งใช้เวลานานกว่าที่คาดการณ์เอาไว้เนื่องจากต้องแปลงข้อมูลให้อยู่ในรูปแบบที่ SNPInt ต้องการ

3. ข้อคิดเห็นและข้อเสนอแนะสำหรับการวิจัยต่อไป

เพื่อให้ซอฟต์แวร์ SNPer ใช้งานได้ง่ายกว่าเดิม อาจจะปรับปรุงซอฟต์แวร์ SNPer ให้สามารถบริการในรูปแบบเว็บเซอร์วิสหรือเว็บไซต์ โดยที่ผู้ทำการวิจัยได้เขียนเอกสารสำหรับเจ้าหน้าที่เทคนิคและเอกสารสำหรับผู้ใช้งานเอกสารท้ายเล่ม เพื่อเป็นความรู้พื้นฐานในการต่อยอด