

***IN SILICO* APPROACH FOR DIFFERENTIATION OF  
HYPOCHROMIC MICROCYTIC ANEMIA AND DYSLIPIDEMIA**



**ABDUL HAFEEZ KANDHRO**

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY (MEDICAL TECHNOLOGY)  
FACULTY OF GRADUATE STUDIES  
MAHIDOL UNIVERSITY**

Copyright by Mahidol University 2017

**COPYRIGHT OF MAHIDOL UNIVERSITY**

*IN SILICO* APPROACHES FOR DIFFERENTIATION OF HYPOCHROMIC MICROCYTIC ANEMIA AND DYSLIPIDEMIA

ABDUL HAFEEZ KANDHRO 5637869 MTMT/D

Ph.D. (MEDICAL TECHNOLOGY)

THESIS ADVISORY COMMITTEE: PORNLADA NUCHNOI, Ph.D. WATSHARA SHOOMBUEATONG, Ph.D., SUREERUT PORNTADAVITY, Ph.D., VIRAPONG PRACHAYASITTIKUL, Ph.D., CHOTIROS PLABPLUENG, Ph.D.

ABSTRACT

Bioinformatics is a powerful research tools for facilitating disease diagnosis, treatment and prevention. In this study, the researcher systematically applied bioinformatics approach for solving common diseases/conditions in Asia. Unfortunately, the lack of special tests has hampered the differential and definitive diagnosis in poor resource setting areas. Several discrimination formulas have been used as simply screening tool to discriminate thalassemia traits from IDA. Therefore, twelve common existing discrimination formulas were applied in this study using Pakistan population as model. Six formulas (MI, EF, G&K, RDWI, R, and HHI) demonstrated the most reliability. Random Forest approach was used to improve cutoff values of the remaining six formulas. Furthermore, the researcher generated two new discrimination formulas with proposed cutoff values. The new formulas demonstrated high accuracies and the Youden's index. This indicated that the new proposed formulas and the six existing formulas with adjusted cutoff values discriminate IDA and thalassemia trait cases. We applied Text-mining approach was applied for constructing molecular network of miRNA related dyslipidemia. In order to comprehensively visualize the interaction of miRNAs and the role the related genes play in dyslipidemia, the researcher retrieved public databases information for microRNA and dyslipidemia associations, a total 227 associations including 148 microRNAs for constructing and visualizing interaction network by Cytoscape. The top 20 significant microRNAs used to derive predicted and validated gene targets by using CyTargetLinker. Predicted gene targets showed significant biological processes for cholesterol, lipid and fatty acids under the gene ontology analysis by BiNGO on Cytoscape. This study uncovered the complex network of miRNA in biological and pathological mechanism related to dyslipidemia. Finally, the impact of bioinformatics approach for laboratory application was demonstrated.

KEY WORDS: DYSLIPIDEMIA / miRNA / TEXT-MINING / THALASSEMIA

Copyright by Mahidol University

128 pages

การประยุกต์ใช้ชีวสารสนเทศศาสตร์เพื่อการศึกษาโรคโลหิตจางชนิด hypochromic microcytic anemia และภาวะไขมันในเลือดสูง  
*IN SILICO* APPROACH FOR DIFFERENTIATION OF HYPOCHROMIC MICROCYTIC ANEMIA AND DYSLIPIDEMIA

ABDUL HAFEEZ KANDHRO 5637869 MTMT/D

ปร.ด. (เทคนิคการแพทย์)

คณะกรรมการที่ปรึกษาวิทยานิพนธ์: พรลดา นุชน้อย, ปร.ด., วีระช รุ่งบัวทอง, ปร.ด. สุวีรัตน์ พรธาดาวีทย์, ปร.ด., วีระพงศ์ ปรัชญาสิทธิกุล, ปร.ด. โชติรส พลับพลึง, ปร.ด.

บทคัดย่อ

ชีวสารสนเทศศาสตร์เป็นเทคโนโลยีประสิทธิภาพสูงในการวิจัยเพื่อวินิจฉัย รักษาและป้องกันโรค การศึกษาได้ประยุกต์ใช้เทคโนโลยีชีวสารสนเทศศาสตร์เพื่อแก้ไขปัญหาของโรคที่พบได้บ่อยในทวีปเอเชีย เช่น โรคโลหิตจางชนิด hypochromic microcytic anemia และภาวะไขมันในเลือดสูง โรคโลหิตจาง ชนิด hypochromic microcytic anemia ที่พบได้บ่อยคือ โรคทาลัสซีเมียและโรคโลหิตจางจากการขาดธาตุเหล็ก จากการศึกษาในประเทศปากีสถาน พบพาหะของโรคทาลัสซีเมียมากที่สุดในเขตอำเภอ Badin (35.27%) และโรคโลหิตจางจากการขาด ธาตุเหล็กพบมากที่สุดในเขตอำเภอ Larkana (30.73%) ลักษณะทางโลหิตวิทยาของทั้งสองโรคนี้อาจคล้ายคลึง กันมาก จึงต้องใช้การทดสอบ serum ferritin เพื่อวินิจฉัยโรคโลหิตจางจากการขาดธาตุเหล็กและใช้การตรวจ วิเคราะห์ชนิดและปริมาณฮีโมโกลบินเพื่อวินิจฉัยโรคทาลัสซีเมีย ซึ่งในพื้นที่ที่ขาดแคลนอุปกรณ์และเครื่องมือที่จำเป็นทำให้การวินิจฉัยทำได้ยาก ในปัจจุบันมีการใช้สูตรคำนวณเพื่อเป็นการคัดกรองแยกทั้งสองโรคออกจากกัน ในการศึกษาได้ทำการวิเคราะห์สูตรคำนวณทั้ง 12 สูตรที่มีอยู่ในปัจจุบัน และสูตรคำนวณทางคณิตศาสตร์ขึ้นมาใหม่เพื่อใช้วินิจฉัยแยกทั้งสองโรคในกลุ่มประชากรนี้ พบว่าสูตรใหม่แสดงผลของความถูกต้อง และค่าการยอมรับในระดับสูง แสดงให้เห็นว่า สูตรใหม่ทางคณิตศาสตร์ที่ถูกเสนอขึ้นและ 6 สูตรที่มีอยู่ในปัจจุบันที่ปรับเปลี่ยนค่าจุดตัดสามารถแยกโรคทาลัสซีเมียออกจากโรคโลหิตจางจากการขาดธาตุเหล็กในกลุ่มประชากรนี้ ได้ นอกจากนี้ผู้วิจัยได้ใช้เทคนิคเหมืองข้อความ (text-mining) ในการสร้างเครือข่ายโมเลกุลไมโครอาร์เอ็นเอที่เกี่ยวข้องกับภาวะไขมันในเลือดสูง โดยการรวบรวมรูปแบบความสัมพันธ์ระหว่างไมโครอาร์เอ็นเอและภาวะไขมันในเลือดสูงจำนวนทั้งสิ้น 227 ความสัมพันธ์ ซึ่งประกอบด้วยไมโครอาร์เอ็นเอ 145 ตัว เพื่อทำการคัดเลือกทางสถิติจนได้ ไมโครอาร์เอ็นเอที่มีความสำคัญที่สุดจำนวน 20 ตัว เพื่อนำมาทำนายและตรวจสอบยืนยันเป้าหมายโดยใช้ CyTargetLinker พบผลเป็นที่น่าสนใจว่าอินเป้าหมายใหม่ที่ผู้วิจัยค้นพบโดยกระบวนการศึกษา gene ontology มีความสำคัญในการระบุการเผาผลาญไขมันของคลอเลสเทอรอล ไขมันและกรดไขมัน การศึกษานี้จึงมีบทบาทสำคัญและแสดงให้เห็นถึงการประยุกต์ใช้เทคนิคทางชีวสารสนเทศศาสตร์ในการอธิบายเครือข่ายโมเลกุลที่มีความซับซ้อน ซึ่งเป็นประโยชน์ในการศึกษาทางห้องปฏิบัติการและการสร้างองค์ความรู้จากข้อมูลที่มีอยู่อย่างมากมายต่อไป

คำสำคัญ: ไมโครอาร์เอ็นเอ/ ภาวะไขมันในเลือดสูง/ เทคนิคเหมืองข้อความ/ โรคทาลัสซีเมีย

128 หน้า