

Genetic Variation of the *CCL2* Promoter Polymorphism (rs1024611) in HIV-Infected Thai Patients

Nualprang Prasartam¹, Chareeporn Akekawatchai², Warisara Sretapunya³,
and Khaimuk Changsri^{2*}

¹Graduate Program in Medical Technology, Faculty of Allied Health Sciences,
Thammasat University, Pathum Thani Province, Thailand

²Department of Medical Technology, Faculty of Allied Health Sciences,
Thammasat University, Pathum Thani Province, Thailand

³Department of Medical Technology and Clinical Pathology,
Nakorn Nayok Hospital, Nakorn Nayok Province, Thailand

Abstract

C-C Motif Chemokine Ligand 2 (CCL2) plays a crucial role in inflammation and immune response. Single-nucleotide polymorphism (SNP) rs1024611 is associated with increased CCL2 protein expression, potentially affecting HIV-related complications. This study analyzes the genetic diversity of the *CCL2* gene promoter region, specifically 378 bp covering SNP rs1024611, among HIV-infected individuals in Thailand. DNA was extracted from whole blood samples, and the *CCL2* gene segment covering rs1024611 was amplified using polymerase-chain reaction (PCR) technique. Nucleotide sequences were analyzed by direct sequencing method. Bioinformatics analysis of 140 samples with qualified chromatograms identified nucleotide variations and examined association and linkage disequilibrium between SNPs. Only one additional SNP, rs2151345396, was found. The allelic frequencies of rs1024611 were 0.511 for allele A and 0.489 for allele G, with genotypic frequencies of 0.193 (AA), 0.636 (AG), and 0.171 (GG). The allelic frequencies of rs2151345396 were 0.907 (G) and 0.093 (C). Statistical analysis revealed no significant association between rs1024611 and rs2151345396, indicating their independence, despite strong linkage disequilibrium ($D' = 0.9985$, $r^2 = 0.0984$). In addition, the study found that the genetic diversity in the studied population differed significantly from the data in the SNP database LiveRef SNPs, dbSNP b156v2 and 1000

*Corresponding author E-mail address: khaimuk.c@allied.tu.ac.th

Genomes Phase 3) from NCBI database. This research underscores the importance of considering population-specific genetic diversity, emphasizing the need for nucleotide sequencing and diversity assessment before utilizing this information for gene expression studies and the development of molecular diagnostic methods targeting the *CCL2* gene.

Keywords: *CCL2* gene promoter, rs1024611, HIV infection, Genetic polymorphism

การวิเคราะห์ความหลากหลายทางพันธุกรรมบริเวณ โปรโมเตอร์ของยีน *CCL2* ที่เกี่ยวข้องกับ SNP rs1024611 ในผู้ติดเชื้อเอชไอวีในประเทศไทย

นवलปรารค์ ประศาสน์ธรรม¹ จริกรณ เอศวณิชชัย² วริศรา ศรีตะปัญญะ³ และ ไช้มุกด์ ช่างศรี^{2*}

¹บัณฑิตศึกษาสาขาเทคนิคการแพทย์ คณะสหเวชศาสตร์ มหาวิทยาลัยธรรมศาสตร์ จังหวัดปทุมธานี

²ภาควิชาเทคนิคการแพทย์ คณะสหเวชศาสตร์ มหาวิทยาลัยธรรมศาสตร์ จังหวัดปทุมธานี

³กลุ่มงานเทคนิคการแพทย์และพยาธิวิทยาคลินิก โรงพยาบาลนครนายก จังหวัดนครนายก

บทคัดย่อ

C-C Motif Chemokine Ligand 2 (CCL2) มีบทบาทสำคัญในกระบวนการอักเสบและการตอบสนองของระบบภูมิคุ้มกันและ single-nucleotide polymorphism (SNP) rs1024611 มีความสัมพันธ์กับการเพิ่มการแสดงออกของโปรตีน CCL2 ซึ่งอาจส่งผลกระทบต่อภาวะแทรกซ้อนที่เกี่ยวข้องกับการติดเชื้อเอชไอวี การศึกษานี้วิเคราะห์ความหลากหลายทางพันธุกรรมของบริเวณโปรโมเตอร์ของยีน *CCL2* โดยเฉพาะ 378 คู่เบสรอบ rs1024611 ในผู้ติดเชื้อเอชไอวีในประเทศไทย โดยทำการสกัดดีเอ็นเอและเพิ่มปริมาณชิ้นส่วนยีน *CCL2* ที่มี rs1024611 ด้วยวิธี polymerase-chain reaction (PCR) และหาลำดับนิวคลีโอไทด์ด้วยเทคนิค direct sequencing แล้วทำการวิเคราะห์ทางชีวสารสนเทศ โดยมี 140 ตัวอย่างที่มีผลโครมาโตแกรมที่ดี โดยดูความหลากหลายของนิวคลีโอไทด์และความสัมพันธ์ระหว่าง SNP ในบริเวณที่ทำการศึกษาพบ SNP เพิ่มเติมอีกหนึ่งชนิด คือ rs2151345396 ใน SNP rs1024611 ความถี่ของอัลลีล A และ G คือ 0.511 และ 0.489 ตามลำดับ โดยมีความถี่ของจีโนไทป์ AA 0.193, AG 0.636, และ GG 0.171 ขณะที่ SNP rs2151345396 พบความถี่ของอัลลีล G และ C คือ 0.907 และ 0.093 ตามลำดับ การวิเคราะห์ทางสถิติพบว่าไม่มีความสัมพันธ์อย่างมีนัยสำคัญระหว่าง rs1024611 และ rs2151345396 ซึ่งบ่งชี้ถึงความเป็นอิสระของ SNP ทั้งสอง แม้ว่าจะมี linkage disequilibrium ที่สูง ($D' = 0.9985$, $r^2 = 0.0984$) นอกจากนี้ ผลการศึกษานี้พบว่าความหลากหลายทางพันธุกรรมในประชากรที่ศึกษานี้มีความแตกต่างอย่างชัดเจนจากข้อมูลในฐานข้อมูลอ้างอิง (LiveRef SNPs, dbSNP b156v2 and 1000 Genomes Phase 3) จากฐานข้อมูล NCBI งานวิจัยนี้ชี้ให้เห็นถึงความสำคัญของความหลากหลายทางพันธุกรรมเฉพาะในกลุ่มผู้ติดเชื้อเอชไอวี เพื่อประเมินความหลากหลายก่อนที่จะนำมาใช้ในการศึกษาการแสดงออกและการพัฒนาวิธีการวินิจฉัยระดับโมเลกุลของยีน *CCL2* ต่อไป

คำสำคัญ: ยีน *CCL2* โปรโมเตอร์ rs1024611 การติดเชื้อเอชไอวี ความหลากหลายทางพันธุกรรม

*ผู้รับผิดชอบบทความ E-mail address: khaimuk.c@allied.tu.ac.th

รับบทความ: 24 มกราคม 2568

แก้ไขบทความ: 10 พฤษภาคม 2568

รับตีพิมพ์บทความ: 4 มิถุนายน 2568

บทนำ

C-C Motif Chemokine Ligand 2 (CCL2) หรือ monocyte chemotactic protein-1 (MCP-1) เป็นโปรตีนที่มีบทบาทสำคัญในกระบวนการอักเสบและการตอบสนองทางภูมิคุ้มกัน ปัจจุบัน CCL2 เป็นเป้าหมายสำหรับการพัฒนายารักษาโรคเรื้อรังหลายชนิด เช่น ข้ออักเสบรูมาตอยด์ (rheumatoid arthritis) หลอดเลือดแดงแข็ง (atherosclerosis) และมะเร็งชนิดต่าง ๆ นอกจากนี้ยังพบว่า CCL2 มีบทบาทในการเพิ่มจำนวนของเชื้อเอชไอวี (human immunodeficiency virus, HIV) และเกี่ยวข้องกับกลไกการตอบสนองทางภูมิคุ้มกันที่เป็นสาเหตุของการเกิดภาวะแทรกซ้อนในตับ⁽¹⁻⁵⁾ ยีน *CCL2* ตั้งอยู่ที่แขนข้างยาว (q) ของโครโมโซมคู่ที่ 17 (q12) มีขนาดประมาณ 1,986 คู่เบส และประกอบด้วย 3 เอกซอน⁽⁶⁾

งานวิจัยในปัจจุบันมุ่งเน้นที่การศึกษาความหลากหลายทางพันธุกรรมเดี่ยว (single-nucleotide polymorphism หรือ SNP) ได้มีการศึกษาความสัมพันธ์ของ SNP rs1024611 ซึ่งอยู่ที่ตำแหน่งโปรโมเตอร์ของยีน *CCL2* กับโรคต่าง ๆ เช่น โรคจุดภาพชัดเสื่อมในผู้สูงอายุ (age-related macular degeneration หรือ AMD) พบว่าอัลลีล G ของ rs1024611 เกี่ยวข้องกับความเสี่ยงของการพัฒนาโรค AMD ในระยะต้น พร้อมกับระดับ CCL2 ในซีรัมที่สูงขึ้น⁽⁷⁾ ส่วนผู้ป่วยโรคพังผืดในไขกระดูก (primary myelofibrosis) ซึ่งอยู่ในกลุ่มโรคไขกระดูกสร้างเม็ดเลือดมาก (myeloproliferative neoplasm) ชนิดที่ไม่พบโครโมโซมฟีลาเดลเฟีย แต่พบจีโนไทป์ (genotype) rs1024611 GG ได้บ่อยขึ้นอย่างมีนัยสำคัญในผู้ป่วย ซึ่งสัมพันธ์กับระดับการแสดงออกของ CCL2 ที่สูงขึ้น⁽⁸⁾ และโรคไขกระดูกเป็นพังผืดปฐมภูมิ (primary myelofibrosis) พบว่าจีโนไทป์

rs1024611 GG เชื่อมโยงกับความเสี่ยงของโรคที่เพิ่มขึ้น และสัมพันธ์กับอัตราการรอดชีวิตของผู้ป่วยที่ลดลง⁽⁹⁾ นอกจากนี้ ความหลากหลายของ rs1024611 A/G ในโรคข้อเข่าเสื่อมสัมพันธ์กับอาการที่รุนแรงขึ้น โดยมีระดับ CCL2 ในซีรัมสูงขึ้นในพาหะของอัลลีล G⁽¹⁰⁾ สุดท้ายการศึกษาโรคพาร์กินสันในประชากรฮั่นของจีน พบว่าจีโนไทป์ AG และ GG เชื่อมโยงกับอาการรุนแรงขึ้นของผู้ป่วย⁽¹¹⁾

rs1024611 มีผลโดยตรงต่อการเพิ่มขึ้นของการแสดงออกของโปรตีน CCL2 และการตอบสนองทางภูมิคุ้มกัน ทำให้มีความเสี่ยงสูงขึ้นในการเกิดโรคอักเสบต่าง ๆ การศึกษาเกี่ยวกับความหลากหลายทางพันธุกรรมนี้จะช่วยให้เข้าใจเพิ่มเติมเกี่ยวกับบทบาทของ rs1024611 ในภาวะอักเสบ และสามารถนำไปใช้ในการวินิจฉัยและการรักษาได้ในอนาคต นอกจากนี้ rs1024611 ยังสามารถใช้เป็นตัวบ่งชี้ทางชีวภาพในการวินิจฉัยและการพยากรณ์โรคในผู้ป่วยได้⁽⁹⁾

rs1024611 เป็นความหลากหลายทางพันธุกรรมพบที่ตำแหน่ง 5' ของยีน *CCL2* ซึ่งตั้งอยู่บนโครโมโซมคู่ที่ 17 ภายในโปรโมเตอร์ของยีน *CCL2* ซึ่งอยู่ห่างจากตำแหน่งเริ่มต้นของการถอดรหัสประมาณ 2,578 คู่เบส โดยมีความหลากหลายนิวคลีโอไทด์ 2 ชนิด ได้แก่ อะดีนีน (A) และ กวานีนีน (G)^(12,13) โดยจะพบสัดส่วนแปรผันไปในประชากรที่แตกต่างกัน⁽¹⁴⁾ อัลลีล G ของ rs1024511 มีผลให้การแสดงออกของยีน *CCL2* เพิ่มขึ้นอย่างมีนัยสำคัญ เนื่องจากการมีอัลลีล G ทำให้เกิดตำแหน่งการจับที่เหมาะสมสำหรับโปรตีน TALE (three amino acid loop extension) ได้แก่ PREP1/PBX2 complex (pre-B cell leukemia homeobox 1/ pre-B B cell leukemia homeobox 2)^(13,15) ซึ่งส่งผลให้กระตุ้นการทำงานของโปรโมเตอร์ *CCL2* ให้สูงขึ้นโดยเฉพาะในระหว่างมีการอักเสบ นอกจากนี้

ตำแหน่งนิวคลีโอไทด์ 22 คู่เบส ที่ตำแหน่ง 5' เหนืออัลลิล G เมื่อพบร่วมกับอัลลิล G ทำให้เพิ่มการแสดงออกของยีน *CCL2* มากขึ้น ในขณะที่ถ้าตำแหน่งนี้เป็นอัลลิล A จะทำให้ TALE ไม่สามารถเกาะได้และขณะเดียวกันก็จะทำให้เพิ่มแรงของการเกาะ IRF1 ซึ่งเป็นโปรตีนที่กีดการแสดงออกของยีน *CCL2* ทำให้การแสดงออกน้อยลง rs1024611 ยังอยู่ใน linkage disequilibrium อย่างสมบูรณ์กับ rs13900 ซึ่งอยู่ที่บริเวณ 3' untranslated region ของ *CCL2* โดยที่อัลลิล T ของ rs13900 สามารถเพิ่มความเสถียรของ mRNA ได้มากขึ้นจากการจับที่มีประสิทธิภาพกับโปรตีน HuR ส่งผลให้ระดับ *CCL2* เพิ่มขึ้นอีก⁽¹⁶⁾

การศึกษาความสัมพันธ์ของ SNP rs1024611 ในยีน *CCL2* กับภาวะแทรกซ้อนในผู้ติดเชื้อเอชไอวีมีความสำคัญ เนื่องจาก SNP นี้มีบทบาทในการควบคุมการอักเสบและการกระตุ้นระบบภูมิคุ้มกัน อัลลิล G ของ rs1024611 ซึ่งอยู่ในตำแหน่ง cis-regulatory ของยีน *CCL2* มีความเกี่ยวข้องกับการเพิ่มการแสดงออกของยีน *CCL2* ทั้งในสภาพแวดล้อมในหลอดทดลองและจากตัวอย่างผู้ป่วยจริง ซึ่งนำไปสู่การเพิ่มการเคลื่อนย้ายของโมโนไซต์และแมคโครฟาจ⁽¹⁷⁾ ซึ่งเป็นเซลล์สำคัญที่ทำหน้าที่เป็นแหล่งกักเก็บเชื้อเอชไอวี และมีส่วนกระตุ้นการอักเสบเรื้อรังอันเป็นปัจจัยหนึ่งที่เกี่ยวข้องกับความผิดปกติทางระบบประสาทในผู้ติดเชื้อเอชไอวี (HIV-associated neurocognitive disorders หรือ HAND)^(1,17) รวมทั้งปัจจัยที่มีผลกระทบจากการแปรผันทางพันธุกรรมอื่น ๆ ร่วมด้วย ผลลัพธ์นี้ช่วยอธิบายความแตกต่างระหว่างบุคคลในแง่ของการตอบสนองการอักเสบ และแสดงให้เห็นถึงศักยภาพของ rs1024611 ในฐานะตัวบ่งชี้ทางชีวภาพหรือเป้าหมายการรักษาในผู้ติดเชื้อเอชไอวี

การตรวจจับ SNP มีความสำคัญในการศึกษาทางพันธุกรรม ได้แก่ การวินิจฉัยโรคทางพันธุกรรม และการประยุกต์ใช้ในหลายด้าน การออกแบบเครื่องมือทางอณูชีวโมเลกุลที่มีความสามารถในการตรวจจับ SNP อย่างแม่นยำและมีประสิทธิภาพขึ้นอยู่กับการระบุความหลากหลายทางพันธุกรรมที่ใกล้เคียง เช่น polymerase-chain reaction (PCR), real-time PCR⁽¹⁸⁾ หรือการใช้ molecular beacon-based OC sensor⁽¹⁹⁾ เทคนิคเหล่านี้แสดงให้เห็นถึงความสำคัญของการออกแบบเครื่องมือทางโมเลกุลที่ใช้ในการตรวจจับ SNP ที่ต้องการ โดยอาศัยผลของความหลากหลายทางพันธุกรรมที่อยู่ใกล้เคียง ในการเพิ่มความแม่นยำและประสิทธิภาพในการตรวจจับ

การใช้สายดีเอ็นเอต้นแบบจากฐานข้อมูล National Center for Biotechnology Information (NCBI) และตรวจสอบความหลากหลายทางพันธุกรรมบริเวณรอบ rs1024611 จากฐานข้อมูล SNP ในฐานข้อมูล NCBI เพื่อออกแบบและตรวจสอบความหลากหลายทางพันธุกรรมพบว่ามีความหลากหลายทางพันธุกรรมมากกว่า 120 ชนิด⁽¹⁴⁾ ทำให้การออกแบบวิธีการตรวจทางอณูชีววิทยาที่จำเพาะต่อชนิดพันธุกรรมเป็นไปได้ยาก นอกจากนี้ ข้อมูลความถี่ของ SNP ในประชากรเฉพาะพื้นที่ยังมีความแตกต่างกัน ซึ่งปรากฏเฉพาะข้อมูลจากยุโรปและเอเชีย แต่ยังไม่มียข้อมูลสำหรับประเทศไทยหรือเอเชียตะวันออกเฉียงใต้ การศึกษาความหลากหลายของยีนรอบ rs1024611 จึงมีความจำเป็นเพื่อใช้ในการออกแบบวิธีการตรวจทางอณูพันธุศาสตร์ที่มีความแม่นยำมากขึ้นและช่วยแสดงผลกระทบของ SNP อื่นที่อาจส่งผลกระทบต่อการทำงานของยีนที่ศึกษาได้ ในการศึกษาครั้งนี้ได้สำรวจความหลากหลายทางพันธุกรรมบริเวณใกล้เคียงกับ rs1024611 โดยการใช้การวิเคราะห์ลำดับนิวคลีโอไทด์

ที่ได้จาก direct sequencing ด้วยวิธีการทางชีวสารสนเทศศาสตร์ในการช่วยระบุ SNP ที่มีความจำเพาะในประชากรผู้ติดเชื้อเอชไอวีในประเทศไทย สำหรับใช้เป็นข้อมูลพื้นฐานในการออกแบบวิธีการตรวจทางอณูชีววิทยา และการทดสอบการทำงานของยีน *CCL2* ที่จำเพาะอย่างแม่นยำต่อไป

วัสดุและวิธีการ

1. ตัวอย่างที่ใช้ในการศึกษา

ตัวอย่างเลือดที่เหลือจากโครงการศึกษาความชุกและผลกระทบของการติดเชื้อไวรัสตับอักเสบบต่อความรุนแรงของการติดเชื้อไวรัสเอชไอวีและการเกิดภาวะพิษต่อตับในผู้ติดเชื้อชาวไทยที่ได้รับยาต้านไวรัส ดำเนินการในระหว่างปี พ.ศ. 2556-2557 และโครงการศึกษาความชุกและปัจจัยเสี่ยงของโรคตับเรื้อรังและความสัมพันธ์ระหว่างสารหลังกลุ่ม CXCL9-11 ในพลาสมากับความรุนแรงของโรคตับในกลุ่มผู้ติดเชื้อเอชไอวีชาวไทยที่ได้รับยาด้านไวรัส ดำเนินการในปี พ.ศ. 2559 กลุ่มตัวอย่างนี้เป็นผู้ติดเชื้อเอชไอวีที่มีข้อบ่งชี้ในการให้ยาด้านไวรัสตามแนวทางการตรวจวินิจฉัยและการดูแลรักษาผู้ติดเชื้อเอชไอวีและผู้ป่วยเอดส์ระดับชาติ ปี พ.ศ. 2553 โครงการวิจัยข้างต้นมีจำนวน 260 คน ใช้เกณฑ์การคัดเลือกตัวอย่างเลือดเข้าโครงการ คือ เป็นผู้มีอายุ 18 ปีบริบูรณ์ขึ้นไปที่มีประวัติการรักษาการติดเชื้อเอชไอวีและได้รับยาด้านไวรัสในเวชระเบียนของโรงพยาบาล นครนายก ส่วนเกณฑ์การคัดตัวอย่างตรวจออกจากโครงการ ได้แก่ ผู้เข้าร่วมโครงการมีการติดเชื้อฉวยโอกาส รวมทั้งเชื้อวัณโรค หรืออยู่ระหว่างการรับประทานยาสมุนไพร ยาสเตียรอยด์ และแอลกอฮอล์ ตัวอย่างได้มีการเก็บข้อมูล เพศ อายุ ระยะเวลาการได้รับยาด้านไวรัส สูตรยาด้านไวรัสที่ได้รับ และข้อมูลทางห้องปฏิบัติการ ได้แก่ ค่า platelet count ระดับ

เอนไซม์ AST และ ALT และ CD4+ cell count ส่วนการศึกษานี้ทำโดยสุ่มตัวอย่างโดยไม่คำนึงถึงเพศและผลทางห้องปฏิบัติการมาจำนวน 150 ตัวอย่างเพื่อนำมาศึกษาลำดับนิวคลีโอไทด์ของ *CCL2* ทั้งนี้เพื่อนำไปใช้พัฒนาวิธีการตรวจทางอณูชีววิทยาของความหลากหลายทางพันธุกรรมของยีน *CCL2* และจะนำไปหาความสัมพันธ์กับพยาธิสภาพตับในผู้ติดเชื้อเอชไอวีต่อไปในอนาคต

การศึกษานี้ได้รับการรับรองจากคณะกรรมการจริยธรรมการวิจัยในคน มหาวิทยาลัยธรรมศาสตร์ สาขาวิทยาศาสตร์ รหัสโครงการวิจัยที่ 125/2565 ลงวันที่ 8 ตุลาคม พ.ศ. 2565

2. การสกัดดีเอ็นเอและการเพิ่มจำนวนชิ้นส่วนยีน *CCL2* ด้วยวิธี PCR

สกัดดีเอ็นเอจากเลือดในสารกันเลือดแข็งตัวชนิด EDTA จำนวน 150 ตัวอย่าง โดยใช้ชุดสกัด DNA QIAamp® DNA Mini Kit (QIAGEN, Hilden, สหพันธ์สาธารณรัฐเยอรมนี) แล้วนำมาเพิ่มจำนวนด้วยคู่ไพรเมอร์ (F- 5' TCCACTCACTTCTCTCACGC-3', R- 5'-CCAGTCATGCTTCGGGTTTTC-3') ที่ครอบคลุมตำแหน่ง SNP-2578 A/G หรือ rs1024611 ของยีน *CCL2* มีขนาด 473 คู่เบสด้วยน้ำยา HotStarTaq Plus DNA Polymerase kit (QIAGEN, Hilden, สหพันธ์สาธารณรัฐเยอรมนี) ปริมาตรทั้งหมด 50 ไมโครลิตร ประกอบด้วย HotStarTaq Plus DNA Polymerase 1 ไมโครลิตร (5 units) ไพรเมอร์ forward และ reverse ชนิดละ 5 มิลลิโมลาร์ 10X CoralLoad PCR buffer 5 ไมโครลิตร ดีเอ็นเอ 400-800 นาโนกรัม และเติมน้ำกลั่นจนได้ปริมาตรทั้งหมด 50 ไมโครลิตร ทำปฏิกิริยาที่อุณหภูมิ 94 องศาเซลเซียส นาน 5 นาที

ตามด้วยขั้นตอน denaturation 94 องศาเซลเซียส นาน 30 วินาที, annealing 60.5 องศาเซลเซียส นาน 30 วินาที และ extension 72 องศาเซลเซียส นาน 45 วินาที 35 รอบ และ 72 องศาเซลเซียส นาน 10 นาที ทำการตรวจสอบขั้น PCR ด้วย 1% agarose 0.5xTBE ภายใต้แรงเคลื่อนไฟฟ้า คงที่ 100 mV 30 นาที จากนั้นนำเจลไปย้อมด้วย ethidium bromide 1 ไมโครลิตร (10 mg/mL) และนำไปอ่านผลภายใต้แสง UV ด้วยเครื่อง UV transilluminator

3. การวิเคราะห์ลำดับนิวคลีโอไทด์ด้วยเทคนิค direct sequencing และวิธีการทางชีวสารสนเทศศาสตร์

ทำผลผลิตขั้น PCR ให้บริสุทธิ์โดยใช้ ชุดทำผลผลิตดีเอ็นเอให้บริสุทธิ์ QIAquick PCR purification kit (QIAGEN, Hilden, สหพันธ์สาธารณรัฐเยอรมนี) ตามวิธีการที่ระบุในชุดน้ำยา และส่งไปวิเคราะห์เพื่อหาลำดับนิวคลีโอไทด์ที่หน่วยบริการวิเคราะห์หาลำดับนิวคลีโอไทด์ บริษัท MacroGen Inc. กรุงโซล ประเทศเกาหลีใต้ โดยเทคนิค direct sequencing และใช้ 5'-TCTCTCACGCCAGCACTGAC-3' เป็นไพรเมอร์สำหรับหาลำดับนิวคลีโอไทด์

นำผลการวิเคราะห์มาคัดสายนิวคลีโอไทด์ ที่มีโครมาโตแกรมคุณภาพดี ได้แก่ ไม่มีสัญญาณรบกวนและสามารถแยกพีคสัญญาณของแต่ละ นิวคลีโอไทด์ได้ชัดเจนระหว่าง 20 ลำดับคู่เบสแรก ของการวิเคราะห์จนถึง 50 คู่เบสก่อนลำดับสุดท้าย 140 รายจากจำนวนตัวอย่างที่วิเคราะห์ 150 ราย นำมาหาลำดับและชนิดนิวคลีโอไทด์ของ SNP rs1024611 โดยการใช้โปรแกรม Unipro UGENE, Novosibirsk ประเทศรัสเซีย⁽²⁰⁾

4. การวิเคราะห์หาความหลากหลายทางพันธุกรรมของสายนิวคลีโอไทด์

นำไฟล์ข้อมูล ab1 จำนวน 140 ไฟล์ มาระบุตำแหน่งความหลากหลายทางพันธุกรรมต่าง ๆ ภายใน rs1024611 ที่ตำแหน่ง 30 คู่เบสหน้า rs1024611 จนถึง 347 คู่เบสหลัง rs1024611 ด้วยการเทียบนิวคลีโอไทด์ด้วยอัลกอริทึม ClustalX⁽²¹⁾ และต่อดำเนิน Muscle⁽²²⁾ จากโปรแกรม Unipro UGENE⁽²⁰⁾ ซึ่งสามารถตรวจสอบตำแหน่งของความหลากหลายทางพันธุกรรมชนิดต่าง ๆ ได้ เมื่อเทียบกับไฟล์อ้างอิง ได้แก่ นิวคลีโอไทด์จาก ตัวอย่าง R (rs1024611 AA) และเมื่อพบว่าตำแหน่งใดเป็น SNP จึงตรวจสอบต่อดำเนินการใช้ function compare motif จากโปรแกรม Unipro UGENE โดย function นี้จะสามารถจำแนกชนิด SNP เป็น heterozygote หรือ homozygote ของนิวคลีโอไทด์ ชนิดใดได้

ตรวจสอบชนิดและตำแหน่งความหลากหลายทางพันธุกรรมที่ตรวจพบกับฐานข้อมูล single nucleotide polymorphism database (dbSNP) ในฐานข้อมูล NCBI โดยอ้างอิงจากตำแหน่งไฟล์ GRCh38.p14, Accession no. NC_000017 ซึ่งเป็นลำดับนิวคลีโอไทด์ในฐานข้อมูล GENBANK (NCBI) โดยมีการตรวจสอบชนิดและตำแหน่งของการกลายพันธุ์ ที่พบทำได้โดยอ้างอิงฐานข้อมูล dbSNP ในฐานข้อมูล NCBI โดยใช้ตำแหน่งไฟล์ GRCh38.p14, Accession no. NC_000017 ซึ่งเป็นลำดับนิวคลีโอไทด์ในฐานข้อมูล GENBANK ข้อมูลที่ได้จากการตรวจสอบ นำมาเปรียบเทียบกับข้อมูลใน LiveRef SNPs, dbSNP b156v2 (ชุดข้อมูลที่มีการปรับปรุงอย่างต่อเนื่องและอยู่ในเวอร์ชันที่ 56 รุ่นที่ 2) และ 1000 Genomes Phase 3

ที่ถูกจัดเก็บในฐานข้อมูล dbSNP รุ่นที่ 156 อับเดคเเวอร์ชันที่ 2

ตรวจสอบความสัมพันธ์ของตำแหน่งของความหลากหลายทางพันธุกรรมที่พบกับอาการทางคลินิกและความเกี่ยวข้องกับการทำงานของโปรตีนสำคัญจาก dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>), Haploreg (<https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php>) และ RegulomeDB (<https://regulome.stanford.edu/regulome-search/>)

5. การวิเคราะห์ข้อมูลทางสถิติ

ความถี่และการกระจายตัวของอัลลีล จีโนไทป์ของ rs1024611 และความหลากหลายทางพันธุกรรมอื่นที่พบ รวมทั้งความสัมพันธ์ของการกลายพันธุ์อื่นที่ตรวจพบกับ rs1024611 ถูกนำมาวิเคราะห์ด้วยสถิติไคสแควร์ โดยใช้ SPSS version 17.0 และวิเคราะห์ Linkage Disequilibrium (LD) และการวิเคราะห์สมดุลฮาร์ดี-ไวน์เบิร์ก (Hardy-Weinberg equilibrium, HWE) โดยใช้โปรแกรม SNPStats⁽²³⁾ (<https://snpstats.net/start.htm?>) โดยค่าไคสแควร์จะแสดงความสอดคล้องของการพบจีโนไทป์ที่พบใหม่กับชนิดของจีโนไทป์ rs1024611 ส่วน Linkage Disequilibrium (LD) นั้นเป็นการศึกษาว่าตำแหน่ง SNPs ที่อยู่ใกล้กันมีโอกาสถ่ายทอดไปด้วยกันมากหรือน้อยกว่าค่าคาดหวังแสดงได้ด้วยค่า D' , r และ r^2 โดยเมื่อ D' มากกว่า 0.8 แสดงถึงการเชื่อมโยงที่แข็งแกร่งระหว่างสองตำแหน่ง ส่วนค่า r เป็นค่าสัมประสิทธิ์การเชื่อมโยง ซึ่งแสดงถึงความสามารถในการทำนายการเชื่อมโยงนั้น โดยค่า r^2 ที่สูง แสดงถึงการเชื่อมโยงที่แข็งแกร่ง ในกรณีนี้ ค่า r มีค่าน้อยกว่า 0.5 และเป็นลบแสดงถึงความเชื่อมโยงที่อ่อนหรือไม่มีความเชื่อมโยงที่ชัดเจนระหว่างสอง SNP

ผลการวิจัย

1. ผลการตรวจสอบ ความหลากหลายทางพันธุกรรมของนิวคลีโอไทด์

ผลการคัดเลือกโครมาโตแกรมจากการหาลำดับนิวคลีโอไทด์ของตัวอย่างดีเอ็นเอของผู้ป่วยเอชไอวีที่มีคุณภาพดี มีจำนวน 140 ตัวอย่าง จาก 150 ตัวอย่าง ซึ่งเมื่อนำมาหาลำดับและชนิดของนิวคลีโอไทด์ของ rs1024611 ด้วยโปรแกรม Unipro UGENE พบความถี่อัลลีล A และ G ของ rs1024611 เท่ากับ 0.511 และ 0.489 โดยมีความถี่จีโนไทป์ AA, GA, GG เท่ากับ 0.193, 0.636, 0.171 ตามลำดับ (Table 1) ลักษณะโครมาโตแกรมในแต่ละจีโนไทป์ (Fig. 1) เมื่อนำสายนิวคลีโอไทด์มาตรวจสอบเทียบกับ NC_000017.11 จะอยู่ที่ตำแหน่ง 34252769 โดยนำสายนิวคลีโอไทด์ทั้งหมด 140 สาย มาเทียบเคียงลำดับนิวคลีโอไทด์ พบว่าสายนิวคลีโอไทด์ระหว่าง 30 คู่เบสหน้า จนถึง 347 คู่เบสหลัง rs1024611 ซึ่งเท่ากับตำแหน่ง 34252739 และ 34253116 ตามลำดับ พบว่ามีระดับที่เหมือนกันทั้งสาย (consensus) ที่สูงมาก เมื่อใช้ข้อมูลลำดับนิวคลีโอไทด์อ้างอิง (R) ของ rs1024611 (NC_000017.11) พบตำแหน่งการแปรผันทางพันธุกรรมจำนวน 2 ตำแหน่ง ได้แก่ rs1024611 และที่ตำแหน่งห่างจาก rs1024611 จำนวน 170 คู่เบสในทิศทาง 3' (34252939) (Fig. 2) เมื่อนำลำดับนิวคลีโอไทด์จากชิ้น PCR ระหว่างตำแหน่ง 34252739 จนถึง 34253116 มาเทียบหาชนิดความหลากหลายทางพันธุกรรมจากฐานข้อมูล dbSNP พบจำนวน SNP และความหลากหลายทางพันธุกรรมอื่น ได้แก่ การขาดหายของนิวคลีโอไทด์ และการเพิ่มนิวคลีโอไทด์สายสั้น ๆ จำนวนสูงสุด 124 ตำแหน่ง (Table 2) โดยเฉพาะอย่างยิ่งการศึกษา Live RefSNPs dbSNP b156 v2 ในฐานข้อมูล dbSNP พบความหลากหลายจำนวน

มากถึง 124 ตำแหน่ง ในขณะที่การศึกษา 1000 Genomes Phase 3 dbSNP b156 v2 ในฐานะข้อมูล dbSNP พบเพียง 11 ตำแหน่ง จากการศึกษาี้ พบ SNP อื่นในบริเวณรอบ rs1024611 เพียง 1 ตำแหน่งเท่านั้น

Table 1 Genotype frequency of rs1024611 and rs2151345396 in HIV infected individuals.

rs1024611	rs2151345396			Total (Ratio)
	GG	GC	CC	
AA	22	5	0	27 (0.193)
AG	68	21	0	89 (0.636)
GG	24	0	0	24 (0.171)
Total (Ratio)	114 (0.814)	26 (0.186)	0	140



Fig. 1 Chromatogram analysis of rs1024611 genotypes. Representative chromatograms of genotypes AA, AG, and GG at the rs1024611 locus. Mutations are indicated by arrows, with Sample R as the reference and S1, S2, S3, and S4 corresponding to genotypes GG, AG, AG, and AA, respectively. Alignment performed using ClustalX and Muscle algorithms.

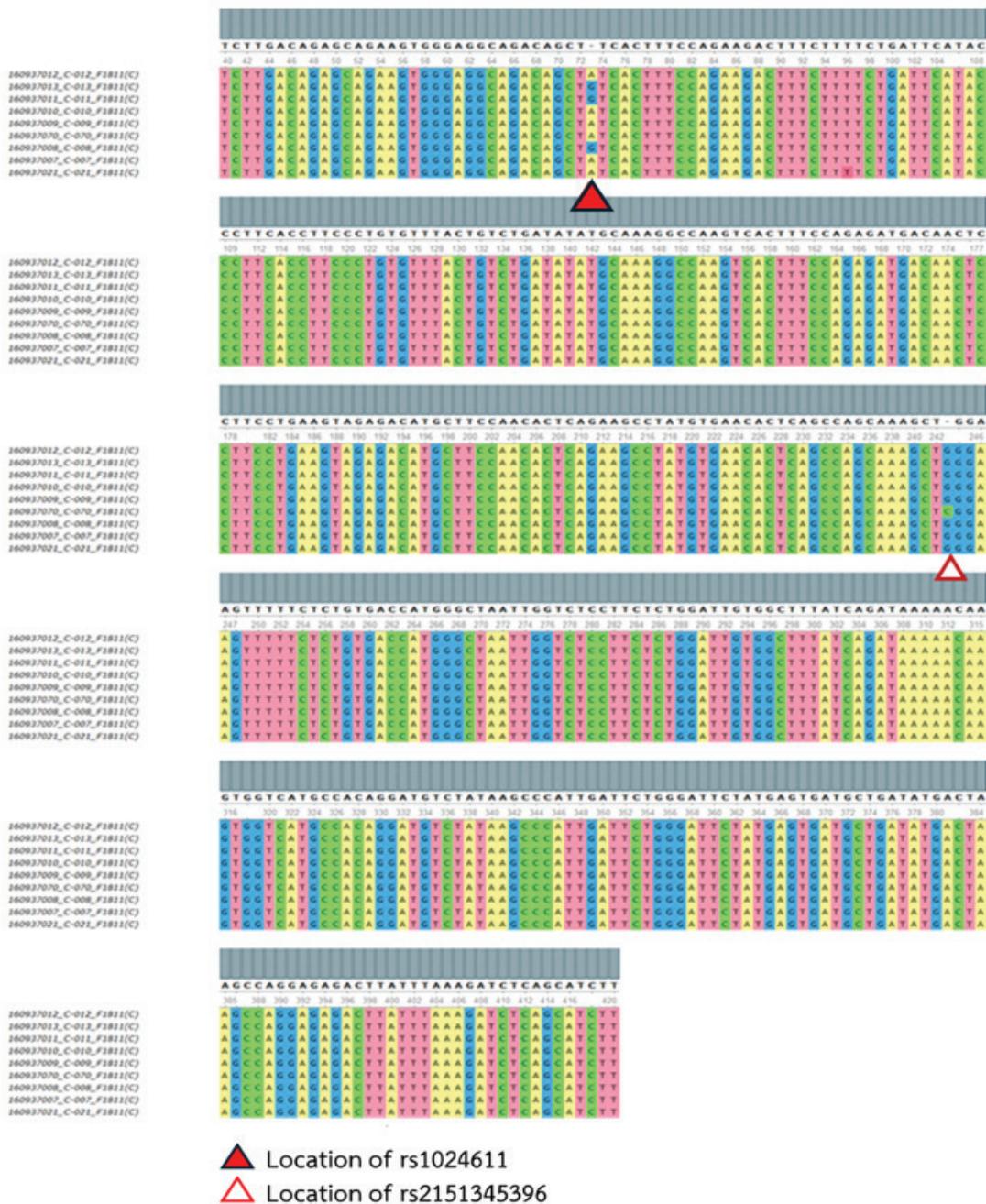


Fig. 2 Nucleotide Variation Analysis Around rs1024611. Positions of nucleotide variations identified within 378 base pairs surrounding rs1024611 from the alignment of 140 samples using the ClustaX and Muscle algorithms within the Unipro UGENE software. The markers ▲ and △ denote the positions of rs1024611 and rs2151345396, respectively.

Table 2 Number of genetic variations identified on GRCh38.p14, Accession Number NC_000017, from dbSNP in NCBI database.

Study in dbSNP (NCBI)	Positions on GRCh38.p14, Accession No. NC_000017		
	34252739 - 34252769	34252769 *	34252770 - 34253116
1000 Genomes Phase 3 dbSNP b156 v2	SNP = 0	SNP = 1	SNP = 10
Live RefSNPs dbSNP b156 v2	SNP = 8	SNP = 1	SNP = 107, D = 8

Remark: D = Other Types of Genetic Variations Excluding SNPs

ตำแหน่ง SNP ที่พบในการศึกษานี้ (34252939) ใน dbSNP สำหรับการศึกษาลiveRef SNPs, dbSNP b156v2 พบ SNP จำนวน 4 ตำแหน่ง ได้แก่ rs2151345395 G/C, rs3917903 G/A/C/T, rs2151345399 G/C และ rs2151345396 G/C/T ที่ตำแหน่งเดียวกัน และถ้ามีการหาจากการศึกษา 1000 Genomes Phase 3 dbSNP b156 v2 พบ 1 SNP ได้แก่ rs2151345396 โดยในการศึกษานี้จะใช้ rs2151345396 เป็นตัวแทนของตำแหน่งที่พบนี้ เนื่องจากมีรายงานความหลากหลายนิวคลีโอไทด์สามชนิด ได้แก่ C, G และ T ซึ่งเพิ่มโอกาสในการเปรียบเทียบความหลากหลายกับตัวอย่างที่ศึกษาในงานวิจัยนี้ รวมทั้งตำแหน่งนี้มีข้อมูลรายงานที่ศึกษาในประชากรมากที่สุดเมื่อเปรียบเทียบกับ SNP อื่นในตำแหน่งเดียวกัน ผลในการศึกษานี้พบอัลลีล 2 ชนิด ได้แก่ C และ G ไม่พบอัลลีล T โดยพบจีโนไทป์ CC และ GG (Fig. 3) เมื่อเทียบตำแหน่งนี้ถึงความสำคัญ

ที่พบจากการศึกษาในฐานข้อมูล dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>), Haploreg (<https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php>) และ RegulomeDB (<https://regulome.stanford.edu/regulome-search/>) ไม่พบรายงานถึงความเกี่ยวข้องทางคลินิกถึงแม้จะเป็นตำแหน่งที่มีการเกาะของ transcription factor c-FOS, c-JUN และ GATA2 บน promoter ของยีน *CCL2* นอกจากนี้รายงานในฐานข้อมูล dbSNP แสดงความถี่ rs2151345396 ของประชากรในโลกนี้เท่ากับ G = 0.940, C = 0.060 เอเชียใต้เท่ากับ G = 0.763 C = 0.237⁽²⁴⁾ ซึ่งยังไม่มีรายงานในประเทศไทยมาก่อน ในส่วนการศึกษานี้ rs2151345396 มีความถี่อัลลีล G เท่ากับ 0.907 และ C เท่ากับ 0.093 ไม่พบอัลลีลชนิดอื่น ซึ่งสอดคล้องในทิศทางเดียวกัน

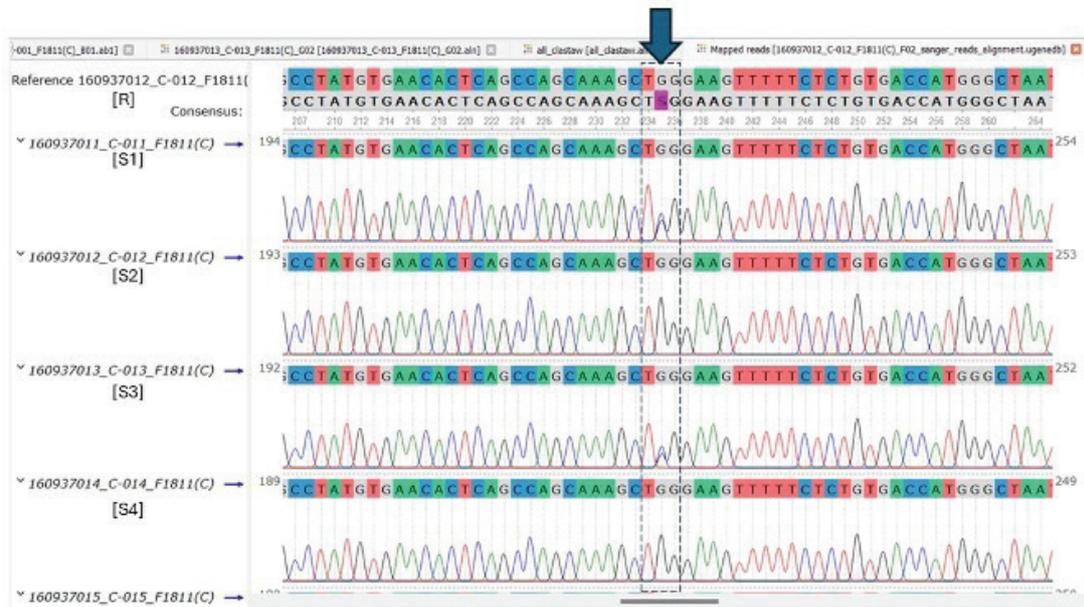


Fig. 3 Chromatogram analysis of rs2151345396 genotypes. Representative chromatograms of genotypes GC, GC at the rs2151345396 locus. Mutations are indicated by arrows, with Sample R as the reference and S1, S2, S3, and S4 corresponding to genotypes GC, GG, GC, and GG, respectively. Alignment performed using ClustalX and Muscle algorithms.

2. การวิเคราะห์สมดุลฮาร์ดี-ไวน์เบิร์ก (Hardy-Weinberg equilibrium: HWE)

การทดสอบสมดุลฮาร์ดี-ไวน์เบิร์กของ SNP rs1024611 จากกลุ่มตัวอย่างทั้งหมด 140 คน ให้ค่าจีโนไทป์ที่พบ ได้แก่ AA 27 ราย AG 89 ราย และ GG 24 ราย ผลการทดสอบแบบ exact test จากโปรแกรม SNPSTATS พบว่า $p = 0.0022$ ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญทางสถิติ ($p < 0.05$) อย่างชัดเจน แสดงให้เห็นว่าการกระจายของจีโนไทป์เบี่ยงเบนจากความคาดหวังภายใต้สมมติฐานของ HWE ซึ่งหมายความว่าอาจมีแรงทางพันธุกรรมหรือสิ่งแวดล้อมบางอย่างที่ส่งผลต่อความถี่ของจีโนไทป์

ในตำแหน่งนี้ ตัวอย่างเช่น การคัดเลือกโดยธรรมชาติ (natural selection) หรือโครงสร้างประชากร (population stratification) หรือจำนวนที่ทำการศึกษาน้อยเกินไป

สำหรับ SNP rs2151345395 พบจีโนไทป์ TT 114 ราย, TC 26 ราย และไม่มีรายใดที่เป็น CC (0 ราย) ค่า $p = 0.61$ ซึ่งสูงกว่าระดับนัยสำคัญ 0.05 หมายความว่า การกระจายของจีโนไทป์ไม่เบี่ยงเบนจากสมดุลฮาร์ดี-ไวน์เบิร์ก อย่างไรก็ตาม การไม่พบจีโนไทป์ CC อาจเกิดจากความถี่ของอัลลีล C ที่ต่ำมากจนไม่ปรากฏในกลุ่มตัวอย่าง ทำให้ไม่สามารถเปรียบเทียบจีโนไทป์ให้ครบทุกแบบได้

3. การศึกษาความสอดคล้องของการพบ ความหลากหลายทางพันธุกรรม rs2151345396 กับ rs1024611

rs1024611 ในการศึกษานี้มีจีโนไทป์ 3 ชนิด ได้แก่ AA, AG, และ GG ส่วน rs2151345396 พบจีโนไทป์ 2 ชนิด ได้แก่ GG และ GC โดยเป็น GG ถึงร้อยละ 90 และไม่พบ CC ผลการศึกษาความสอดคล้องของการพบจีโนไทป์ rs1024611 และจีโนไทป์ rs2151345396 แต่ไม่พบจีโนไทป์ rs2151345396 GC ร่วมกับจีโนไทป์ rs1024611 GG (Table 1) และเมื่อคำนวณทางสถิติถึงความสัมพันธ์ด้วยไคสแควร์ค่าไคสแควร์ 5.595 โดยไม่เกินค่าที่สำคัญ (9.488) ซึ่งบ่งชี้ว่าไม่มีความสัมพันธ์ที่มีนัยสำคัญระหว่าง rs1024611 และ rs2151345396 ผลการศึกษานี้แสดงให้เห็นว่า SNP ทั้งสองมีความเป็นอิสระต่อกันในกลุ่มตัวอย่างนี้ถึงแม้ว่าจะห่างเพียง 170 คู่เบส ส่วนผลการศึกษากการวิเคราะห์ความเชื่อมโยงทางพันธุกรรม (Linkage Disequilibrium) ระหว่าง rs1024611 และ rs2151345396 มีค่า $D' = 0.9985$ และ $r^2 = 0.0984$ และ $r = -0.3137$ ($p < 0.05$) ซึ่งสะท้อนให้เห็นถึงระดับของความเชื่อมโยงทางพันธุกรรมที่สูงมาก ($D' \approx 1$) แม้ว่าค่า r^2 จะอยู่ในระดับต่ำ โดยหมายความว่า SNP ทั้งสองมีแนวโน้มที่จะถูกถ่ายทอดร่วมกันในโครโมโซมเดียวกัน แต่ไม่ได้บ่งชี้ว่ามีความสัมพันธ์กันในเชิงหน้าที่หรือส่งผลต่อฟีโนไทป์โดยตรง ผลลัพธ์นี้บ่งชี้ว่า SNP ทั้งสองอาจอยู่ในแฮพลอไทป์บล็อก (haplotype block) เดียวกัน และอาจได้รับอิทธิพลจากแรงคัดเลือกร่วมกัน อย่างไรก็ตาม การมีค่าความเชื่อมโยงทางพันธุกรรมสูงไม่ได้หมายความว่าจำเป็นต้องมีความสัมพันธ์ทางสถิติในระดับจีโนไทป์หรือฟีโนไทป์เสมอ

วิจารณ์

rs1024611 เป็น SNP ที่มีการศึกษาความสัมพันธ์กับฟีโนไทป์กับโรคที่เกี่ยวข้องโดยเฉพาะกระบวนการอักเสบที่ก่อให้เกิดพยาธิสภาพในผู้ป่วยหลายชนิด⁽¹⁻⁵⁾ ตำแหน่งดังกล่าวอยู่ในขอบเขตของโปรโมเตอร์ซึ่งควบคุมการแสดงออกของยีน *CCL2* และมีรายงานความหลากหลายของยีนในเชื้อชาติที่แตกต่างกัน⁽⁴⁾ มีการศึกษาความสัมพันธ์ของ SNP นี้กับระดับการแสดงออกของยีน *CCL2* และระดับของ *CCL2* ยังพบว่ามีความสัมพันธ์กับความรุนแรงของพยาธิสภาพในผู้ติดเชื้อเอชไอวี⁽⁴⁾ แต่ยังไม่มีการศึกษาความหลากหลายทางพันธุกรรมนี้ในผู้ติดเชื้อเอชไอวีโดยตรง ความหลากหลายของสายนิวคลีโอไทด์ในบริเวณนี้จึงมีความสำคัญสำหรับการศึกษาการทำงานของยีน *CCL2* และอาจมีผลในการออกแบบวิธีการตรวจทางอณูชีววิทยา ในส่วนของการออกแบบวิธีการตรวจทางอณูชีววิทยาเพื่อตรวจว่า SNP ตำแหน่งที่สนใจนั้นเป็นนิวคลีโอไทด์ชนิดใด ผู้ออกแบบจำเป็นต้องมีการนำต้นแบบของสายนิวคลีโอไทด์จากฐานข้อมูล NCBI มาประกอบกับข้อมูลการปรากฏของ SNP นั้น รวมทั้ง SNP อื่นที่อาจปรากฏในบริเวณใกล้เคียงร่วมด้วย และผู้ออกแบบต้องทำการประเมินเบื้องต้นถึงผลกระทบของ SNP ใกล้เคียงต่อความแม่นยำของผลการตรวจที่จะเกิดขึ้น

ในการทดลองนี้เมื่อทำการตรวจสอบจำนวนและตำแหน่ง SNP ในบริเวณที่สนใจจากฐานข้อมูล dbSNP พบการแปรผันทางพันธุกรรมจำนวนมากกว่า 120 ชนิด และความถี่แต่ละ SNP ในฐานข้อมูลนี้แตกต่างกันในประชากรภูมิภาคต่าง ๆ⁽¹⁴⁾ ซึ่งข้อมูลนี้แสดงความชุกในประชากรที่ศึกษาหลากหลายเชื้อชาติและส่วนใหญ่ไม่ได้จำเพาะกับการศึกษาในประเทศไทย

เมื่อผู้วิจัยได้ทำการศึกษาลำดับนิวคลีโอไทด์เบื้องต้น โดยการใช้เทคนิค direct sequencing พบว่า สายนิวคลีโอไทด์ในพื้นที่บริเวณ rs1024611 มีลำดับ นิวคลีโอไทด์เหมือนกันสูงมากซึ่งพบ SNP เพียง ชนิดเดียว ได้แก่ ตำแหน่ง rs2151345396 ซึ่ง อยู่ห่างเพียง 170 คู่เบสในทิศทาง 3' จากตำแหน่ง rs1024611

ตำแหน่ง rs2151345396 ในฐานะข้อมูล dbSNP ปรากฏเป็นชื่อ SNP ที่หลากหลายแต่ไม่มี รายงานที่มีความสัมพันธ์กับพีโนไทป์ได้อย่างชัดเจน ถึงแม้จะพบว่าเป็นตำแหน่งเกาะของ transcription factors ที่สำคัญ ได้แก่ GATA2, c-FOS และ c-JUN โดยความสำคัญของ GATA2 นั้น จำเป็น ต่อการเพิ่มจำนวนและการอยู่รอดของเซลล์ต้นกำเนิด โลหิตในระยะเริ่มต้น และควบคุมการแสดงออกของ ยีนโดยเฉพาะสายเม็ดเลือด⁽²⁵⁾ ส่วน c-FOS, c-JUN เป็นหนึ่งในกลุ่มของ transcription factor AP-1 ซึ่ง มีความสำคัญในการควบคุมการแสดงออกของยีนที่ เกี่ยวข้องกับการเพิ่มจำนวน การเปลี่ยนแปลง และ การตายของเซลล์⁽²⁶⁾

ในรายงาน ฐานข้อมูล dbSNP มีความ หลากหลายของนิวคลีโอไทด์ SNP rs2151345396 น้อยโดยพบ G เกือบร้อยละ 100 โดยพบ C ได้บ้าง แต่ไม่พบ A และ T⁽¹⁴⁾ ส่วนผลจากการศึกษานี้ สัดส่วนที่พบมีความสอดคล้องกันคือพบจีโนไทป์ CG ร้อยละ 1 เมื่อทำการศึกษาความสอดคล้อง ของ rs1024611 และ rs2151345396 จีโนไทป์ พบว่าทั้งสอง SNP เป็นอิสระต่อกันด้วยการทดสอบ ทางสถิติไคสแควร์ และความสัมพันธ์อัลลีลในแง่ของ linkage disequilibrium แสดงความเชื่อมโยงที่ แข็งแรงระหว่าง rs1024611 และ rs2151345396 ซึ่งความเชื่อมโยงนี้อาจมีความผันแปรและไม่ได้บอก ถึงความสัมพันธ์ที่ชัดเจนเสมอไป หรืออาจอยู่ใน

แฮพลอไทป์บล็อกเดียวกันได้ ประกอบกับการทดสอบ สมดุลฮาร์ดี-ไวน์เบิร์ก ของ rs1024611 ไม่พบความ สมดุลในประชากรที่ศึกษา ส่วนการทดสอบ SNP rs2151345396 จะมีความสมดุลในประชากรที่ศึกษา แต่มีข้อจำกัดเนื่องจากข้อมูลไม่พบจีโนไทป์ครบทุก ชนิดทำให้ความสัมพันธ์มีความน่าเชื่อถือน้อยลง หาก มีตัวอย่างที่มากขึ้นอาจทำให้เห็นความสัมพันธ์ได้ อย่างชัดเจน อย่างไรก็ตามข้อสังเกตในการศึกษานี้ ตัวอย่างที่มีจีโนไทป์ ของ rs1024611 GG จะไม่พบ จีโนไทป์ rs2151345396 GC ในผู้ติดเชื้อเอชไอวี ในประเทศไทย

จากการศึกษาที่ผ่านมา พบว่าจีโนไทป์ rs1024611 AG และ GG มีความสัมพันธ์กับความ เสี่ยงที่เพิ่มขึ้นในการเกิดโรค AMD⁽²⁾ นอกจากนี้ rs1024611 ยังมีความสัมพันธ์กับความเสี่ยงในการเกิด เส้นเลือดขอดปฐมภูมิ (primary varicose veins, PVVs)⁽²⁷⁾ rs1024611 ซึ่งเป็น SNP ในยีน CCL2 มีความสัมพันธ์กับระดับที่สูงขึ้นของ CCL2 ในซีรัม และเพิ่มความเสี่ยงในการเกิดโรคต่าง ๆ เช่น โรคทาง ระบบประสาทที่เกี่ยวข้องกับการติดเชื้อ เชื้อเอชไอวี-1 และโรคหลอดเลือดแดงแข็ง (atherosclerosis) ข้อมูลนี้บ่งชี้ว่า rs1024611 อัลลีล G อาจมีผลต่อ ความไม่สมดุลในการแสดงออกของยีน CCL2 และมีบทบาทสำคัญในกระบวนการทางพยาธิวิทยา⁽¹²⁾ ส่วนการศึกษาในผู้ป่วย PMF พบว่าระดับการแสดงออก ของคีโมคัยล์ CCL2 ที่สูงขึ้นนั้นมีความสัมพันธ์กับ อัตราการรอดชีวิตที่ลดลง⁽³⁾

จากการศึกษาตัวอย่างผู้ติดเชื้อเอชไอวีใน ประเทศไทยนี้ ทำให้เกิดข้อสังเกตเพิ่มเติมว่าถ้าหาก มีการศึกษา rs1024611 ร่วมกับ rs2151345396 อาจสามารถแบ่งกลุ่มผู้ป่วยและอาจพบความสัมพันธ์ ที่ชัดเจนมากขึ้นได้ และพบว่าเป็นลักษณะจำเพาะ ของตัวอย่างการศึกษานี้ โดยในภูมิภาคอื่นในโลก

อาจไม่พบความหลากหลายในลักษณะดังกล่าว ซึ่งจะพบได้เมื่อทำการศึกษายานิวคลีโอไทด์ในกลุ่มตัวอย่างที่สนใจเท่านั้น โดยข้อมูลเบื้องต้นนี้จะนำไปใช้ในการออกแบบวิธีการตรวจทางอณูชีววิทยาที่เหมาะสมกับการศึกษาผู้ติดเชื้อเอชไอวีต่อไป

ดังนั้นหากจะพิจารณาสายนิวคลีโอไทด์ที่จะนำมาเป็นต้นแบบเพื่อศึกษา rs1024611 มีความจำเป็นที่ต้องพิจารณาชนิดความหลากหลายทางพันธุกรรมอื่นรอบ ๆ ประกอบด้วย ข้อมูลชนิดของ SNP ที่ได้จากการศึกษา Live RefSNPs dbSNP b156 v2, dbSNP อาจไม่เหมาะสม เพราะให้จำนวน SNP ที่มากเกินไป จากการศึกษานี้พบว่าการใช้ลำดับนิวคลีโอไทด์ในประชากรที่ศึกษาและความหลากหลายที่พบจริงจะให้ความเหมาะสมของข้อมูลมากกว่า ซึ่งอาจจะนำไปสู่การออกแบบวิธีการตรวจความหลากหลายทางพันธุกรรมที่จำเพาะที่ดีกว่าต่อไป อีกประการหนึ่ง หากจะศึกษาความสัมพันธ์ระหว่างความหลากหลายทางพันธุกรรมนี้กับภาวะแทรกซ้อนในผู้ป่วยเอชไอวี กลุ่มนี้มีข้อจำกัดเนื่องจากผู้ป่วยเอชไอวีจะได้รับยาต้านไวรัสตามโปรแกรมการรักษาปกติและพบว่าบางส่วนมีการติดไวรัสตัวกลับซ้ำปี ไปด้วย โดยปัจจัยเหล่านี้จะทำให้อาการแทรกซ้อนตัวรุนแรงขึ้นได้ ข้อจำกัดของงานวิจัยนี้อีกประการหนึ่งคือ ไม่ได้ทำการตรวจ rs1024611 ในกลุ่มคนปกติ จึงไม่อาจทราบว่ามีแตกต่างจากตัวอย่างหรือไม่ ทั้งนี้จึงไม่สามารถนำไปวิเคราะห์เปรียบเทียบเพื่อหาความสัมพันธ์ระหว่าง genotype ของ rs1024611 กับลักษณะทางคลินิกของผู้ป่วยเอชไอวีได้ จำนวนตัวอย่างมีข้อจำกัดโดยต้องมีการรักษาด้วยยาต้านร่วมด้วยตามเกณฑ์จริยธรรมการวิจัยในมนุษย์ ซึ่งมีผลต่อความสามารถในการตรวจจับความแตกต่างที่มีนัยสำคัญทางสถิติ และไม่พบความสมดุลของประชากรจากการทดสอบสมมูลฮาร์ดี-ไวน์เบิร์ก

สรุป

ความหลากหลายทางพันธุกรรมของยีน *CCL2* ที่ตำแหน่งรอบบริเวณ rs1024611 ซึ่งเป็นตำแหน่งโปรโมเตอร์ของยีน ในผู้ติดเชื้อเอชไอวีในประเทศไทย ขนาด 378 คู่เบส ศึกษาโดยการเพิ่มจำนวนชิ้น PCR แล้วนำไปวิเคราะห์ลำดับนิวคลีโอไทด์ด้วยเทคนิค direct sequencing ร่วมกับการวิเคราะห์ทางชีวสารสนเทศศาสตร์ พบ SNP 1 ตำแหน่ง ได้แก่ rs2151345396 ซึ่งมีความแตกต่างอย่างมากเมื่อเทียบกับ SNP ที่ได้จากฐานข้อมูล dbSNP ซึ่งมีมากถึง 116 ตำแหน่ง และเมื่อรวมความหลากหลายทางพันธุกรรมอื่นจะมีความหลากหลายถึง 124 ตำแหน่ง rs2151345396 พบอัลลีล G และ C โดยไม่พบจีโนไทป์ CC ซึ่งสอดคล้องกับรายงานความถี่ในประชากรโลก โดยที่ไม่เคยมีรายงานการตรวจสอบในประเทศไทยและไม่พบการศึกษาที่เกี่ยวข้องกับอาการทางคลินิกหรือการแสดงออกของยีน *CCL2* จากการศึกษานี้พบความสัมพันธ์ระหว่าง SNP ระหว่าง rs1024611 และ rs2151345396 เป็นอิสระต่อกัน

การวิจัยนี้เน้นย้ำถึงความสำคัญของการศึกษาความหลากหลายทางพันธุกรรมในกลุ่มศึกษาในภูมิภาคจำเพาะ เพื่อสร้างข้อมูลที่จำเพาะเจาะจงและใช้เป็นพื้นฐานในการพัฒนาวิธีการวินิจฉัยการศึกษายีน *CCL2* ในผู้ติดเชื้อเอชไอวีในประเทศไทยที่มีประสิทธิภาพยิ่งขึ้นในอนาคต

กิตติกรรมประกาศ

ขอขอบคุณเจ้าหน้าที่ห้องปฏิบัติการเทคนิคการแพทย์ คณะสหเวชศาสตร์ มหาวิทยาลัยธรรมศาสตร์ และนักวิทยาศาสตร์กลุ่มวิจัย Thammasat Research Unit in diagnostic molecular biology of chronic diseases related to cancer ที่สนับสนุนการดำเนินงานจนทำให้งานวิจัยนี้สำเร็จลงด้วยดี

เอกสารอ้างอิง

1. Gonzalez E, Rovin BH, Sen L, *et al.* HIV-1 infection and AIDS dementia are influenced by a mutant MCP-1 allele linked to increased monocyte infiltration of tissues and MCP-1 levels. *Proc Natl Acad Sci U S A* 2002; 99: 13795-800.
2. Sharma P, Patel S, Hasan N. Association of *CCL2* gene polymorphism with age-related macular degeneration. *Gene* 2018; 642: 251-6.
3. Masselli E, Venegoni A, Fisco G, *et al.* Chemokine *CCL2* expression is related to survival in patients with primary myelofibrosis. *Blood Adv* 2021; 5: 1290-301.
4. Wahid A, Ansari R, Schmidt E, *et al.* Immuno-pathomechanism of liver fibrosis: targeting chemokine *CCL2*-mediated HIV nexus. *J Transl Med* 2014; 12: 341. doi:10.1186/s12967-014-0341-8.
5. Da LS, Zhang Y, Zhang S, *et al.* Association between MCP-1 -2518A/G polymorphism and cancer risk: evidence from 19 case-control studies. *PLoS One* 2013; 8: e82855. doi:10.1371/journal.pone.0082855.
6. National Center for Biotechnology Information. *CCL2* C-C motif chemokine ligand 2 [Homo sapiens (human)] [Internet]. 2024 [cited 2024 Jun 21]. Available from: <https://www.ncbi.nlm.nih.gov/gene?Db=gene&Cmd=Details-Search&Term=6347>.
7. Gudauskiene G, Vilkeviciute A, Gedvilaite G, *et al.* *CCL2*, *CCR2* gene variants and *CCL2*, *CCR2* serum levels association with age-related macular degeneration. *Life (Basel)* 2022; 12: 1038. doi:10.3390/life12071038.
8. Hodeib H, Abd El Hai D, Tawfik M, *et al.* *CCL2* rs1024611 gene polymorphism in Philadelphia-negative myeloproliferative neoplasms. *Genes (Basel)* 2022; 13: 492. doi:10.3390/genes13030492.
9. Masselli E, Carubbi C, Pozzi G, *et al.* Impact of the rs1024611 polymorphism of *CCL2* on the pathophysiology and outcome of primary myelofibrosis. *Cancers (Basel)* 2021; 13: 2552. doi:10.3390/cancers13112552.
10. Sharma AC, Srivastava RN, Srivastava SR, *et al.* AB1291 association of *CCL2* gene polymorphisms and their serum levels with susceptibility to knee osteoarthritis. *Ann Rheum Dis* 2019. doi:10.1136/annrheumdis-2019-eular.6642.
11. Shen R, Lin S, He L, *et al.* Association of two polymorphisms in *CCL2* with Parkinson's disease: a case-control study. *Front Neurol* 2019; 10: 35. doi:10.3389/fneur.2019.00035.
12. Pham MHT, Bonello GB, Castiblanco J, *et al.* The rs1024611 regulatory region polymorphism is associated with *CCL2* allelic expression imbalance. *PLoS One* 2012; 7: e49498. doi:10.1371/journal.pone.0049498.

13. Wright EK, Page SH, Barber SA, *et al.* Prep1/Pbx2 complexes regulate *CCL2* expression through the -2578 guanine polymorphism. *Genes Immun.* 2008; 9: 419-30. doi:10.1038/gene.2008.33.
14. National Center for Biotechnology Information. rs1024611 [Internet]. dbSNP Short Genetic Variations. 2022 [cited 2024 Jun 21]. Available from: <https://www.ncbi.nlm.nih.gov/snp/rs1024611>.
15. Page SH, Wright EK, Gama L, *et al.* Regulation of *CCL2* expression by an upstream TALE homeodomain protein-binding site that synergizes with the site created by the A-2578G SNP. *PLoS One* 2011; 6: e22052. doi:10.1371/journal.pone.0022052.
16. Akhtar F, Hernández Ruiz J, Liu YG, *et al.* Functional characterization of the disease-associated *CCL2* rs1024611G-rs13900T haplotype: the role of the RNA-binding protein HuR. *bioRxiv* 2023. doi:10.1101/2023.10.31.564937.
17. Oven J, Coll B, Tous M, *et al.* The influence of HIV infection on the correlation between plasma concentrations of monocyte chemoattractant protein-1 and carotid atherosclerosis. *Clin Chim Acta* 2006; 368: 114-9.
18. Zhao F, Maren NA, Kosentka PZ, *et al.* An optimized protocol for stepwise optimization of real-time RT-PCR analysis. *Hortic Res* 2021; 8: 161. doi:10.1038/s41438-021-00616-w.
19. Mueller BL, Liberman MJ, Kolpashchikov DM. OWL2: a molecular beacon-based hybridization sensor for highly selective SNV detection in folded nucleic acids. *Nanoscale* 2023; 15: 158-64. doi:10.1039/d2nr05590b.
20. Okonechnikov K, Golosova O, Fursov M, *et al.* Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* 2012; 28: 1166-7. doi:10.1093/bioinformatics/bts091.
21. Thompson JD, Gibson TJ, Higgins DG. Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics* 2003; Chapter 2: Unit 2.3. doi:10.1002/0471250953.bi0203s00.
22. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004; 32: 1792-7. doi:10.1093/nar/gkh340.
23. Solé X, Guinó E, Valls J, *et al.* SNPStats: a web tool for the analysis of association studies. *Bioinformatics* 2006; 22: 1928-9. doi:10.1093/bioinformatics/btl268.
24. National Center for Biotechnology Information. rs2151345396 [Internet]. dbSNP Short Genetic Variations. 2022 [cited 2024 Jun 21]. Available from: <https://www.ncbi.nlm.nih.gov/snp/rs2151345396>
25. Gao J, Chen Y, Peterson L. GATA family transcriptional factors: emerging suspects in hematologic disorders. *Exp Hematol Oncol* 2015; 4: 24. doi:10.1186/s40164-015-0024-z.

26. Turner R, Tjian R. Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFos-cJun heterodimers. *Science* 1989; 243: 1689-94. doi:10.1126/science.2494701.
27. Shadrina AS, Kashtalap VV, Slominsky PA. Polymorphisms in the CCL2 gene are associated with the risk of primary varicose veins. *Mol Biol (Mosk)* 2017; 36: 445-53.