



การวิเคราะห์ความสัมพันธ์ในครอบครัว จากสลิป 44 ตำแหน่ง ในคนไทย Parents-child analysis with 44 single nucleotide polymorphisms (SNPs) in Thai individuals

สุนิษา โอบอ้อม*
Sunisa Aobaom

■ บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์ความสัมพันธ์ในครอบครัว (parent-child) จากสลิป (SNPs) ทั้งหมด 44 ตำแหน่ง จากนั้นคำนวณหาค่าดัชนีความเป็นบิดา (Paternity index, PI) ในการวิจัยครั้งนี้แบ่งตัวอย่างออกเป็น 2 กลุ่มตัวอย่าง ได้แก่ กลุ่มที่มีความสัมพันธ์พ่อ-แม่-ลูกที่แท้จริง (True parent-child group) และกลุ่มที่ไม่มีความสัมพันธ์กัน (unrelated group) จำนวนกลุ่มละ 38 ครอบครัว จากผลการวิจัยพบว่าในกลุ่มมีความสัมพันธ์พ่อ-แม่-ลูกที่แท้จริงนั้น มีการใช้ SNPs ร่วมกันทั้งหมด 44 ตำแหน่ง ทั้งหมด 38 ครอบครัว คิดเป็นร้อยละ 100 และกลุ่มที่ไม่มีความสัมพันธ์กัน มีตำแหน่งที่ไม่ตรงกันตั้งแต่ 2 ถึง 13 ตำแหน่ง โดยในกลุ่มนี้ พบ 10 ครอบครัวที่มีตำแหน่งที่ไม่ตรงกันถึง 4 ตำแหน่ง ซึ่งเป็นจำนวนตำแหน่งที่ไม่ตรงกันมากที่สุดคิดเป็นร้อยละ 26.32 ค่า PI ในกลุ่มที่มีความสัมพันธ์พ่อ-แม่-ลูกที่แท้จริง มีค่าอยู่ระหว่าง 83.9371 ถึง 1.0120×10^7 และกลุ่มที่ไม่มีความสัมพันธ์ มีค่าอยู่ระหว่าง 7.0750 ถึง 5.0065×10^4 เมื่อสังเกตค่ากระจายตัวของ PI พบว่า ค่าต่ำสุดของกลุ่มที่มีความสัมพันธ์พ่อ-แม่-ลูกที่แท้จริงติดกับค่าสูงสุดของกลุ่มที่ไม่มีความสัมพันธ์กันที่มีค่าระหว่าง 83.9371 ถึง 5.0065×10^4 ซึ่งในช่วงนี้จะเป็นช่วงที่เรียกว่า “gray zone” กล่าวคือ ไม่สามารถระบุความสัมพันธ์พ่อ-แม่-ลูกที่แท้จริงได้

* อาจารย์ ภาควิชาเทคนิคการแพทย์ คณะสหเวชศาสตร์ มหาวิทยาลัยธรรมศาสตร์

Lecturer, Department of Medical Technology Faculty of Allied Health Sciences Thammasat University

Received: July 12, 2022 Revised: September 16, 2022 Accepted: October 5, 2022

ซึ่งค่า PI ของกลุ่มที่ไม่มีความสัมพันธ์ เกิดจากคำนวณโดยการตัดตำแหน่งที่ไม่ตรงกันออก และคำนวณค่า PI ดังกล่าว งานวิจัยนี้แสดงให้เห็นว่าพบความแตกต่างจากการใช้ SNPs 44 ตำแหน่ง ระหว่างกลุ่มที่มีความสัมพันธ์ พ่อแม่-ลูกที่แท้จริง (True parent-child group) และกลุ่มที่ไม่มีความสัมพันธ์กัน (unrelated group) สามารถช่วยทำนายความสัมพันธ์ของมนุษย์และนำไปประยุกต์ใช้กับความสัมพันธ์ในเครือญาติอื่น ๆ ในอนาคตได้

คำสำคัญ: พ่อแม่ลูก, สนิปส์, ค่าดัชนีความเป็นบิดา

■ Abstract

The aim of research is to determine the parent-child relationship based upon 44 SNPs markers using statistic values of paternity index (PI). Samples of this study were in two groups: true parent-child (n = 38 families) and unrelated groups (n = 38 families). The results showed that the true parent-child group shared 44 loci (38 families: 100%), and the unrelated groups had loci mismatch between 2 to 13 loci. The four mismatch loci were the highest number of families (10 families: 26.32%). The maximum and minimum of PI values were 1.0120×10^7 and 83.9371 were among the true parent-child groups. The unrelated group's maximum and minimum PI values were 5.0065×10^4 and 7.0750, respectively. The lower limit

of the distribution of the true parent-child intersected with the upper limit of the distribution of the unrelated between the values of 83.9371 to 5.0065×10^4 . In this range of values, the distribution of the PI values could not be used to determine the relationship and therefore was recognized as a "gray zone", whereas the non-relationship PI is from the calculation of mismatch elimination. This research showed that the difference between true parent-child and unrelated groups using 44 SNPs could help to predict human relationships and can be further applied in other relationships.

Keyword: SNPs, Parents-child, Paternity index

■ บทนำ/ความเป็นมาของปัญหา

งานทางด้านนิติพันธุศาสตร์เป็นส่วนหนึ่งของการนำมาช่วยในการตัดสินในกระบวนการยุติธรรม เช่น การตรวจเพื่อยืนยันบุคคลที่กระทำ ความผิด หรือการพิสูจน์บุคคลในกรณีที่ถูกสงสัย เป็นต้น ซึ่งรูปแบบในการคลี่คลายคดีที่เป็นหลักฐานชิ้นสำคัญและมีความถูกต้องแม่นยำที่สุด ทั้งนี้การตรวจความสัมพันธ์ทางสายเลือดก็เป็นอีกหนึ่งรูปแบบที่มีส่วนช่วยในกระบวนการยุติธรรมเช่นเดียวกัน ด้วยหลักการที่ว่ามนุษย์ทุกคนมีสารพันธุกรรมที่แตกต่างกัน ยกเว้นฝาแฝดที่เกิดจากไข่ใบเดียวกันที่จะมีสารพันธุกรรมที่เหมือนกันทุกประการ ซึ่งแต่ละบุคคลจะได้รับการถ่ายทอดมาจากพ่อและแม่อย่างละครึ่ง

โดยมีรหัสดีเอ็นเอมากกว่าสามพันล้านรหัส แต่ไม่จำเป็นจะต้องตรวจสอบทุกรหัสหากแต่เลือกเพียงบางตำแหน่งเท่านั้นในการตรวจสอบ ซึ่งในการพิสูจน์ลักษณะทางสายเลือดในรูปแบบ พ่อ-แม่-ลูก จะต้องมียาลพิมพ์ดีเอ็นเอเหมือนกันทุกตำแหน่ง หากไม่ตรงกันเพียง 1 ตำแหน่งเท่านั้น ก็แสดงถึงบุคคลเหล่านั้นอาจไม่มีความสัมพันธ์กันแบบพ่อ-แม่-ลูกจริง โดยในปัจจุบันมีการนำเทคโนโลยีทางพันธุศาสตร์มาตรวจสอบ เช่นเครื่องหมายพันธุกรรมบริเวณดีเอ็นเอลำดับซ้ำ (Short Tandem Repeat : STRs) ทั้ง 16 หรือ 24 ตำแหน่ง ซึ่งต้องใช้เวลานานและมีราคาที่สูง จึงมีการนำ Single Nucleotide Polymorphisms (SNPs) มาใช้ในรูปแบบของการศึกษาในกลุ่มของการตรวจพิสูจน์เอกลักษณ์บุคคล ของสองวัตถุประสงค์ว่าเป็นคนคนเดียวกันหรือไม่ หากแต่ยังไม่มีการศึกษาลงไปถึงระดับของความสัมพันธ์ของมนุษย์ว่าสามารถใช้วิธีการดังกล่าวตรวจสอบได้หรือไม่ ด้วยเหตุผลทั้งหมดที่กล่าวมาข้างต้นจึงทำให้เกิดงานวิจัยนี้ขึ้น

การตรวจลายพิมพ์ดีเอ็นเอในปัจจุบันที่เป็นที่ยอมรับนั้น ใช้วิธีการทดสอบด้วย STRs ซึ่งในแต่ความสัมพันธ์มีการคำนวณค่าอัตราส่วนความน่าจะเป็น หรือ Likelihood Ratio (LR) ซึ่งจะมีประโยชน์ในการกำหนด และคัดแยกความสัมพันธ์ของมนุษย์ ไม่ว่าจะเป็น พ่อ-แม่-ลูก (Parent-Child) หรือความสัมพันธ์พี่น้อง (Kinship relationship) (Buckleton, Bright, & Taylor, 2016; Butler, 2005; Elkins, 2013) ซึ่งความสัมพันธ์แบบพ่อ-แม่-ลูก (parents-child) เป็นลักษณะความสัมพันธ์ของมนุษย์หรือสิ่งมีชีวิตที่ถูกควบคุมโดยดีเอ็นเอที่มีการถ่ายทอดจากพ่อและแม่อย่างละครึ่ง ยกเว้นในแฝดไข่

ใบเดียวกันที่จะมีลักษณะการถ่ายทอดพันธุกรรมเหมือนกันทุกประการ ปัจจุบันการตรวจความสัมพันธ์แบบพ่อ-แม่-ลูก ใช้วิธีการตรวจสอบโดยใช้วิธีการทดสอบด้วย STRs ยกตัวอย่างเช่น เมื่อครอบครัวมี genotype แบบ 11,14 (พ่อ) และ 8,12 (แม่) จะมีโอกาสถ่ายทอดดีเอ็นเอไปสู่ลูกคือ 11,8 11,12 14,8 หรือ 14,12 เป็นต้น (Butler, 2005)

SNPs เป็นการแปรผันของลำดับดีเอ็นเอชนิดหนึ่งซึ่งมีการเปลี่ยนแปลงของนิวคลีโอไทด์หนึ่งตัวในจีโนม ทำให้แตกต่างจากจีโนมของสิ่งมีชีวิตอื่นในสปีชีส์เดียวกัน หรือต่างจากโครโมโซมอีกแท่งหนึ่งในสิ่งมีชีวิตเดียวกัน ในมนุษย์ SNPs เกิดขึ้นครั้งเดียวโดยเฉลี่ยในทุก 300 นิวคลีโอไทด์ ซึ่งหมายความว่า จะมี SNPs ประมาณ 10 ล้านในจีโนมมนุษย์ โดยทั่วไปแล้วรูปแบบนี้จะพบได้ในดีเอ็นเอระหว่างยีน ซึ่งสามารถเป็นเครื่องหมายทางชีวภาพสำหรับการตรวจหาโรค นอกจากนี้ยังสามารถใช้เพื่อตรวจสอบการถ่ายทอดดีเอ็นเอภายในกลุ่มครอบครัวได้ (Borsting et al., 2008)

งานวิจัยหนึ่งกล่าวว่า SNPs มีตำแหน่งที่เหมาะสมที่เหมาะสมต่อการระบุลักษณะตัวบุคคลอันดับแรกคือ มีอัตราการกลายพันธุ์ที่ต่ำกว่า STRs (Lou et al., 2011) ซึ่งเหล่านี้จะเหมาะกับการตรวจสอบในกลุ่มนักโทษ และการตรวจคนเข้าเมือง สองการวิเคราะห์ SNPs สามารถทำได้เมื่อมีการเพิ่มปริมาณดีเอ็นเอ (PCR) สายสั้น ๆ ในตัวอย่างที่เสื่อมสภาพ ซึ่งเหมาะกับการสืบค้นทางมานุษยวิทยา และอาชญากรรม และสุดท้ายคือ SNPs เป็น binary polymorphisms ซึ่งจะเป็นวิธีที่ง่ายในการตรวจสอบ เนื่องจาก

มีค่าประมาณความถี่ของแอลลีลที่แม่นยำ ซึ่งจำเป็นต่อการตีความข้อมูลทางนิติเวชได้อย่างถูกต้อง และทำให้การวิเคราะห์โดยใช้ความเข้มข้นของดีเอ็นเอในตัวอย่างที่น้อยลงเมื่อเทียบกับการประมาณความถี่แอลลีลของ STRs และเมื่อผู้วิจัยทำการตรวจเอกลักษณ์บุคคลด้วย SNPs ทั้งหมด 52 ตำแหน่ง โดยทำการเปรียบเทียบผลระหว่างการวิเคราะห์ด้วย SNPs และ STRs ในดีเอ็นเอที่เสื่อมสภาพนั้น พบว่า SNPs สามารถให้ผลการวิเคราะห์ได้ครบทุกตำแหน่ง ในขณะที่การวิเคราะห์ด้วย STRs สามารถให้ผลได้เพียงแค่บางตำแหน่งเท่านั้น ทำให้เห็นว่าการใช้ SNPs สามารถระบุเอกลักษณ์ส่วนบุคคลได้ โดยมีความน่าจะเป็นที่แต่ละบุคคลจะมีตำแหน่ง SNPs ซ้ำกันอยู่ที่ 5.00×10^{-19} (Borsting et al., 2008; Sanchez et al., 2006) เช่นเดียวกับการศึกษาพัฒนาการตรวจเอกลักษณ์บุคคลด้วย SNPs ทั้งหมด 30 ตำแหน่ง ในประชากรเกาหลีทั้งหมด 8,842 คน ที่ไม่มีความสัมพันธ์กัน พบความน่าจะเป็นที่แต่ละบุคคลจะมีตำแหน่ง SNPs ซ้ำกันอยู่ที่ 4.83×10^{-13} และเมื่อลดโอกาสการที่จะมีตำแหน่ง SNPs ซ้ำอีก 10 เท่า เมื่อเพิ่มตำแหน่ง SNPs ทุกๆ 5 ตำแหน่ง (Kim et al., 2010) และเมื่อไม่นานมานี้ก็ยังพบการศึกษาพัฒนาการตรวจเอกลักษณ์บุคคลด้วย SNPs ทั้งหมด 50 ตำแหน่ง ในกลุ่มคนที่ไม่มีความสัมพันธ์กันและกลุ่มที่มีความสัมพันธ์กันแบบพี่น้องที่มีพ่อและแม่เดียวกัน (Full-sibling) โดยพบความน่าจะเป็นที่แต่ละบุคคลจะมีตำแหน่ง SNPs ซ้ำกันอยู่ที่ 6.90×10^{-20} ในกลุ่มที่ไม่มีความสัมพันธ์กันเลย ส่วนในกลุ่มที่มีความสัมพันธ์กันแบบ พี่น้อง

พ่อแม่เดียวกัน (full-sibling) มีค่าความน่าจะเป็นอยู่ที่ 1.20×10^{-10} (Yousefi et al., 2018) นอกจากนี้ยังมีรายงานการทดสอบความเป็นเครือญาติแบบคู่ด้วยการรวมกันของ STRs และ SNPs พบว่า 40 STRs ที่มี 91 SNPs พบว่า สามารถแยกแยะความสัมพันธ์แบบ full-sibling ออกจากพี่น้องที่มีพ่อหรือแม่เดียวกัน (Half-sibling) ได้ แต่ไม่สามารถแยกแยะญาติที่อยู่ในเครือญาติเดียวกันอื่นๆ เช่น ลูกพี่ลูกน้องได้ (Zhang et al., 2020) เหล่านี้ยิ่งเพิ่มความเชื่อมั่นในการใช้ SNPs ในการตรวจเอกลักษณ์บุคคลให้มากยิ่งขึ้น และเมื่อกลุ่มคนมีความสัมพันธ์กันโอกาสที่จะเกิดการใช้ SNPs ร่วมกันยิ่งมากขึ้น

งานวิจัยนี้ทำการศึกษาการแปรจำนวนเบสบนโครโมโซมในความสัมพันธ์พ่อ-แม่-ลูก (Parents-child) ที่ตำแหน่ง Single Nucleotide Polymorphisms (SNPs) ทั้งหมด 44 ตำแหน่ง และวิเคราะห์ด้วยเครื่อง MassARRAY® System (Agena Bioscience, USA.) เพื่อสามารถนำมาปรับใช้ในการตัดสินความสัมพันธ์ทางสายเลือดในรูปแบบพ่อ-แม่-ลูก ในคนไทยได้

■ วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาจำนวนเบสบนโครโมโซมที่ได้จากการตรวจ SNPs 44 ตำแหน่งที่มีรูปแบบเหมือนกันของกลุ่มที่มีความสัมพันธ์พ่อ-แม่-ลูกที่แท้จริง
2. เพื่อศึกษาความสัมพันธ์ของเบสบนโครโมโซมจากการตรวจด้วย SNPs ที่จำแนกความเป็นพ่อ-แม่-ลูกที่แท้จริงและไม่แท้จริง

■ วิธีการดำเนินการวิจัย

1. การเก็บตัวอย่างและข้อมูล

กลุ่มตัวอย่าง : คนไทยกลุ่มที่มีความสัมพันธ์ พ่อ-แม่-ลูกอย่างแท้จริงที่อาศัยอยู่ในประเทศไทย

เกณฑ์การคัดเข้า: อาสาสมัครคนไทย เพศชายหรือเพศหญิง ที่มีสุขภาพดี ที่มีอายุตั้งแต่ 10 ปีขึ้นไป ไม่เป็นโรคธาลัสซีเมีย และไม่โรคทางพันธุกรรมที่ทำให้เกิดความผิดปกติของโปรตีนฮีโมโกลบินอื่น ๆ และมีการตรวจสอบสัมภาษณ์ว่ากลุ่มคนดังกล่าวมีความสัมพันธ์ทางสายเลือดจริง โดยข้อมูลทั้งหมดจะเก็บเป็นความลับ

เกณฑ์การคัดออก: อาสาสมัครที่เป็นโรคธาลัสซีเมีย และโรคทางพันธุกรรมที่ทำให้เกิดความผิดปกติของโปรตีนฮีโมโกลบินอื่น ๆ และผู้วิจัยเห็นว่าในกลุ่มที่มีความเสี่ยงว่าจะไม่มีความสัมพันธ์กันทางสายเลือดอย่างแท้จริง

เกณฑ์ยุติการวิจัย: จะทำการยุติเมื่อมีการไม่ร่วมมือในขั้นตอนการวิจัยในเด็กที่มีอายุต่ำกว่า 18 ปี

การถอนอาสาสมัครออกจากกรวิจัย: จะทำการถอดถอนอาสาสมัครทันทีเมื่อเกิดอันตรายในอาสาสมัครอย่าง เช่น การเกิดแผลในปากหรืออาเจียน ที่เกิดขึ้นระหว่างการเก็บเยื่อกระดูกพุงแก้ม

ระยะเวลาและจำนวนครั้งที่นัดอาสาสมัคร: อาสาสมัครจะถูกสัมภาษณ์เป็นระยะเวลา 5 นาที และเก็บตัวอย่าง เป็นระยะเวลา 25 นาที เพียงครั้งเดียว ตลอดงานวิจัย และไม่มีนัดฟังผลการตรวจสอบใด ๆ

ขนาดตัวอย่าง : จำนวน 38 ครอบครัว รวม 99 คน โดยสุ่มจากสูตรการคำนวณของ

Yamane (1973) (จำนวนประชากร (N) อ้างอิงจากจำนวนผู้ที่มีสัญชาติไทย และมีชื่ออยู่ในทะเบียนบ้าน จากระบบสถิติทางการทะเบียน ปี พ.ศ. 2560 : 64,627,515 คน)

ตัวอย่างที่เก็บ : สำลีพันปลายไม้เก็บบริเวณเยื่อกระดูกพุงแก้ม (Buccal swab) คนละ 3 ก้าน

การสัมภาษณ์ : สอบถามชื่อ-สกุล อายุ เพศ ภูมิลำเนา โดยจะแทนความสัมพันธ์ตามรหัสต่อไปนี้ พ่อ = FOOX แม่ = MOOX และลูก = COOX โดย X ลำดับตัวเลขตั้งแต่ 1-38 ครอบครัว

2. การสกัดและเพิ่มปริมาณดีเอ็นเอ

นำสำลีพันปลายไม้ที่เก็บบริเวณเยื่อกระดูกพุงแก้มทำการสกัดดีเอ็นเอด้วยชุด kit จากนั้นทำการเพิ่มปริมาณดีเอ็นเอด้วยเทคนิค multiplex PCR โดยใช้ primer ทั้งหมด 44 ตำแหน่ง โดยทำการวิเคราะห์ ณ ห้องปฏิบัติการมหาวิทยาลัยธรรมศาสตร์

3. การวิเคราะห์ด้วยเครื่อง MassARRAY® System

นำตัวอย่างที่ทำการเพิ่มปริมาณดีเอ็นเอด้วยเทคนิค multiplex PCR ทำการวิเคราะห์ความแตกต่างและความจำเพาะของน้ำหนักโมเลกุล (mass) ของเบสบนสายนิวคลีโอไทด์ (nucleotide) ด้วยเครื่อง MassARRAY® System (Agena Bioscience, USA.) ซึ่งจะสามารถตรวจวิเคราะห์สารพันธุกรรมได้คราวละหลายตำแหน่ง เป็น multiplexed assays

4. การวิเคราะห์ผล

ทำการวิเคราะห์จำนวนเบสที่เกิดจากการถ่ายทอดตามลักษณะพันธุกรรม โดยนำลำดับเบสของ SNPs ทั้ง 44 ตำแหน่ง สรุปเป็นจำนวน และค่าเฉลี่ยของการถ่ายทอดลักษณะทางพันธุกรรม จากนั้นคำนวณความสัมพันธ์

ตามสูตร ตารางที่ 1 โดยใช้ความถี่แอลลีลตามตารางที่ 2 โดยแบ่งเป็น 2 กลุ่มใหญ่ 2 กลุ่มย่อย ที่มีความสัมพันธ์ ดังนี้

- กลุ่มที่มีความสัมพันธ์กัน (True Parents-Child : TPC) โดยจะเป็นกลุ่มตัวอย่างทั้งหมด 38 ครอบครัว ที่มีความสัมพันธ์พ่อ-แม่-ลูกอย่างแท้จริง

- กลุ่มที่ไม่มีความสัมพันธ์กัน โดยนำกลุ่มที่มีความสัมพันธ์กันแท้จริงจาก 99 ตัวอย่างมาทำการสลับกันแบบสุ่ม (random pair) โดยการจับฉลากขึ้นมาทั้งหมด 3 คนต่อ 1 กลุ่ม ทั้งหมด 38 ครอบครัว

ตารางที่ 1 สูตรการคำนวณความสัมพันธ์พ่อ - แม่ - ลูก

พ่อต้องสงสัย	แม่	ลูก	PI*
ab	aa	aa	$0.5/p_a$
ab	aa	ab	$0.5/p_b$
ab	ab	aa	$0.5/p_a$
ab	bb	ab	$0.5/p_a$
aa	aa	aa	$1/p_a$
aa	ab	aa	$1/p_a$
aa	bb	ab	$1/p_a$
aa	ab	ab	$1/(p_a + p_b)$
ab	ab	ab	$1/(p_a + p_b)$

ที่มา. จาก *Forensic DNA evidence interpretation* (2nd ed.), by Buckleton, Bright, & Taylor, (2016), CRC Press, United States.

Bayesian statistics : An introduction (4th ed.), by Lee, P. M. (2012), Wiley, United State.

ตารางที่ 2 ความถี่แอลลีลของ SNPs 44 ตำแหน่ง (Chro.: Chromosome)

SNPs maker	Chro. (Agena Bioscience, 2021)	Allele frequency				Study (NCBI, 2021)
		A	T	C	G	
rs10495407	1	0.2390			0.7610	The PAGE Study
rs1413212	1		0.4940	0.5060		The PAGE Study
rs1982986	1		0.5350	0.4650		HapMap
rs891700	1	0.5005			0.4995	The PAGE Study
rs876724	2		0.4625	0.5375		The PAGE Study
rs1357617	3	0.2200	0.7800			HapMap
rs3819854	3		0.6060	0.3940		HapMap
rs2046361	4	0.3300	0.6700			HapMap
rs251934	5	0.8415			0.1585	The PAGE Study
rs717302	5	0.8851			0.1149	The PAGE Study
rs214054	6	0.5160			0.4840	HapMap
rs727811	6		0.7289		0.2711	The PAGE Study
rs2714854	7		0.5080	0.4920		HapMap
rs11781516	8		0.5739	0.4261		ExAC
rs2056277	8		0.1430	0.8570		The PAGE Study
rs1360288	9		0.3693	0.6307		The PAGE Study
rs1463729	9		0.3922	0.6078		The PAGE Study
rs826472	10		0.2349	0.7651		The PAGE Study
rs964681	10		0.7079	0.2921		The PAGE Study
rs2076848	11	0.5900	0.4100			HapMap
rs901398	11		0.6730	0.3270		HapMap
rs10771010	12		0.5097	0.4903		The PAGE Study
rs1994997	12	0.6540			0.3460	HapMap
rs1335873	13	0.6886	0.3114			The PAGE Study
rs354439	13	0.5980	0.4020			HapMap
rs9583190	13		0.5280	0.4720		HapMap
rs1454361	14	0.5230	0.4770			The PAGE Study
rs2518968	15			0.6220	0.3780	HapMap

SNPs maker	Chro. (Agena Bioscience, 2021)	Allele frequency				Study (NCBI, 2021)
		A	T	C	G	
rs8037429	15		0.4984	0.5016		The PAGE Study
rs1382387	16	0.6127		0.3873		The PAGE Study
rs729172	16		0.1594		0.8406	The PAGE Study
rs1468118	17		0.6520	0.3480		HapMap
rs2010253	17		0.5750	0.4250		HapMap
rs740910	17	0.9024			0.0976	The PAGE Study
rs1024116	18		0.1006	0.8994		The PAGE Study
rs1493232	18	0.3580		0.6420		The PAGE Study
rs2247221	18		0.4840	0.5160		HapMap
rs1005533	20	0.2929			0.7071	The PAGE Study
rs13050660	21		0.5650	0.4350		HapMap
rs2831700	21	0.5400			0.4600	The PAGE Study
rs914165	21	0.3334			0.6666	The PAGE Study
rs1028528	22	0.6510			0.3490	The PAGE Study
rs136337	22	0.3944		0.6056		The PAGE Study
rs2040411	22	0.3447			0.6553	The PAGE Study

■ ผลการวิจัย

1. การแชร์กันของแอลลีลในกลุ่มพ่อ-แม่-ลูก (Parentage allele sharing)

จากการทดสอบกลุ่มที่มีความสัมพันธ์พ่อ-แม่-ลูกที่แท้จริง (True parent-child group) และกลุ่มที่ไม่มีความสัมพันธ์กัน (unrelated group) กลุ่มละ 38 ครอบครัว ทำการวิเคราะห์จำนวนเบสที่เกิดจากการถ่ายทอดตามลักษณะพันธุกรรม โดยนำลำดับเบสของ SNPs ทั้ง 44

ตำแหน่ง และศึกษาการการใช้เบสร่วมกัน พบว่าในกลุ่มที่มีความสัมพันธ์พ่อ-แม่-ลูกที่แท้จริง มีการแชร์แอลลีลร่วมกันทั้งหมด 44 ตำแหน่ง คิดเป็นร้อยละ 100 และกลุ่มที่ไม่มีความสัมพันธ์กัน มีตำแหน่งที่ไม่ตรงกันตั้งแต่ 2-13 ตำแหน่ง (ตารางที่ 3) โดยในกลุ่มนี้ พบ 10 ครอบครัว ที่มีตำแหน่งที่ไม่ตรงกันถึง 4 ตำแหน่ง ซึ่งเป็นจำนวนตำแหน่งที่ไม่ตรงกันมากที่สุดคิดเป็นร้อยละ 26.32

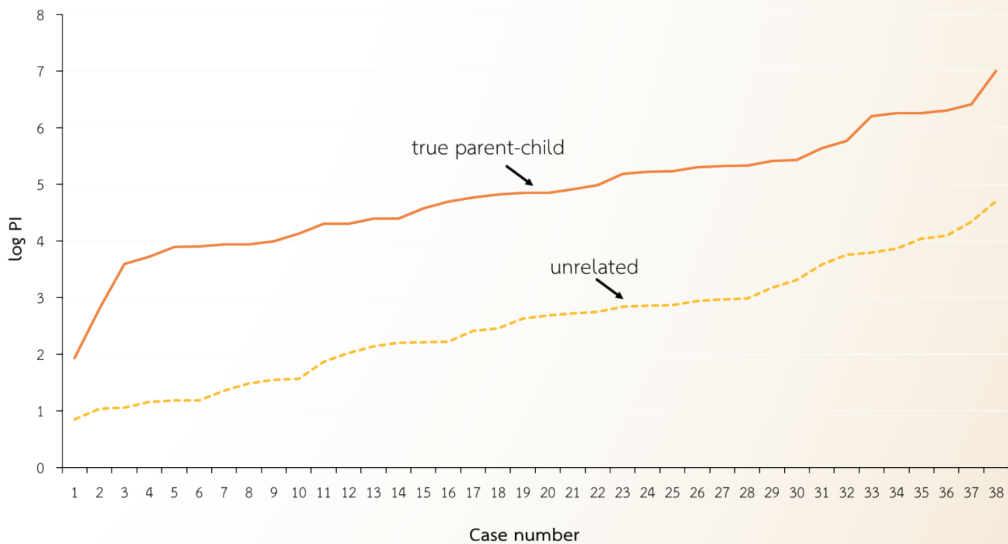
ตารางที่ 3 สรุปผลการใช้แอลลีลร่วมกันของกลุ่มที่มีความสัมพันธ์พ่อแม่-ลูกที่แท้จริง (True parent-child group) และกลุ่มที่ไม่มีความสัมพันธ์กัน (unrelated group) กลุ่มละ 38 ครอบครัว

ความสัมพันธ์	ผลการแชร์กันของตำแหน่ง SNPs					รวม
	แชร์กันทุกตำแหน่ง	ไม่ตรงกัน 1 ตำแหน่ง	ไม่ตรงกัน 2 ตำแหน่ง	ไม่ตรงกัน 3 ตำแหน่ง	ไม่ตรงกันมากกว่า 3 ตำแหน่ง	
พ่อแม่-ลูกที่แท้จริง	38	0	0	0	0	38
ไม่มีความสัมพันธ์กัน	0	0	1	4	33	38

2. ค่าดัชนีความเป็นบิดา (Paternity Index :PI)

ค่า PI จากทั้ง 2 กลุ่ม ถูกคำนวณและได้ผลออกมาในรูปแบบที่ 1 โดยกลุ่มที่มีความสัมพันธ์พ่อแม่-ลูกที่แท้จริง (True parent-child group) มีค่ามัธยฐาน เท่ากับ 6.9982×10^4 (log 4.8450)

มีค่าสูงสุดและต่ำสุด เท่ากับ 1.0120×10^7 และ 83.9371 (log 7.0052 และ 1.9240) ตามลำดับ และกลุ่มที่ไม่มีความสัมพันธ์กันมีค่ามัธยฐาน เท่ากับ 454.8812 (log 2.6570) มีค่าสูงสุดและต่ำสุด เท่ากับ 5.0065×10^4 และ 7.0750 (log 4.6995 และ 0.8497) (ตารางที่ 4)



ภาพที่ 1 ค่า PI values (แสดงค่าเป็น log) ระหว่างกลุ่มที่มีความสัมพันธ์พ่อแม่-ลูกที่แท้จริง (true parent-child group) และกลุ่มที่ไม่มีความสัมพันธ์กัน (unrelated group)

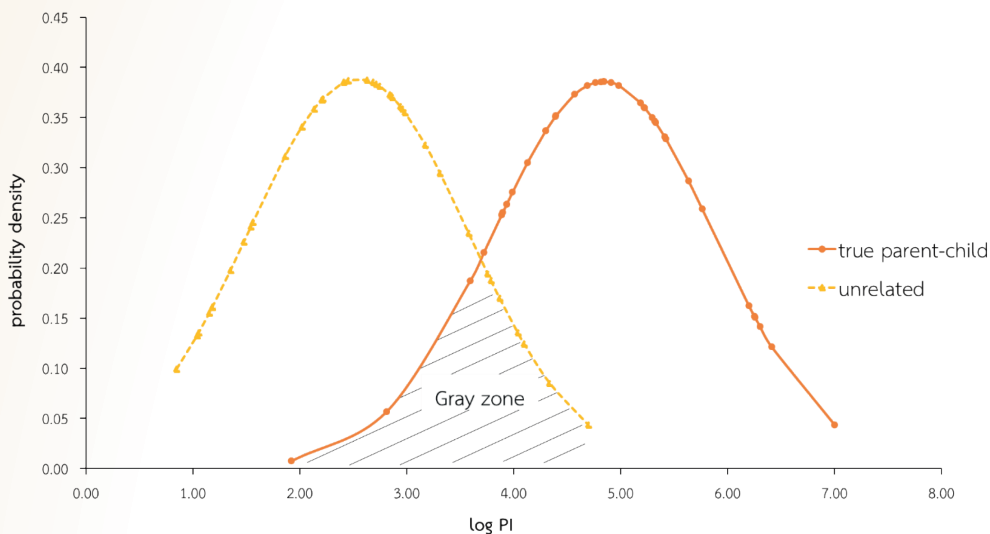
ตารางที่ 4 ค่าสถิติของกลุ่มที่มีความสัมพันธ์พ่อแม่-ลูกที่แท้จริง (true parent-child group) และกลุ่มที่ไม่มีความสัมพันธ์กัน (unrelated group)

ความสัมพันธ์	N	Mean	SD	Min	Max	Median
พ่อแม่-ลูกที่แท้จริง	38	6.1286×10^5	1.7162×10^6	83.9371	1.0120×10^7	6.9982×10^4
ไม่มีความสัมพันธ์กัน	38	3.4292×10^3	8.9449×10^3	7.0750	5.0065×10^4	454.8812

รูปและอภิปรายผล

เมื่อทำการเปรียบเทียบการกระจายตัวของค่า PI ของทั้ง 2 กลุ่ม พบว่า ค่าต่ำสุดของกลุ่มที่มีความสัมพันธ์พ่อแม่-ลูกที่แท้จริง ต่ำกว่าค่าสูงสุดของกลุ่มที่ไม่มีความสัมพันธ์กัน โดยจุดตัดมีค่าระหว่าง 83.9371 ถึง 5.0065×10^4 ($\log 1.9240$ ถึง 4.6995) ซึ่งในค่าช่วงนี้เป็นส่วนที่

ไม่สามารถตัดสินความสัมพันธ์ได้ซึ่งจะถือว่าเป็นส่วน “gray zone” หมายถึง กลุ่มตัวอย่างพ่อแม่-ลูก อยู่ในช่วงที่ซ้อนทับกันอยู่ ไม่สามารถตัดสินความสัมพันธ์พ่อแม่-ลูกที่แท้จริงได้ ซึ่งจะต้องทำการทดสอบด้วยวิธีอื่นเพื่อทำการยืนยันผลของความสัมพันธ์ (ภาพที่ 2)



ภาพที่ 2 การกระจายตัวของค่า PI values (แสดงค่าเป็น log) ของกลุ่มที่มีความสัมพันธ์พ่อแม่-ลูกที่แท้จริง (true parent-child group) และกลุ่มที่ไม่มีความสัมพันธ์กัน (unrelated group)

เกณฑ์ความแข็งแกร่งของค่า PI ในแง่ของการคำนวณโดยคำแนะนำของ John Butler ที่กำหนดเกณฑ์ออกเป็นทั้งหมด 4 เกณฑ์ ดังนี้ limited support ค่า PI อยู่ระหว่าง, moderate support ค่า PI อยู่ระหว่าง 10 ถึง 100, strong support ค่า PI อยู่ระหว่าง 100 ถึง 1000, และ very strong support ค่า PI อยู่ระหว่าง 1000 หรือมากกว่า (Butler, 2005) โดยในกลุ่มที่มีความสัมพันธ์พ่อ-แม่-ลูกที่แท้จริง สามารถแบ่งออกได้เป็น 3 เกณฑ์ ได้แก่ moderate support 1 ครอบครัว, strong support 1 ครอบครัว และ very strong

เกณฑ์การตัดสินความเป็นพ่อ-แม่-ลูก โดยใช้ค่า PI จะต้องมีความตั้งแต่ 5.0065×10^4 ($\log 4.6995$) หรือมากกว่า จากงานวิจัยก่อนหน้านี้ที่เปรียบเทียบค่า PI ในกลุ่มที่มีความสัมพันธ์พ่อ-แม่-ลูกที่แท้จริง จำนวน 1,209 ครอบครัว และกลุ่มที่ไม่มีความสัมพันธ์กัน จำนวน 16 ครอบครัว จากการวิเคราะห์โดยใช้ SNPs ทั้งหมด 136 ตำแหน่ง แต่ใช้เกณฑ์การตัดสินความสัมพันธ์พ่อ-แม่-ลูกที่แท้จริงนั้นจะต้องมีการแชร์กันของแอลลีลจำนวน 50 ตำแหน่ง พบว่า ค่า PI ของกลุ่มที่มีความสัมพันธ์พ่อ-แม่-ลูกที่แท้จริงนั้น

ตารางที่ 5 เกณฑ์ความแข็งแกร่งของค่า PI ในแง่ของการคำนวณโดยคำแนะนำของ Butler JM ของกลุ่มที่มีความสัมพันธ์พ่อ-แม่-ลูกที่แท้จริง และกลุ่มที่ไม่มีความสัมพันธ์กัน

ความสัมพันธ์/ เกณฑ์	limited support (PI :1-10)	moderate support (PI :10-100)	strong support (PI :100-1000)	very strong support (PI :>1000)	รวม (ครอบครัว)
พ่อ-แม่-ลูกที่แท้จริง	-	1 (2.63%)	1 (2.63%)	36 (94.74%)	38
ไม่มีความสัมพันธ์กัน	1 (2.63%)	10 (26.32%)	17 (44.73%)	10 (26.32%)	38

support 36 ครอบครัว และในกลุ่มที่ไม่มีความสัมพันธ์กันสามารถจัดได้ทั้งหมด 4 เกณฑ์ ได้แก่ limited support (1 ครอบครัว), moderate support (10 ครอบครัว), strong support (17 ครอบครัว) และ very strong support (10 ครอบครัว) (ตารางที่ 5) ซึ่งค่า PI ที่สูงกว่า 1,000 นั้น เกิดจากการที่ตัดตำแหน่งที่ไม่มีการแชร์กันของแอลลีลออกไป และนำค่าทั้งหมดมาคูณกันจึงทำให้เกิดค่าที่สูง อีกทั้งในกลุ่มการแชร์กันของแอลลีลเป็น homologous alleles จำนวนมากทำให้ค่า PI มีค่าสูงขึ้นตามไปด้วย

อยู่ระหว่าง $\log 1.25$ ถึง 6.39 ในขณะที่กลุ่มที่ไม่มีความสัมพันธ์กัน มีค่าระหว่าง 0.63 ถึง 3.57 (Cho et al., 2017) การศึกษานี้พบว่า มีส่วนที่ตัดกันของทั้ง 2 กลุ่มเช่นเดียวกับการศึกษาวิจัยนี้ นอกจากนี้ การยังพบว่ามีส่วน gray zone เช่นเดียวกับงานวิจัยอื่น ๆ ก่อนหน้า (Cho et al., 2017; Sanchez et al., 2006; Sun et al., 2020) อย่างไรก็ตาม หากตัดความไม่ตรงกันของแอลลีลเพื่อที่จำคำนวณค่า PI จะทำให้สามารถตกอยู่ในความสัมพันธ์ที่สามารถตัดสินได้และเกิดผลบวกลงขึ้น

ในการศึกษานี้ จะเห็นได้ว่า SNPs สามารถใช้แทนการคำนวณค่าความสัมพันธ์โดยใช้ STRs ได้เช่นเดียวกัน เนื่องจากลักษณะการคำนวณมีความคล้ายคลึงการคำนวณ PI ใน STRs ซึ่งในการวิเคราะห์ความสัมพันธ์โดยใช้ SNPs นั้น จำเป็นจะต้องมีการแชร์กันของแอลลีลในทุกตำแหน่ง ซึ่งจะช่วยให้แยกความแตกต่างระหว่างกลุ่มที่มีความสัมพันธ์และกลุ่มที่ไม่มีความสัมพันธ์ได้ เพื่อเพิ่มความน่าเชื่อถือ และสามารถวิเคราะห์ร่วมกัน STRs และตัวบ่งชี้พันธุกรรมอื่นๆ เพิ่มเติมได้ (Liu, Zhao, Xu, Zhang, & Li, 2019; Zhang et al., 2020)

ขอบเขตของการวิจัยนี้แสดงให้เห็นถึงจำนวนเบสบนโครโมโซมที่ได้จากการตรวจ SNPs 44 ตำแหน่ง ที่มีรูปแบบเหมือนกันของกลุ่มที่มีความสัมพันธ์พ่อ-แม่-ลูกที่แท้จริงนั้นจะต้องมีการใช้แอลลีลร่วมกันทุกตำแหน่ง และเมื่อศึกษาความสัมพันธ์ของเบสบนโครโมโซมจากการตรวจด้วย SNPs ที่จำแนกความเป็นพ่อ-แม่-ลูกที่แท้จริงและไม่แท้จริงนั้นจะต้องคำนึงถึงการตัดตำแหน่งของแอลลีลที่ไม่ได้แชร์กัน เพราะอาจเกิดความไม่สัมพันธ์กัน หรือทำให้เกิดการไม่สามารถตัดสินความสัมพันธ์นั้นได้

■ ข้อคิดเห็น/ข้อเสนอแนะ

การศึกษาวิจัยครั้งนี้พบข้อจำกัดของการศึกษาในเรื่องไม่มีการเปรียบเทียบกับค่าความสัมพันธ์แบบ STRs หากมีการศึกษาร่วมกันอาจจะทำให้ค่าความสัมพันธ์น่าเชื่อถือมากยิ่งขึ้น และการเพิ่มจำนวนตำแหน่งอาจจะช่วยให้ค่าความสัมพันธ์มีความแข็งแกร่งมากยิ่งขึ้น

■ กิตติกรรมประกาศ

งานวิจัยนี้เป็นผลงานวิจัยที่ได้รับทุนสนับสนุนการวิจัย ประเภททุนนักวิจัยรุ่นใหม่ มหาวิทยาลัย ธรรมศาสตร์ ตามสัญญาเลขที่ TUYSR62/2562

■ บรรณานุกรม

- Agena Bioscience. (2021). *Variant List—Sample ID Panel*. 13 มกราคม 2021, จาก <https://agenabio.com/products/panel/plex-pro-sample-id-panel/>
- Børsting, C., Sanchez, J. J., Hansen, H. E., Hansen, A. J., Bruun, H. Q., & Morling, N. (2018). Performance of the SNPforID 52 SNP-plex assay in paternity testing. *Forensic Science International: Genetics*. 2(4), 292-300. <https://doi.org/10.1016/j.fsigen.2008.03.007>
- Buckleton, J. S., Bright, J. A., & Taylor, D. (2016). *Forensic DNA evidence interpretation* (2nd ed.), CRC Press, United States.
- Butler, J. M. (2005). *Forensic DNA typing : biology, technology, and genetics of STR markers* (2nd ed.), Elsevier Academic Press, United States.
- Cho, S., Shin, E. S., Yu, H. J., Lee, J. H., Seo, H. J., & Kim, M. Y., et.al. (2017). Set up of cutoff thresholds for kinship determination using SNP loci. *Forensic Science International: Genetics*. 29, 1-8. <https://doi.org/10.1016/j.fsigen.2017.03.009>

- Elkins, K. M. (2013). Chapter 18 - Human Genetic Analysis: Paternity or Missing Persons Cases and Statistics. In: Elkins KM, editor. *Forensic DNA Biology*. San Diego, Academic Press, United States.
- Kim, J. J., Han, B. G., Lee, H. I., Yoo, H. W., & Lee, J. K. (2010). Development of SNP-based human identification system. *International Journal of Legal Medicine*. 124(2), 125-131. <https://doi.org/10.1007/s00414-009-0389-9>
- Lee, P. M. (2012). *Bayesian statistics : An introduction* (4th ed.), Wiley, United State
- Liu, X., Zhao, Z., Xu, Q., Zhang, J., & Li, C. (2019). Analysis of full- and half-siblings using a combined system of STR, InDel and SNP markers. *Forensic Science International: Genetics Supplement*. 7(1), 349-350. <https://doi.org/10.1016/j.fsigss.2019.10.007>.
- Lou, C., Cong, B., Li, S., Fu, L., Zhang, X., & Feng, T., et.al. (2011). A SNaPshot assay for genotyping 44 individual identification single nucleotide polymorphisms. *Electrophoresis*. 32(3-4), 368-78. <https://doi.org/10.1002/elps.201000426>
- NCBI. (2021). *dbSNP Allele Frequency of Homo sapiens (Asian)*. ค้นเมื่อ 26 ตุลาคม 2021, จาก <https://www.ncbi.nlm.nih.gov/snp/>
- Sanchez, J. J., Phillips, C., Børsting, C., Balogh, K., Bogus, M., & Fondevila, M., et.al. (2006). A multiplex assay with 52 single nucleotide polymorphisms for human identification. *Electrophoresis*. 27(9), 1713-24. <https://doi.org/10.1002/elps.200500671>
- Sun, S., Liu, Y., Li, J., Yang, Z., Wen, D., & Liang, W., et.al. (2020). Development and application of a nonbinary SNP-based microhaplotype panel for paternity testing involving close relatives. *Forensic Science International: Genetics*. 46102255. <https://doi.org/10.1016/j.fsigen.2020.102255>
- Yamane, T. (1973). *Statistics: an introductory analysis*. New York: New York: Harper & Row.
- Yousefi, S., Abbassi-Daloi, T., Kraaijenbrink, T., Vermaat, M., Mei, H., & van't Hof, P., et.al. (2018). A SNP panel for identification of DNA and RNA specimens. *BMC Genomics*. 19(1), 90. <https://doi.org/10.1186/s12864-018-4482-7>
- Zhang, Q., Zhou, Z., Wang, L., Quan, C., Liu, Q., & Tang Z., et.al. (2020). Pairwise kinship testing with a combination of STR and SNP loci. *Forensic Science International: Genetics*. 46, 102265. <https://doi.org/10.1016/j.fsigen.2020.102265>



