

การวิเคราะห์โครงสร้างทางพันธุกรรมของประชากรชาวไทยอีสาน จังหวัดบุรีรัมย์ ด้วยไมโครแซทเทลไลต์บนออโตโซม

Genetic structure of the Thai-Isan from Buriram Province as revealed by autosomal microsatellites analysis

กนกพร ศรีทองแดง, ศุภรัตน์ ศรีทะวงษ์, วิภู กุดะนันท์*

Kanokpohn Srithongdaeng, Suparat Srithawong, Wibhu Kutanan*

ภาควิชาชีววิทยา คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น ขอนแก่น 40002

Department of Biology, Faculty of Science, Khon Kaen University, Khon Kaen 40002, Thailand

*Corresponding author: wibhu@kku.ac.th

บทคัดย่อ

ชาวไทยอีสานเป็นกลุ่มชาติพันธุ์ขนาดใหญ่ที่สุด อาศัยอยู่ในภาคตะวันออกเฉียงเหนือของประเทศไทย ชาวไทยอีสานอพยพมาจากประเทศลาวในระหว่างศตวรรษที่ 14 ถึง 18 งานวิจัยที่ผ่านมาได้มีการศึกษาโครงสร้างทางพันธุกรรมของชาวไทยอีสาน ด้วยเครื่องหมายทางพันธุกรรมชนิดเดียวกันเท่านั้น คือ ดีเอ็นเอไมโทคอนเดรีย งานวิจัยนี้จึงต้องการขยายการศึกษาโครงสร้างทางพันธุกรรมของชาวไทยอีสาน จำนวน 53 คน จากจังหวัดบุรีรัมย์ ด้วยไมโครแซทเทลไลต์บนออโตโซมจำนวน 15 ตำแหน่ง คือ D8S1179, D21S11, D7S820, CSF1PO, D3S1358, THO1, D13S317, D16S539, vWA, TPOX, D18S51, D5S818, FGA, D19S433 และ D2S1338 ผลการศึกษาพบว่าประชากรที่ศึกษามีความหลากหลายทางพันธุกรรมสูง (จำนวนแอลลีลทั้งหมด 122 แอลลีล และความหลากหลายของยีน เท่ากับ 0.7094 ± 0.4406) เมื่อวิเคราะห์ความสัมพันธ์ทางเชื้อสายระหว่างประชากร โดยเปรียบเทียบกับประชากรอื่นที่อยู่ในภาคตะวันออกเฉียงเหนือของประเทศไทย พบว่าชาวไทยอีสาน จังหวัดบุรีรัมย์ มีความใกล้ชิดทางเชื้อสายกับชาวไทยอีสาน จังหวัดร้อยเอ็ด และชาวไทยอีสานทั้งสองประชากรแสดงลักษณะทางพันธุกรรมของประชากรลูกผสม เนื่องจากมีตำแหน่งอยู่กึ่งกลางระหว่างประชากรอื่นในแผนภูมิต้นไม้แสดงความสัมพันธ์ทางวิวัฒนาการและแผนภูมิ multidimensional scaling เมื่อคำนวณสัดส่วนการ

ผสมผสานทางพันธุกรรมของประชากรพบว่า ชาวไทยอีสาน จังหวัดบุรีรัมย์ ได้รับสัดส่วนทางพันธุกรรมจากประชากรพ่อแม่ที่พูดภาษาตระกูลกระได (กะเลิงและญ้อ) ร้อยละ 65 มากกว่าประชากรที่พูดภาษาตระกูลออสโตรเอเชียติก (ชาวนน เขมร และส่วย) ร้อยละ 35

ABSTRACT

The Thai-Isan is the largest ethnic group in Northeastern Thailand who migrated from Laos during the 14th to 18th centuries C.E. Previous literatures have reported genetic structure in the Thai-Isan populations using mitochondrial DNA as the only genetic marker. To extend this line of genetic investigation, this study analysed 15 autosomal microsatellites (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, THO1, D13S317, D16S539, vWA, TPOX, D18S51, D5S818, FGA, D19S433 and D2S1338) of 53 Thai-Isan subjects from Buriram Province. The studied population showed high genetic diversity (122 total alleles and gene diversity = 0.7094 ± 0.4406). When genetic relationship of the studied population was compared with other northeastern Thai populations, the Thai-Isan from Buriram and Roi-Ed exhibited the greatest genetically relatedness. Both of them displayed a signal of an admixed population owing

to intermediated positions between the other populations as depicted in neighbor joining tree and multidimensional scaling plot. Admixture proportions observed in Thai-Isan from Buriram Province (Isan-Bu) exhibited a higher contribution from the parental Kra-Dai speaking group (Nyo and Kaleung) (65%) than the Austroasiatic speaking group (Chaobon Khmer and Suay) (35%).

คำสำคัญ: ไทยอีสาน; ไมโครแซทเทลไลต์; โครงสร้างทางพันธุกรรม; กระได; ออสโตรเอเชียติก

Keywords: Thai-Isan; microsatellite; genetic structure; Kra-Dai; Austroasiatic

บทนำ

“อีสาน” เป็นคำที่คนไทยใช้เรียกคนหรือพื้นที่ที่ตั้งอยู่ในภาคตะวันออกเฉียงเหนือของประเทศไทย คนไทยอีสาน (Thai-Isan) หมายถึงกลุ่มชาติพันธุ์ลาวที่อาศัยอยู่ในภาคตะวันออกเฉียงเหนือของประเทศไทย และมีสัญชาติไทย นอกจากชาวไทยอีสานซึ่งเป็นกลุ่มชาติพันธุ์ที่มีจำนวนมากที่สุดแล้ว ในภาคอีสานยังประกอบด้วยประชากรอีกจำนวนมากทั้งประชากรดั้งเดิมที่พูดภาษาตระกูลออสโตรเอเชียติก (Austroasiatic family) เช่น ชาวบน เขมร ส่วย และไล่ และประชากรที่พูดภาษาตระกูลกระได (Kra-Dai family) เช่น ภูไท แสกญ้อ และกะเลิง (Lewis, 2009) จากหลักฐานทางประวัติศาสตร์ ชาวไทยอีสานถูกกวาดต้อนให้อพยพมาจากประเทศลาว เมื่อประมาณกลางศตวรรษที่ 14 และปลายศตวรรษที่ 18 (Keyes, 1988) แต่ระยะเวลาก่อนหน้าที่ชาวลาวจะอพยพโยกย้าย ดินแดนแห่งนี้ได้มีชาวเขมรออาศัยอยู่ก่อนดังมีหลักฐานทางโบราณคดีจำนวนมากที่เป็นของชาวเขมร ซึ่งยังคงปรากฏในปัจจุบัน เช่น ปราสาทหิน โดยเฉพาะบริเวณภาคอีสานตอนใต้ นอกจากนั้นหลักฐานทางวัฒนธรรมและภาษาได้ระบุถึงการผสมผสานวัฒนธรรมระหว่างชาวไทยอีสานและชาวเขมร (biculturalism) โดยเฉพาะในเขตจังหวัดสุรินทร์ ศรีสะเกษ และบุรีรัมย์ (Vail, 2007) จึงทำให้นักวิชาการจากหลากหลายสาขา เช่น ภาษาศาสตร์ (Smalley, 1988; 1994) โบราณคดี (Welch, 1998; Talbot and

Chutima, 2001) วัฒนธรรมและประวัติศาสตร์ (Khanittanan, 2001; Wyatt, 2003) สนใจเข้ามาศึกษาในพื้นที่บริเวณนี้จำนวนมาก

ที่ผ่านมาได้มีการศึกษาพันธุศาสตร์ในประชากรที่อาศัยอยู่ในบริเวณภาคอีสานตอนใต้ (Boonsoda *et al.*, 2013; Kutanan *et al.* 2014a; 2014b) แต่การศึกษาดังกล่าวอาศัยเพียงเครื่องหมายทางพันธุกรรมดีเอ็นเอไมโทคอนเดรีย แม้ว่าดีเอ็นเอไมโทคอนเดรียจะมีประสิทธิภาพในการแยกประชากร แต่เนื่องจากการมีขนาดของประชากรในอดีต (effective population size) เพียง 1/4 เท่าของดีเอ็นเอไมโทโซม ทำให้เครื่องหมายทางพันธุกรรมชนิดนี้ไวต่อการเกิดเจเนติกดริฟท์ จึงส่งผลให้ประชากรเกิดความแตกต่างทางพันธุกรรม (genetic differentiation) มากกว่าความเป็นจริง (Kayser, 2007) ดังนั้นการศึกษาครั้งนี้จึงต้องการศึกษาพันธุกรรมของประชากรเพิ่มเติม โดยอาศัยเครื่องหมายพันธุกรรมชนิดไมโครแซทเทลไลต์บนออโตโซม (autosomal microsatellite) ซึ่งมีรูปแบบการถ่ายทอดที่แตกต่างออกไปจากดีเอ็นเอไมโทคอนเดรีย กล่าวคือดีเอ็นเอไมโทโซมมีการถ่ายทอดผ่านทางฝ่ายพ่อและแม่ และเกิดรีคอมบิเนชัน ดังนั้นจึงไม่ไวต่อเจเนติกดริฟท์ นอกจากนั้นวิธีการที่ใช้วิเคราะห์โครงสร้างทางพันธุกรรมด้วยดีเอ็นเอไมโทโซม จะมีวิธีที่ซับซ้อนและให้ข้อมูลมากกว่าดีเอ็นเอไมโทคอนเดรียอีกด้วย ไมโครแซทเทลไลต์บนออโตโซมที่เลือกใช้เป็นชุดเครื่องหมายทางพันธุกรรมที่นิยมใช้ในงานด้านนิติวิทยาศาสตร์ ซึ่งประกอบด้วยไมโครแซทเทลไลต์ จำนวน 15 ตำแหน่ง คือ D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, vWA, TPOX, D18S51, D5S818, FGA, D19S433 และ D2S1338 โดยชุดเครื่องหมายพันธุกรรมนี้ได้พิสูจน์แล้วว่าสามารถแยกความแตกต่างของประชากรในระดับภูมิภาคได้ (Crossetti *et al.*, 2008; Gaibar *et al.*, 2010; Babiker *et al.*, 2011; Kraaijenbrink *et al.*, 2014) ผู้วิจัยจึงสนใจที่จะนำชุดเครื่องหมายทางพันธุกรรมดังกล่าวมาใช้วิเคราะห์โครงสร้างทางพันธุกรรมของประชากรชาวไทยอีสาน จากจังหวัดบุรีรัมย์ และยังใช้ศึกษาความสัมพันธ์ทางเชื้อสายระหว่างประชากรที่ศึกษากับประชากรอื่นที่ใช้เปรียบเทียบ

อุปกรณ์และวิธีการ

กลุ่มตัวอย่างที่ศึกษา

อาสาสมัครชาวไทยอีสาน จากบ้านท่างาม ตำบล พุทไธสง อำเภอพุทไธสง จังหวัดบุรีรัมย์ จำนวน 53 คน โดยมีเกณฑ์การคัดเลือกอาสาสมัคร คือ อาสาสมัครและบรรพบุรุษ 2 ชั่วรุ่นจะต้องเป็นชาวไทยอีสานที่สามารถพูดภาษาไทยอีสาน และจะต้องมีภูมิลำเนาอยู่ที่อำเภอพุทไธสง นอกจากนั้นอาสาสมัครทั้งหมดจะต้องไม่มีความสัมพันธ์กันทางเครือญาติอย่างน้อย 2 ชั่วรุ่น แล้วเก็บตัวอย่างเยื่อ buccal cells โดยใช้แปรงเก็บตัวอย่าง (Qiagen, Germany) ถูเนื้อเยื่อบริเวณกระพุ้งแก้มประมาณ 30 ครั้ง งานวิจัยนี้ได้ผ่านการรับรองจริยธรรมการวิจัยในมนุษย์ จากคณะกรรมการจริยธรรมการวิจัยในมนุษย์ มหาวิทยาลัยขอนแก่น (รหัสโครงการ HE55167)

การสกัดดีเอ็นเอและการหาจีโนไทป์

นำตัวอย่างเยื่อ buccal cells มาสกัดดีเอ็นเอ ด้วยชุดสกัดสำเร็จรูป (Qiagen) วัดปริมาณดีเอ็นเอด้วยเครื่อง Nanodrop 2000 spectrophotometer (Thermo scientific, USA) จากนั้นปรับลดปริมาณดีเอ็นเอให้มีความเข้มข้นประมาณ 1 ถึง 5 นาโนกรัมต่อไมโครลิตร เพิ่มปริมาณดีเอ็นเอด้วยวิธีมีลติเพล็กซ์ พีซีอาร์ (multiplex PCR) โดยใช้ชุดน้ำยา AmpFSTR Identifier kit (Applied Biosystem, USA) วิธีการเพิ่มปริมาณดีเอ็นเอ จะทำตามที่บริษัทแนะนำ แต่ผู้วิจัยได้ทำการปรับลด ปริมาตรในการทำพีซีอาร์จากเดิมลงครึ่งหนึ่ง เป็น ปริมาตรรวม 12.5 ไมโครลิตร นำผลผลิตพีซีอาร์ไปวิเคราะห์ขนาดของไมโครแซทเทลไลต์แต่ละตำแหน่ง (fragment analysis) ด้วยเครื่องหาลำดับเบสอัตโนมัติ รุ่น 3130xl Genetic Analyzer (Applied Biosystem) โดยเปรียบเทียบกับดีเอ็นเอที่ทราบขนาด (GeneScan-500 LIZ dye และ Identifier allelic ladder) (Applied Biosystem) อ่านผลที่ได้ด้วยโปรแกรม Gene Mapper v.3.7 (Applied Biosystem)

การวิเคราะห์ข้อมูล

คำนวณค่าความหลากหลายทางพันธุกรรม ภายในประชากร เช่น จำนวนแอลลีลทั้งหมด ความถี่ แอลลีล ค่า observed และ expected heterozygosity

(H_o และ H_E ตามลำดับ) ค่าความหลากหลายของยีน (gene diversity, GD) (Nei, 1987) ด้วยโปรแกรม ARLEQUIN 3.5 (Excoffier and Lischer, 2010) คำนวณ ค่าพารามิเตอร์สำหรับงานนิติวิทยาศาสตร์ เช่น power of discrimination (PD), polymorphism information content (PIC), combined matching probability (CMP), combined power of exclusion (CPE) และ combined power of discrimination (CPD) ด้วยโปรแกรม PowerStats (www.promega.com)

ประเมินความสัมพันธ์ระหว่างประชากรชาวไทยอีสาน จังหวัดบุรีรัมย์ (Isan-Bu) กับประชากรอื่นที่อาศัยอยู่ในภาคอีสานของประเทศไทย จำนวน 8 ประชากร ประกอบด้วยประชากรที่พูดภาษาตระกูลออสโตรเอเชียติก จำนวน 4 ประชากร คือ เขมร (Khmer) ส่วย (Suay) ชาวนน (Chaobon) และโส้ (So) (Unpublished data) และประชากรที่พูดภาษาตระกูลกระไดอีก 4 ประชากร คือ ภูไท (Phu Thai) ญ้อ (Nyo) กะเลิง (Kaleung) และ ไทยอีสาน จากจังหวัดร้อยเอ็ด (Isan-Ro) (Srithawong *et al.*, 2015) ผู้วิจัยเลือกใช้การวิเคราะห์การจัดกลุ่ม (clustering analysis) เพื่อประเมินความสัมพันธ์ระหว่างประชากร โดยใช้วิธี distance based clustering และ model based clustering ในวิธี distance based clustering คำนวณระยะห่างทางพันธุกรรมระหว่างประชากร (genetic distance) แบบ F_{st} ด้วยโปรแกรม ARLEQUIN 3.5 จากนั้นนำค่าระยะห่างทางพันธุกรรมไปสร้างเป็นแผนภูมิ multidimensional scaling (MDS) ด้วยโปรแกรม STATISTICA v.10 (Statsoft.) และยังใช้ค่าระยะห่างทางพันธุกรรมดังกล่าวมาสร้างเป็นแผนภูมิต้นไม้แสดงความสัมพันธ์ทางวิวัฒนาการ (phylogenetic tree) แบบ neighbor joining (NJ) ด้วยโปรแกรม MEGA v.5 (Tamura *et al.*, 2011) ในวิธี model based clustering จะใช้วิธีการจัดกลุ่มที่อาศัยทฤษฎีของเบย์ (Bayesian clustering method ด้วยโปรแกรม STRUCTURE 2.3 (Pritchard *et al.*, 2000; Falush *et al.*, 2003; Hubisz *et al.*, 2009) ในขั้นตอนการตั้งค่า เพื่อคำนวณค่า posterior probability ในแต่ละครั้ง จะใช้จำนวนซ้ำของ Markov Chain Monte Carlo (MCMC) ในช่วง burn-in จำนวน 50,000 ครั้ง ตามด้วยจำนวนซ้ำ (replication) อีก 50,000 ครั้ง และเลือกใช้โมเดล 3

ชนิดคือ admixture, correlated allele frequency และ LOCPRIOR (Hubisz *et al.*, 2009) ทำการทดลองจำนวน 10 ซ้ำ ในทุกๆ จำนวนกลุ่ม (cluster, K) ตั้งแต่ 1 ถึง 9 รวมทั้งหมด 90 ครั้ง จากนั้นคำนวณหาจำนวนกลุ่มที่เหมาะสมที่สุด ด้วยวิธีของ Evanno *et al.* (2005) ด้วยโปรแกรม Structure Harvester (Earl and vonHoldt, 2011)

วิเคราะห์การผสมผสานทางพันธุกรรม (admixture coefficient, m_y) ด้วยโปรแกรม ADMIX 2.0 เนื่องจากการคำนวณ m_y อาศัยค่าความถี่แอลลีลของไมโครแซทเทลไลต์และเวลาในการเกิดการผสมผสานทางพันธุกรรมระหว่างประชากรในภาคตะวันออกเฉียงเหนือไม่ยาวนานมากนัก ดังนั้นในการคำนวณจะไม่ใช้โมเดลการกลายพันธุ์ (no mutation model) และโมเดล divergence of allele (Bertorelle and Excoffier, 1998; Dupanloup and Bertorelle, 2001; Belle *et al.*, 2006) เพื่อให้ได้ค่า m_y แม่นยำมากที่สุด

ผลการทดลองและวิจารณ์

ความหลากหลายทางพันธุกรรมภายในประชากร

จากการศึกษาไมโครแซทเทลไลต์ จำนวน 15 ตำแหน่ง ของประชากรชาวไทยอีสาน จากจังหวัดบุรีรัมย์ จำนวน 53 คน พบว่ามีจำนวนแอลลีลทั้งหมด 122 แอลลีล ค่า H_o และ H_E เท่ากับ 0.6806 และ 0.7862 ตามลำดับ และค่า GD เท่ากับ 0.7094 ± 0.4406 โดยค่าดังกล่าวอยู่ในช่วงเดียวกับประชากรที่พูดภาษาตระกูลกระไดที่เคยมีรายงานมาก่อนหน้านี้ (Srithawong *et al.*, 2015) เนื่องจากเครื่องหมายทางพันธุกรรมที่เลือกใช้เป็นเครื่องหมายทางพันธุกรรมด้านนิติวิทยาศาสตร์ ผลการศึกษาพบว่า ค่า CMP, CPE และ CPD มีค่าเท่ากับ 1 ใน 1.15×10^{-15} , 0.999049 และ 0.9999999977 ตามลำดับ

เมื่อพิจารณาไมโครแซทเทลไลต์แต่ละตำแหน่ง พบว่า ค่า PD มีค่าสูงที่สุด (0.944) และต่ำสุด (0.765) ใน FGA และ D3S1358 ตามลำดับ ค่า PIC มีค่าระหว่าง 0.55 (D3S1358 และ TPOX) ถึง 0.83 (FGA) ผลดังกล่าวแสดงให้เห็นว่า FGA เป็นตำแหน่งที่มีความสามารถในการแยกความแตกต่างระหว่างบุคคลได้ดีที่สุด (Table 1) หลังจากการปรับระดับนัยสำคัญด้วย

วิธีของ Bonferroni ($P = 0.05/15$) พบว่ามีเพียง D18S51 ตำแหน่งเดียวเท่านั้นที่ไม่อยู่ในสมมูลของฮาร์ดีและไวน์เบิร์ก

ความสัมพันธ์ระหว่างประชากร

จากค่าระยะห่างทางพันธุกรรมแบบ F_{st} ระหว่างคู่ประชากรจำนวน 36 คู่ มีเพียง 1 คู่เท่านั้นที่ไม่แตกต่างอย่างมีนัยสำคัญ หลังจากปรับระดับนัยสำคัญด้วยวิธีของ Bonferroni ($P < 0.001388$) คือ คู่ของชาวไทยอีสาน จากจังหวัดบุรีรัมย์ (Isan-Bu) และชาวไทยอีสาน จากจังหวัดร้อยเอ็ด (Isan-Ro) แสดงถึงการมีความสัมพันธ์ทางเชื้อสายที่ใกล้ชิดกันมากระหว่าง 2 ประชากรนี้ เมื่อนำค่าระยะห่างทางพันธุกรรมดังกล่าวไปสร้างเป็นแผนภูมิ MDS และ แผนภูมิต้นไม้แสดงความสัมพันธ์ทางวิวัฒนาการ แบบ NJ พบว่าจากแผนภูมิ MDS (Figure 1) ประชากรที่พูดภาษาตระกูลกระได (Isan-Bu, Isan-Ro, Phu Thai, Nyo, และ Kaleung) มีตำแหน่งที่ใกล้ชิดกันตรงกลางของแผนภูมิ ในขณะที่ประชากรที่พูดภาษาตระกูลออสโตรเอเชียติก (Khmer, Suay, So และ Chaobon) มีตำแหน่งกระจายอยู่บริเวณขอบของแผนภูมิ และจากแผนภูมิต้นไม้พบว่าประชากรที่พูดภาษาตระกูลกระไดมีระยะห่างระหว่างกันบนแผนภูมิต้นไม้ไม่น้อยกว่าประชากรที่พูดภาษาตระกูลออสโตรเอเชียติก จากผลของการวิเคราะห์การจัดกลุ่มแบบ distance based clustering (Figure 1) แสดงให้เห็นว่าประชากรที่พูดภาษาตระกูลกระได มีความเหมือนกันทางพันธุกรรม (genetic homogeneity) ซึ่งอาจเกิดจากการมีบรรพบุรุษร่วมกันในอดีต (ได้อภิปรายไว้ในการศึกษาของ Srithawong *et al.* (2015)) ในขณะที่ประชากรที่พูดภาษาตระกูลออสโตรเอเชียติกมีความสัมพันธ์ที่แตกต่างไปจากประชากรกระไดและแตกต่างกันเองด้วย แสดงให้เห็นถึงความแตกต่างกันทางพันธุกรรม (genetic heterogeneity) ซึ่งอาจเกิดมาจากหลายปัจจัย เช่น การเกิดเจเนติกดริฟท์ ซึ่งเคยมีรายงานจากการศึกษาดีเอ็นเอไมโทคอนเดรียในประชากรเดียวกัน (Kutanan *et al.*, 2014b) หรืออาจเกิดจากอิทธิพลของประเพณีนิยมในการแต่งงาน และ/หรือ การผสมเลือดชิด (Kutanan *et al.*, 2011a and 2011b) รูปแบบการมีความใกล้ชิดทางพันธุกรรมในประชากรที่พูดภาษาตระกูลกระได และความแตกต่างระหว่างกัน

ของประชากรที่พูดภาษาตระกูลออสโตรเอเชียติก นอกจากจะพบในประชากรที่อาศัยอยู่ในภาคตะวันออกเฉียงเหนือของประเทศไทยแล้ว ยังมีรายงานในประชากรที่อาศัยอยู่ในภาคเหนือของประเทศไทยอีกด้วย (Kutanen *et al.*, 2011a; 2011b) นอกจากนี้ จาก Figure 2 ยังพบว่าคู่ของชาวไทยอีสาน (Isan-Bu และ

Isan-Ro) จัดอยู่บนกิ่งก้านสาขาเดียวกัน แสดงถึงการมีความใกล้ชิดทางพันธุกรรมกันมากที่สุด และทั้งสองประชากรมีตำแหน่งบนแผนภูมิต้นไม้ที่อยู่ตรงกลางระหว่างประชากรอื่น แสดงถึงการมีพันธุกรรมของประชากรลูกผสม (admixed population) (Di Benedetto *et al.*, 2001)

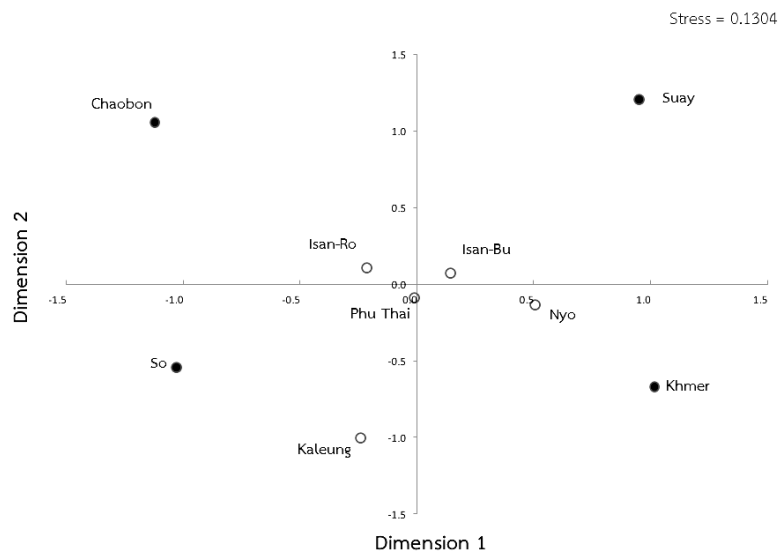


Figure 1 Multidimensional scaling scatter plot (MDS) two dimensions based on the F_{st} values between the studied population and comparable populations (○: Kra-Dai speaking populations, ●: Austroasiatic speaking populations).

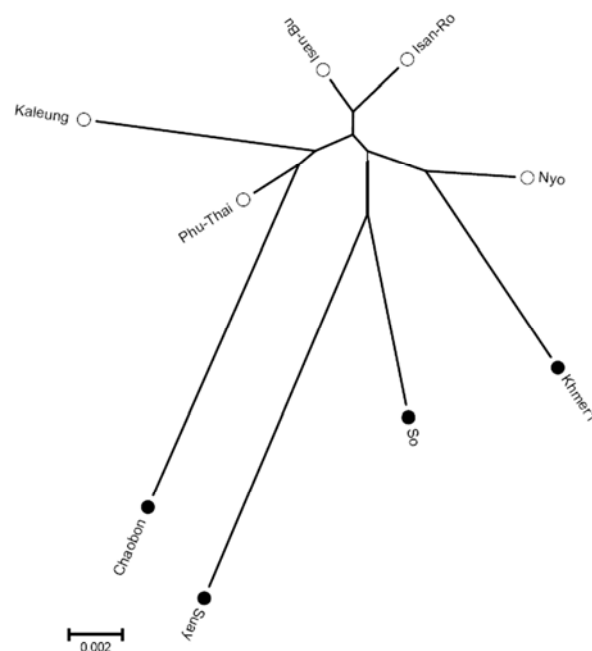


Figure 2 Unrooted neighbor joining tree (NJ) based on the F_{st} values between the studied population and comparable populations (○: Kra-Dai speaking populations, ●: Austroasiatic speaking populations).

Table 1 Microsatellite allele frequency of 15 loci in the Thai-Isan from Buriram Province ($n = 53$).

Allele	<i>D8S1179</i>	<i>D21S11</i>	<i>D7S820</i>	<i>CSF1PO</i>	<i>D3S1358</i>	<i>TH01</i>	<i>D13S317</i>	<i>D16S539</i>	<i>D2S1338</i>	<i>D19S433</i>	<i>vWA</i>	<i>TPOX</i>	<i>D18S51</i>	<i>D5S818</i>	<i>FGA</i>
5.3	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
6	--	--	--	--	--	10.0	--	--	--	--	--	--	--	1.0	--
6.3	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
7	--	--	--	--	--	40.0	1.0	--	--	--	--	--	--	6.0	--
8	--	--	18.9	--	--	6.0	39.8	1.0	--	--	--	56.7	--	1.0	--
8.2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
9	--	--	8.9	--	--	28.0	12.2	21.0	--	--	--	16.3	--	3.0	--
9.2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
9.3	--	--	--	--	--	6.0	--	--	--	--	--	--	--	--	--
10	12.1	--	17.8	28.1	--	10.0	10.2	28.0	--	--	--	4.8	--	29.0	--
11	15.5	--	30.0	34.4	--	--	26.5	16.0	--	1.1	1.0	21.2	--	28.0	--
12	15.5	--	21.1	27.1	--	--	9.2	22.0	--	11.1	--	--	9.6	14.0	--
13	32.8	--	3.3	9.4	2.0	--	--	11.0	--	26.7	2.0	1.0	15.4	15.0	--
13.2	--	--	--	--	--	--	--	--	--	4.4	--	--	--	--	--
14	20.7	--	--	1.0	2.0	--	1.0	1.0	--	22.2	28.4	--	32.7	3.0	--
14.2	--	--	--	--	--	--	--	--	--	18.9	--	--	--	--	--
15.2	--	--	--	--	2.0	--	--	--	--	14.4	--	--	--	--	--
16	3.4	--	--	--	56.0	--	--	--	3.1	--	18.6	--	25.0	--	1.1
16.2	--	--	--	--	--	--	--	--	--	1.1	--	--	--	--	--
17	--	--	--	--	26.0	--	--	--	14.6	--	25.5	--	3.8	--	--
18	--	--	--	--	12.0	--	--	--	7.3	--	16.7	--	--	--	1.1
19	--	--	--	--	--	--	--	--	22.9	--	7.8	--	7.7	--	5.6
20	--	--	--	--	--	--	--	--	5.2	--	--	--	3.8	--	6.7
20.2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	2.2
21	--	--	--	--	--	--	--	--	3.1	--	--	--	--	--	30.0
22	--	--	--	--	--	--	--	--	3.1	--	--	--	--	--	13.3
22.2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	3.3
23	--	--	--	--	--	--	--	--	16.7	--	--	--	--	--	10.0
23.2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	3.3
24	--	--	--	--	--	--	--	--	24.0	--	--	--	--	--	14.4
24.2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	1.1
25	--	--	--	--	--	--	--	--	--	--	--	--	1.9	--	4.4
25.2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
26	--	--	--	--	--	--	--	--	--	--	--	--	--	--	2.2
26.2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	1.1
27	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
28	--	4.3	--	--	--	--	--	--	--	--	--	--	--	--	--
28.2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
29	--	31.5	--	--	--	--	--	--	--	--	--	--	--	--	--
30	--	25.0	--	--	--	--	--	--	--	--	--	--	--	--	--
30.2	--	5.4	--	--	--	--	--	--	--	--	--	--	--	--	--
31	--	7.6	--	--	--	--	--	--	--	--	--	--	--	--	--
31.2	--	7.6	--	--	--	--	--	--	--	--	--	--	--	--	--
32	--	5.4	--	--	--	--	--	--	--	--	--	--	--	--	--
32.2	--	10.9	--	--	--	--	--	--	--	--	--	--	--	--	--
33.2	--	2.2	--	--	--	--	--	--	--	--	--	--	--	--	--
H_o	0.833	0.820	0.667	0.646	0.620	0.560	0.755	0.620	0.646	0.804	0.654	0.442	0.617	0.816	0.708
H_e	0.835	0.835	0.798	0.728	0.745	0.742	0.745	0.799	0.838	0.837	0.796	0.610	0.826	0.795	0.864
HWE	0.908	0.153	0.030	0.243	0.024	0.028	0.809	0.004	0.050	0.201	0.046	0.013	0.003	0.331	0.043
PM	0.089	0.077	0.087	0.140	0.235	0.114	0.113	0.091	0.060	0.081	0.088	0.213	0.107	0.092	0.056
PD	0.911	0.923	0.913	0.860	0.765	0.886	0.887	0.909	0.940	0.919	0.912	0.787	0.893	0.908	0.944
PIC	0.75	0.78	0.76	0.67	0.55	0.70	0.70	0.76	0.81	0.78	0.75	0.55	0.76	0.76	0.83
PE	0.525	0.607	0.379	0.350	0.140	0.246	0.519	0.316	0.350	0.558	0.351	0.142	0.264	0.637	0.411
TPI	2.07	2.56	1.50	1.41	0.89	1.14	2.04	1.32	1.41	2.25	1.42	0.90	1.18	2.78	1.61
Het	75.90	80.40	66.70	64.60	44.00	56.00	75.50	62.00	64.60	77.80	64.70	44.20	57.70	82.00	68.90

H_o = observed heterozygosity, H_e = expected heterozygosity, HWE = Hardy-Weinberg p value, MP = matching probability, PD = power of discrimination, PIC = polymorphic information content, PE = power of exclusion, TPI = paternity index, Het = percent heterozygote

ผลการวิเคราะห์การจัดกลุ่มแบบ model based clustering ด้วยโปรแกรม STRUCTURE โดยประชากรที่นำมาวิเคราะห์ด้วย STRUCTURE ในครั้งนี้ มีค่า overall F_{st} เท่ากับ 0.017 ดังนั้นจึงต้องอาศัยโมเดล LOCPRIOR ซึ่งถูกสร้างขึ้นมาสำหรับการวิเคราะห์ประชากรที่มีระยะทางพันธุกรรมรวมค่อนข้างน้อย (overall $F_{st} < 0.02$) พบว่าจำนวนกลุ่ม (K) ที่เหมาะสมที่สุด มีค่าเท่ากับ 2 ตามวิธีของ Evanno *et al.* (2005) (Figure 3) โดย Figure 4 แสดงถึงพันธุกรรมที่แตกต่างระหว่าง 2 กลุ่ม (สีเขียวและสีแดง) โดยประชากรที่จัดอยู่ในกลุ่มเดียวกันจะมีสีที่เหมือนกัน จากภาพเห็นได้ว่าชาวส่วยมีสีแดงในสัดส่วนที่มากที่สุด แตกต่างจากประชากรอื่นๆ ที่เหลือซึ่งมีสีเขียว แสดงถึงการมีโครงสร้างทางพันธุกรรมของ

ชาวส่วยที่แตกต่างจากประชากรอื่นมากที่สุด นอกจากนี้พบว่าชาวโล่มีสัดส่วนของสีแดงและเขียวประมาณร้อยละ 34 และ 66 ตามลำดับ แสดงถึงการมีพันธุกรรมที่แตกต่างจากประชากรอื่นรองลงมาจากชาวส่วย อย่างไรก็ตามเนื่องจากประชากรที่ศึกษามี overall F_{st} ค่อนข้างน้อย การอาศัยไมโครแซทเทลไลต์เพียง 15 ตำแหน่ง ในครั้งนี้ ยังไม่สามารถจัดกลุ่มประชากรด้วยวิธี model based clustering จึงทำให้ไม่สามารถเห็นโครงสร้างทางพันธุกรรมที่แท้จริงของประชากร ในอนาคตการศึกษาไมโครแซทเทลไลต์ตำแหน่งที่เพิ่มมากขึ้น ดังเช่นตัวอย่างในการศึกษาก่อนหน้านี้ (Listman *et al.*, 2007) จะสามารถวิเคราะห์โครงสร้างของประชากรอย่างแม่นยำได้มากขึ้น

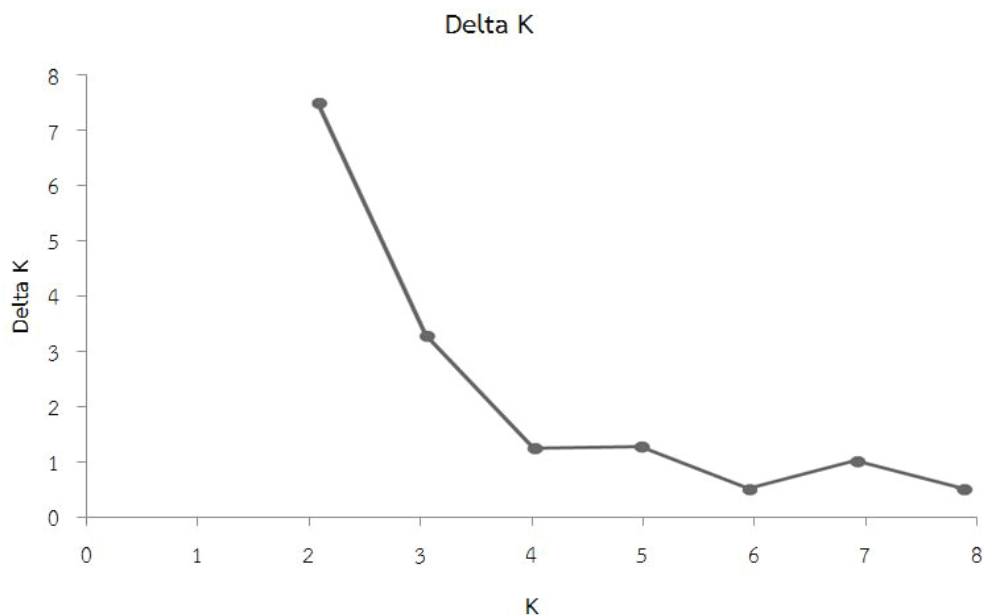


Figure 3 Optimal K in the data was identified by a second order rate of change in logarithmic probability between the subsequent K values (delta K).

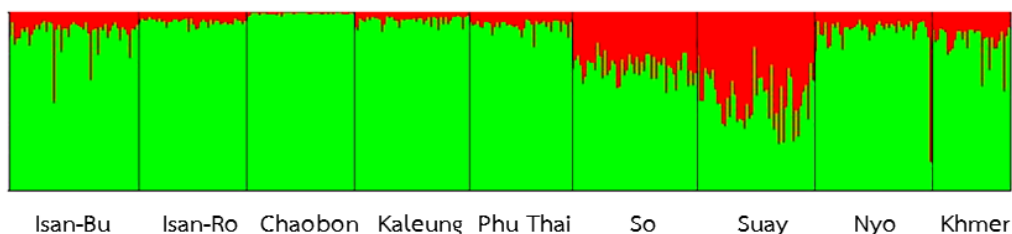


Figure 4 Bar plot estimation of the studied and compared populations, with $K = 2$, inferred from STRUCTURE 2.3 analyses. Each individual is represented by a thin vertical line, which is partitioned into K colored segments that represent the individual's estimated membership fractions in each of the K . r value is 0.8926.

การผสมผสานทางพันธุกรรม

จากผลการศึกษาแผนภูมิต้นไม้แสดงความสัมพันธ์ทางวิวัฒนาการ แบบ NJ (Figure 2) ที่แสดงถึงการมีลักษณะทางพันธุกรรมของประชากรลูกผสมในชาวไทยอีสาน บุรีรัมย์ และร้อยเอ็ด (Isan-Bu และ Isan-Ro) ดังนั้นผู้วิจัยจึงทำการวิเคราะห์การผสมผสานทางพันธุกรรม โดยใช้ประชากรที่ใช้เปรียบเทียบความสัมพันธ์ระหว่างประชากร เป็นประชากรพ่อแม่ (parental population) ดังนั้น เขมร (Khmer) ส่วย (Suay) ชาวนน (Chaobon) และโส้ (So) ภูไท (Phu Thai) ญ้อ (Nyo) และกะเลิง (Kaleung) ผลการศึกษาพบว่าค่า m_v ซึ่งเป็นพารามิเตอร์ที่แสดงถึงสัดส่วนทางพันธุกรรมที่ประชากรลูกผสมได้รับจากประชากรพ่อแม่ มีค่าติดลบในประชากรชาวภูไทและโส้ ผลดังกล่าวแสดงว่าชาวภูไทและโส้ ไม่ได้เป็นประชากรพ่อแม่ของชาวไทยอีสาน บุรีรัมย์ และร้อยเอ็ด ผู้วิจัยจึงได้วิเคราะห์การผสมผสานทางพันธุกรรมอีกครั้งโดยใช้ประชากรพ่อแม่เพียง 5 ประชากร คือ เขมร ส่วย ชาวนน

ญ้อ และกะเลิง ผลการศึกษาพบความแตกต่างในสัดส่วนพันธุกรรมที่ได้รับจากประชากรพ่อแม่ในประชากรชาวไทยอีสานทั้งสองจังหวัด (Figure 5) ชาวไทยอีสานบุรีรัมย์ได้รับสัดส่วนพันธุกรรม ร้อยละ 65 จากชาวนน และกะเลิง (พูดภาษาตระกูลกระได) ซึ่งมากกว่าสัดส่วน ร้อยละ 35 จากชาวนน ส่วย และชาวนน (พูดภาษาตระกูลออสโตรเอเชียติก) ในขณะที่ชาวไทยอีสานร้อยเอ็ดได้รับสัดส่วนพันธุกรรมจากกลุ่มประชากรออสโตรเอเชียติก ร้อยละ 58 ซึ่งมากกว่ากลุ่มประชากรกระได (ร้อยละ 42) เป็นที่น่าสนใจว่าผลการวิเคราะห์การผสมผสานทางพันธุกรรมไม่สอดคล้องกับหลักฐานทางประวัติศาสตร์และโบราณคดี จังหวัดบุรีรัมย์และบางส่วนของจังหวัดร้อยเอ็ดเป็นจังหวัดที่มีอารยธรรมดั้งเดิมของชาวนนและมีประชากรที่พูดภาษาเขมรจำนวนมาก แต่ประชากรดังกล่าวกลับไม่ได้ผสมผสานกับชาวไทยอีสานมากนัก ทั้งนี้อาจเกิดจากชาวนนที่ใช้เปรียบเทียบเป็นชาวนนจากจังหวัดสุรินทร์ ซึ่งอาจมีพันธุกรรมไม่เหมือนกับชาวนนจังหวัดบุรีรัมย์ ดังนั้นในอนาคตการศึกษาพันธุกรรม

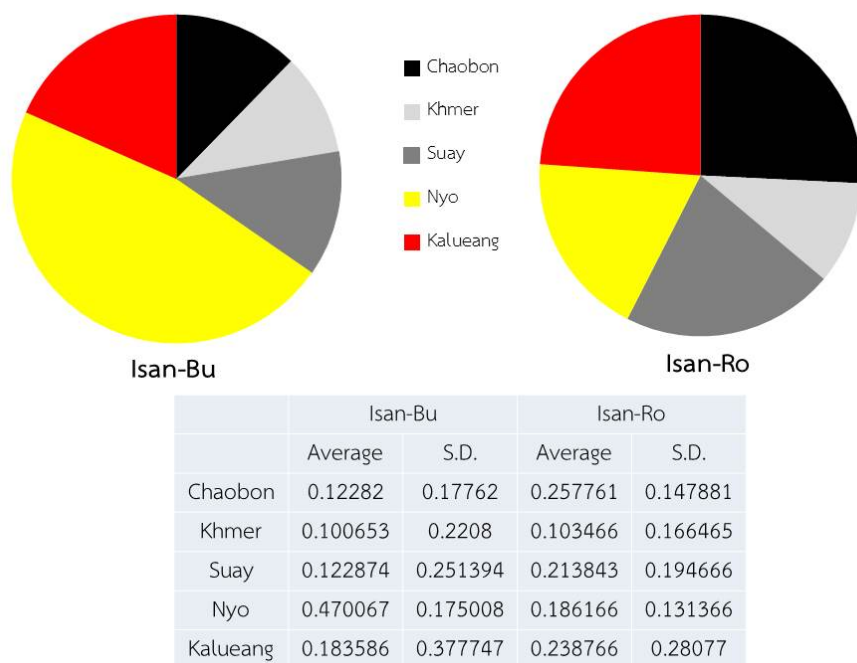


Figure 5 Pie diagrams showing the distribution of Chaobon (black), Khmer (light grey), Suay (dark grey), Nyo (yellow) and Kalueang (red) contributing to the Thai-Isan populations. Table under the pie diagram shows weighted average across loci and standard deviation (s.d.), generated by ADMIX 2, of the estimated contributions of five parental populations to the Thai-Isan populations.

ของชาวเขมรจากหลายพื้นที่ในภาคอีสานตอนใต้จะช่วยให้การวิเคราะห์การผสมผสานทางพันธุกรรมชัดเจนมากยิ่งขึ้น นอกจากนั้นผลการศึกษานี้ไม่สอดคล้องกับงานวิจัยก่อนหน้านี้ที่ศึกษาการผสมผสานทางพันธุกรรมโดยใช้ดีเอ็นเอไมโทคอนเดรียเป็นเครื่องหมายทางพันธุกรรม (Kutanan *et al.*, 2014b) อาจเนื่องมาจากเครื่องหมายทางพันธุกรรมที่เลือกใช้มีรูปแบบการถ่ายทอดที่ไม่เหมือนกัน รวมถึงความแตกต่างของประชากรพ่อแม่ที่ใช้ในการคำนวณ โดยงานวิจัยก่อนหน้านี้ได้ใช้ชาวลาว จากประเทศลาว เป็นประชากรพ่อแม่เนื่องจากชาวไทยอีสานเป็นกลุ่มชาติพันธุ์เดียวกับชาวลาวและได้อพยพเข้าสู่ประเทศไทย เริ่มตั้งแต่ประมาณ 500 ปีที่ผ่านมา (Keyes, 1988) แต่งานวิจัยครั้งนี้ข้อมูลไมโครแซทเทลไลต์ของชาวลาว ยังไม่มีรายงานการศึกษา ดังนั้นการศึกษาพันธุกรรมของชาวลาวในอนาคตจะสามารถประเมินสัดส่วนทางพันธุกรรมที่มีอยู่ในประชากรชาวไทยอีสานได้แม่นยำมากขึ้น

สรุปผลการวิจัย

การวิเคราะห์ความสัมพันธ์ทางพันธุกรรมระหว่างประชากรชาวไทยอีสาน จังหวัดบุรีรัมย์และประชากรอื่นที่อาศัยอยู่ในภาคตะวันออกเฉียงเหนือของประเทศไทย ด้วยชุดไมโครแซทเทลไลต์บนออโทโซมจำนวน 15 ตำแหน่ง พบว่าประชากรชาวไทยอีสาน จังหวัดบุรีรัมย์ มีความใกล้ชิดกับประชากรชาวไทยอีสานจากจังหวัดร้อยเอ็ดมากที่สุด โดยประชากรชาวไทยอีสานจากทั้งสองจังหวัดมีลักษณะพันธุกรรมของประชากรลูกผสมและเมื่อวิเคราะห์ระดับการผสมผสานทางพันธุกรรม พบความแตกต่างสัดส่วนทางพันธุกรรมของทั้งสองประชากร อย่างไรก็ตามการศึกษานี้ยังไม่สามารถทราบโครงสร้างทางพันธุกรรมของประชากรได้มากนักเนื่องจากข้อจำกัดของเครื่องหมายทางพันธุกรรมที่เลือกใช้ เพราะมีเพียง 15 ตำแหน่ง ในอนาคตอาจมีการศึกษาตำแหน่งไมโครแซทเทลไลต์บนออโทโซมที่ใช้ในงานนิติวิทยาศาสตร์เพิ่มเติม เช่น จากชุดตรวจ GlobalFiler PCR Amplification kit หรือการวิเคราะห์จีโนมของพหุนิวคลีโอไทด์เดี่ยว จะช่วยให้การศึกษาโครงสร้างของประชากรในภาคตะวันออกเฉียงเหนือของประเทศไทย มีความชัดเจนมากขึ้น

กิตติกรรมประกาศ

คณะผู้วิจัยขอขอบคุณอาสาสมัครและผู้ใหญ่บ้าน จากบ้านท่างาม ตำบลพุทไธสง อำเภอพุทไธสง จังหวัดบุรีรัมย์ ที่เสียสละเวลาเข้าร่วมงานวิจัย และกลุ่มงานตรวจชีววิทยาและดีเอ็นเอ สำนักงานศูนย์พิสูจน์หลักฐาน 3 จังหวัดนครราชสีมา งานวิจัยนี้ได้รับการสนับสนุนจากมหาวิทยาลัยขอนแก่น ภายใต้โครงการทุนบ่มเพาะนักวิจัย

เอกสารอ้างอิง

- Babiker HM, Schlebusch CM, Hassan HY, Jakobsson M (2011) Genetic variation and population structure of Sudanese populations as indicated by 15 Identifiler sequence-tagged repeat (STR) loci. *Investig Genet* 2: 12.
- Belle EM, Landry PA, Barbujani G (2006) Origins and evolution of the Europeans' genome: evidence from multiple microsatellite loci. *Proc R Soc B Biol Sci* 273: 1595–1602.
- Bertorelle G, Excoffier L (1998) Inferring admixture proportions from molecular data. *Mol Biol Evol* 15: 1298–1311.
- Boonsoda P, Srithawong S, Srikuka S, Kutanan W (2013) Mitochondrial DNA variation of the Khmer in Surin Province. *Thai J Genet* 6: 40–48.
- Crossetti SG, Demarchi DA, Raimann PE, Salzano FM, Hutz MH, Callegari-Jacques SM (2008) Autosomal STR genetic variability in the Gran Chaco native population: Homogeneity or heterogeneity?. *Am J Hum Biol Off Hum Biol Counc* 20: 704–711.
- Di Benedetto G, Ergüven A, Stenico M, Castri L, Bertorelle G, Togan I, Barbujani G (2001) DNA diversity and population admixture in Anatolia. *Am J Phys Anthropol* 115: 144–156.
- Dupanloup I, Bertorelle G (2001) Inferring admixture proportions from molecular data: extension to any number of parental populations. *Mol Biol Evol* 18: 672–675.

- Earl DA, vonHoldt BM (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4: 359–361.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14: 2611–2620.
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10: 564–567.
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.
- Gaibar M, Esteban E, Moral P, Gómez-Gallego F, Santiago C, Bandrés F, Luna F, Fernández-Santander A (2010) STR genetic diversity in a Mediterranean population from the south of the Iberian Peninsula. *Ann Hum Biol* 37: 253–266.
- Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour* 9: 1322–1332.
- Kayser M (2007) Uni-parental marker in parental identity testing including forensic DNA analysis. *Biotechniques* 43(6): S16–S21.
- Keyes CF (1988) Isan: Regionalism in Northeastern Thailand. Dalley Book Service.
- Khanittanan W (2001) Khmero-Thai: The great change in the history of Thai language in the Chao Praya basin. *Journal of Language and Linguistics* 19: 35–50.
- Kraaijenbrink T, van der Gaag KJ, Zuniga SB, Xue Y, Carvalho-Silva DR, Tyler-Smith C, Jobling MA, Parkin EJ, Su B, Shi H, *et al.* (2014) A Linguistically Informed Autosomal STR Survey of Human Populations Residing in the Greater Himalayan Region. *PLoS ONE* 9: e91534.
- Kutanan W, Kampuansai J, Colonna V, Nakbunlung S, Lertvicha P, Seielstad M, Bertorelle G, Kangwanpong D (2011a) Genetic affinity and admixture of northern Thai people along their migration route in northern Thailand: evidence from autosomal STR loci. *J Hum Genet* 56: 130–137.
- Kutanan W, Kampuansai J, Fuselli S, Nakbunlung S, Seielstad M, Bertorelle G, Kangwanpong D (2011b) Genetic structure of the Mon-Khmer speaking groups and their affinity to the neighbouring Tai populations in Northern Thailand. *BMC Genet* 12: 56.
- Kutanan W, Ghirotto S, Bertorelle G, Srithawong S, Srithongdaeng K, Pontham N, Kangwanpong D (2014a) Geography has more influence than language on maternal genetic structure of various northeastern Thai ethnicities. *J Hum Genet* 59: 512–520.
- Kutanan W, Srithawong S, Kamlao A, Kampuansai J (2014b) Mitochondrial DNA-HVR1 Variation Reveals Genetic Heterogeneity in Thai-Isan Peoples from the Lower Region of Northeastern Thailand. *Adv Anthropol* 4: 7–12.
- Lewis MP (2009) *Ethnologue: Languages of the World*. 16th ed. SIL International, Dallas.
- Listman JB, Malison RT, Sughondhabirrom A, Yang BZ, Raaum RL, Thavichachart N, Sanichwankul K, Kranzler HR, Tangwonchai S, Mutirangura A, *et al.* (2007) Demographic changes and marker properties affect detection of human population differentiation. *BMC Genet* 8: 21.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

- Smalley WA (1988) Multilingualism in the Northern Khmer population of Thailand. *Lang Sci* 10: 395–408.
- Smalley WA (1994) Linguistic diversity and national unity: language ecology in Thailand. University of Chicago Press, Chicago.
- Srithawong S, Srikummool M, Pittayaporn P, Ghirotto S, Chantawannakul P, Sun J, Eisenberg A, Chakraborty R, Kutanan W (2015) Genetic and linguistic correlation of the Kra–Dai-speaking groups in Thailand. *J Hum Genet* 60: 371–380.
- Talbot S, Chutima J (2001) Northeast Thailand before Angkor: Evidence from an Archaeological Excavation at the Prasat Hin Phimai. *Asian Perspect* 40: 179–194.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731–2739.
- Vail P (2007) Thailand's Khmer as “invisible minority”: Language, ethnicity and cultural politics in north-eastern Thailand. *Asian Ethn* 8: 111–130.
- Welch DJ (1998) Archaeology of Northeast Thailand in Relation to the Pre-Khmer and Khmer Historical Records. *Int J Hist Archaeol* 2: 205–233.
- Wyatt DK (2003) Thailand: A Short History. 2nd ed. Yale University Press, New Haven.