



ใบรับรองวิทยานิพนธ์
บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

วิทยาศาสตร์มหาบัณฑิต (สัตวศาสตร์)

ปริญญา

สัตวศาสตร์

สัตวบาล

สาขา

ภาควิชา

เรื่อง ความผันแปรทางพันธุกรรมจีโนมสำหรับลักษณะการให้ผลผลิตของโคนมในประชากร
โคนมหลากหลายพันธุ์ในประเทศไทย

Genomic Variation for Dairy Production Traits in a Multibreed Dairy Cattle Population
in Thailand

นามผู้วิจัย นายทวี เหล่าดีมี

ได้พิจารณาเห็นชอบโดย

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(ผู้ช่วยศาสตราจารย์ศกร คุณวุฒิฤทธิธรรม, วท.ด.)

อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม

(Professor Mauricio A. Elzo, Ph.D.)

อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม

(อาจารย์ธนาทิพย์ สุวรรณโสภี, วท.ด.)

หัวหน้าภาควิชา

(รองศาสตราจารย์สมเกียรติ ประสานพานิช, วท.ด.)

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์รับรองแล้ว

(รองศาสตราจารย์กัญจนา วีระกุล, D.Agr.)

คณบดีบัณฑิตวิทยาลัย

วันที่ เดือน พ.ศ.

วิทยานิพนธ์

เรื่อง

ความผันแปรทางพันธุกรรมจีโนมสำหรับลักษณะการให้ผลผลิตของโคนม
ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

Genomic Variation for Dairy Production Traits in a Multibreed
Dairy Cattle Population in Thailand

โดย

นายทวี เหล่าดีม

เสนอ

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

เพื่อความสมบูรณ์แห่งปริญญาวิทยาศาสตรมหาบัณฑิต (สัตวศาสตร์)

พ.ศ. 2558

ทวี เหล่าดี 2558: ความผันแปรทางพันธุกรรมจีโนมสำหรับลักษณะการให้ผลผลิตของโคนม
ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทยปริญญาวิทยาศาสตรมหาบัณฑิต
(สัตวศาสตร์) สาขาสัตวศาสตร์ ภาควิชาสัตวบาล อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก:
ผู้ช่วยศาสตราจารย์ศกร คุณวุฒิฤทธิธ, วท.ค. 83 หน้า

ข้อมูลสปีส์ (8,810 ตำแหน่ง) ของโคนมเพศเมียจำนวน 1,413 ตัว ของเกษตรกรจำนวน 195 ราย ที่ตั้งอยู่ในพื้นที่ภาคกลาง (78 ราย) ภาคใต้ (67 ราย) และภาคตะวันออกเฉียงเหนือ (50 ราย) ของประเทศไทย มีสปีส์ทั้งหมด 8,573 ตำแหน่งที่ผ่านการตรวจสอบคุณภาพของสปีส์ โดยพิจารณาจากสัดส่วนการปรากฏของสปีส์แต่ละตำแหน่งในประชากรโคนม (มากกว่าร้อยละ 90) และมีตำแหน่งทางกายภาพบนโครโมโซมแต่ละคู่ จากนั้น สปีส์ที่ผ่านเกณฑ์คุณภาพถูกนำมาพิจารณาความผันแปรของสปีส์และระดับ Linkage disequilibrium (LD) ในระดับจีโนมในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย สปีส์ทั้งหมดถูกวิเคราะห์การกระจายตัวระหว่างสปีส์ ความหนาแน่นของสปีส์ และค่า Minor allele frequency (MAF) บนโครโมโซมแต่ละคู่ จากนั้น สปีส์ที่ทราบตำแหน่งทางกายภาพบนโครโมโซมร่างกายแต่ละคู่ ที่มีค่า $MAF \geq 0.05$ ถูกนำมาพิจารณาเชื่อมโยงกับข้อมูลแผนที่สปีส์และข้อมูลพันธุ์ประวัติของโคนมด้วยซอฟต์แวร์ Haploview และประมาณค่า LD โดยใช้สองตัวแปรคือ D' และ r^2 ผลการศึกษาชี้ให้เห็นว่า สปีส์ของโคนมไทยหลากหลายพันธุ์มีการกระจายตัวแตกต่างกันบนทั้งโครโมโซมร่างกายและโครโมโซมเพศ และอยู่ร่วมกันเป็นกลุ่มบนพื้นที่บางส่วนของโครโมโซม รวมทั้งการกระจายตัวของ LD บนโครโมโซมร่างกายแต่ละคู่แตกต่างกันขึ้นอยู่กับระยะห่างระหว่างคู่สปีส์ ค่า MAF และขนาดประชากรที่ใช้ในการศึกษาครั้งนี้ ระดับ LD ที่มีค่าสูงบนโครโมโซมร่างกายโดยเฉพาะระหว่างคู่สปีส์ที่มีระยะห่าง 40 ถึง 50 kb ($D' = 0.694 \pm 0.007$ และ $r^2 = 0.202 \pm 0.003$) จำนวนโคนมน้อยที่สุดสำหรับการประมาณค่า LD ในประชากรโคนมไทยหลากหลายพันธุ์ คือ มากกว่า 177 ตัว สำหรับประมาณค่า D' และมากกว่า 89 ตัว สำหรับประมาณ r^2 ผลการศึกษาชี้เป็นนัยถึงความจำเพาะทางพันธุกรรมในระดับจีโนมของโคนมไทยหลากหลายพันธุ์ และความจำเป็นในการศึกษาความสัมพันธ์ทางพันธุกรรมในระดับจีโนม เพื่อหาแนวทางการปรับปรุงศักยภาพการผลิตและพันธุกรรมโคนมไทยในเขตร้อนชื้นต่อไป

ลายมือนิติ

ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

Thawee Laodim 2015: Genomic Variation for Dairy Production Traits in a Multibreed Dairy Cattle Population in Thailand. Master of Science (Animal Science), Major Field: Animal Science, Department of Animal Science. Thesis Advisor: Assistant Professor Skorn Koonawootrittriron, Ph.D. 83 pages.

Single nucleotide polymorphisms (8,810 SNPs) were obtained from 1,413 cows representing 195 farms located in Central (78 farms), Southern (67 farms) and Northeastern (50 farms) regions of Thailand. A total of 8,573 SNPs were chosen based on their call rate ($> 90\%$) and known physical location in each chromosome (autosomes and sex chromosome). The selected SNPs were analyzed in terms of genomic variation and extent of linkage disequilibrium (LD) in a Multibreed dairy cattle population in Thailand. Analyses included distribution of SNPs, distance between SNPs, density of SNPs, and minor allele frequencies (MAF) within chromosomes. The SNPs that had known physical location in autosomes and had $MAF \geq 0.05$ were linked using an SNP map and the pedigree of cows with the Haploview software and LD estimated using parameters D' and r^2 . The results indicated that SNPs of Thai multibreed dairy cattle were unevenly distributed within autosomes and the sex chromosome and they tended to be clustered in some chromosome regions. The distribution of LD in each autosome was affected by distance between pairs of SNPs, MAF, and sample size. High levels of the LD were found in autosomes, particularly between SNP pairs at distances between 40 to 50 kb ($D' = 0.694 \pm 0.007$ and $r^2 = 0.202 \pm 0.003$). The minimum sample size of Thai multibreed dairy cattle required to obtain accurate measurement of LD was more than 177 for D' , and 89 for r^2 . Differences between results obtained here and from other cattle populations support the need to conduct genome-wide association studies in Thai multibreed populations to improve their genetic potential under local environmental conditions.

Student's signature

Thesis Advisor's signature

กิตติกรรมประกาศ

ผู้วิจัยขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.ศกร คุณวุฒิฤทธิธรม อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก Professor Dr. Mauricio A. Elzo และ อาจารย์ ดร.ธนาทิพย์ สุวรรณโสภี อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วมที่ให้ความรู้และคำปรึกษา ทั้งด้านการเรียนและการค้นคว้าวิจัยจนสามารถนำความรู้ที่ได้มาประยุกต์ใช้ในการทำวิทยานิพนธ์ฉบับนี้ และขอขอบคุณ โครงการพัฒนาศักยภาพบุคลากรเพื่อการวิจัยและพัฒนาสำหรับภาคอุตสาหกรรม (NUI-RC) สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ (สวทช.) สำหรับทุนการศึกษาของผู้วิจัย

ผู้วิจัยขอขอบพระคุณ คณาจารย์ทุกท่าน ที่ให้ความรู้และอบรมสั่งสอนอันเป็นประโยชน์ต่อการศึกษาวิจัยและการดำเนินชีวิตประจำวัน ขอขอบคุณสมาชิกหน่วยพันธุศาสตร์สัตว์เขตร้อนชื้น (Tropical Animal Genetic Unit; TAGU) ภาควิชาสัตวบาล คณะเกษตร มหาวิทยาลัยเกษตรศาสตร์ทุกคน ที่ให้ความช่วยเหลือ ให้คำแนะนำ และให้กำลังใจมาโดยตลอด ขอขอบพระคุณ โครงการพัฒนาศักยภาพทางพันธุกรรมและการผลิตโคนมในเขตภาคกลางของประเทศไทย [พ-ท(ช)7.55] สถาบันวิจัยและพัฒนาแห่งมหาวิทยาลัยเกษตรศาสตร์ สำหรับความอนุเคราะห์ข้อมูลเบื้องต้นสำหรับการศึกษาวิจัย และ โครงการพัฒนาระบบประเมินความสามารถทางพันธุกรรมจีโนมโคนมในประเทศไทย [P-11-00116] ซึ่งเป็นการลงทุนร่วมระหว่าง สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ (สวทช.) มหาวิทยาลัยเกษตรศาสตร์ และ องค์การส่งเสริมกิจการโคนมแห่งประเทศไทย (อ.ส.ค.) สำหรับความอนุเคราะห์ข้อมูลความแตกต่างทางพันธุกรรมจีโนมโคนม และงบประมาณหลักในการศึกษาวิจัย และ ขอขอบคุณเกษตรกรผู้เลี้ยงโคนมทุกท่าน สำหรับความอนุเคราะห์ข้อมูลและตัวอย่างเลือดโคนมอันมีค่ายิ่งต่อการศึกษาวิจัยในครั้งนี้

สุดท้ายนี้ ผู้วิจัยขอขอบพระคุณ คุณพ่อ คุณแม่ และสมาชิกทุกคนในครอบครัว ที่คอยให้การสนับสนุน ให้คำปรึกษา และเป็นกำลังใจมาโดยตลอด ทั้งนี้วิทยานิพนธ์ฉบับนี้จะไม่สามารถสำเร็จลงได้เลยหากปราศจากบุคคลทุกท่านดังที่กล่าวมา ความดีหรือประโยชน์อันใดที่เกิดขึ้นจากวิทยานิพนธ์เล่มนี้ผู้วิจัยขอมอบให้บุคคลทุกท่านดังที่กล่าวมาข้างต้น

นายทวี เหล่าดี้ม

พฤศจิกายน 2557

สารบัญ

หน้า

สารบัญ	(1)
สารบัญตาราง	(2)
สารบัญภาพ	(3)
คำอธิบายสัญลักษณ์และคำย่อ	(6)
คำนำ	1
วัตถุประสงค์	3
การตรวจเอกสาร	4
อุปกรณ์และวิธีการ	18
ผลและวิจารณ์	29
สรุปและข้อเสนอแนะ	54
สรุป	54
ข้อเสนอแนะ	55
เอกสารและสิ่งอ้างอิง	56
ภาคผนวก	64
ประวัติการศึกษาและการทำงาน	82

สารบัญตาราง

ตารางที่	หน้า
1 จำนวนสปีชีส์ที่พบในประชากร โคนมหลากหลายพันธุ์ในประเทศไทย	22
2 รูปแบบสปีชีส์จีโนมไทป์ในประชากร โคนมหลากหลายพันธุ์ในประเทศไทย	22
3 จำนวนสปีชีส์ และคู่สปีชีส์ที่ผ่านเกณฑ์พิจารณาในแต่ละกลุ่มของค่า MAF	27
4 จำนวนสปีชีส์ ความหนาแน่นสปีชีส์ ค่าเฉลี่ยตำแหน่งสปีชีส์บนโครโมโซม ส่วนเบี่ยงเบนมาตรฐาน และค่า Minor allele frequency (MAF) บนโครโมโซมร่างกายและโครโมโซมเพศ ในประชากร โคนมหลากหลายพันธุ์ในประเทศไทย	30
5 จำนวนคู่สปีชีส์ ค่าต่ำสุด ค่าสูงสุด ค่าเฉลี่ยของระยะห่างระหว่างสปีชีส์ และค่า LD (D' และ r^2) บนโครโมโซมร่างกายที่ระยะห่างระหว่างสปีชีส์ไม่เกิน 5 Mb ในประชากร โคนมหลากหลายพันธุ์ในประเทศไทย	36
6 จำนวนคู่สปีชีส์ ค่าต่ำสุด ค่าสูงสุด ค่าเฉลี่ยลีสแควร์ ความคลาดเคลื่อนมาตรฐาน ของ LD (D' และ r^2) ที่ระยะห่างระหว่างสปีชีส์บนจีโนมที่แตกต่างกันในประชากร โคนมหลากหลายพันธุ์ในประเทศไทย	42
ตารางผนวกที่	
1 จำนวนคู่สปีชีส์ที่มีระดับ LD สูงบนโครโมโซมร่างกายแต่ละคู่ของโคนม ในประชากร โคนมหลากหลายพันธุ์ในประเทศไทย	80

สารบัญภาพ

ภาพที่	หน้า
1 กระบวนการในการทำไมโครแอเรย์	6
2 การใช้ Photolithography ในการทำให้ดีเอ็นเอติดบนแผ่นกระจก	7
3 เทคนิค Bead arrays โดยใช้เม็ดซิลิกาขนาด 3 ไมโครเมตร ที่สุ่มประกอบลงบนพื้นผิวของ Sentrix array matrix (SAM) หรือ Sentrix BeadChip	8
4 การกระจายตัวของสปีส์ Minor allele frequency ในระดับจีโนมของโคนม	34
5 การกระจายของค่า D' บนระยะห่างระหว่างคู่สปีส์ไม่เกิน 5 Mb บนจีโนมในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย	39
6 การกระจายของค่า r^2 บนช่วงระยะห่างระหว่างสปีส์ไม่เกิน 5 Mb บนจีโนมในประชากรโคนมหลากหลายพันธุ์ ในประเทศไทย	39
7 ค่าเฉลี่ย D' ในกลุ่มสปีส์ที่มีระดับ MAF แตกต่างกัน บนช่วงห่างระหว่างคู่สปีส์ไม่เกิน 100 kb	46
8 ค่าเฉลี่ย D' ในกลุ่มสปีส์ที่มีระดับ MAF แตกต่างกัน บนช่วงห่างระหว่างคู่สปีส์ไม่เกิน 1,000 kb และ 5 Mb	47
9 ค่าเฉลี่ย r^2 ในกลุ่มสปีส์ที่มีระดับ MAF แตกต่างกัน บนช่วงห่างระหว่างคู่สปีส์ไม่เกิน 100 kb	47
10 ค่าเฉลี่ย r^2 ในกลุ่มสปีส์ที่มีระดับ MAF แตกต่างกัน บนช่วงห่างระหว่างคู่สปีส์ไม่เกิน 1,000 kb และ 5 Mb	48
11 ความแตกต่างของระดับ D' ในกลุ่มประชากรที่มีขนาด แตกต่างกัน บนช่วงห่างระหว่างคู่สปีส์ไม่เกิน 100 kb	49
12 ความแตกต่างของระดับ D' ในกลุ่มประชากรที่มีขนาด แตกต่างกัน บนช่วงห่างระหว่างคู่สปีส์ไม่เกิน 1,000 kb และ 5 Mb	49
13 ความแตกต่างของระดับ r^2 ในกลุ่มประชากรที่มีขนาด แตกต่างกัน บนช่วงห่างระหว่างคู่สปีส์ไม่เกิน 100 kb	51
14 ความแตกต่างของระดับ r^2 ในกลุ่มประชากรที่มีขนาด แตกต่างกัน บนช่วงห่างระหว่างคู่สปีส์ไม่เกิน 1,000 kb และ 5 Mb	51

สารบัญภาพ (ต่อ)

ภาพที่	หน้า
15 ความสัมพันธ์ระหว่างค่า LD ในกลุ่มประชากร โคนมทั้งหมด (1,413 ตัว) และประชากรกลุ่มทดสอบ (1,059, 707 และ 354 ตัว)	52
16 ความสัมพันธ์ระหว่างค่า LD ในกลุ่มประชากร โคนมทั้งหมด (1,413 ตัว) และประชากรกลุ่มทดสอบ (177, 89 และ 45 ตัว)	53
ภาพผนวกที่	
1 การกระจายตัวของสปีส์ MAF บนโครโมโซมคู่ที่ 1 ถึง 6 ในประชากร โคนมหลากหลายพันธุ์ในประเทศไทย	65
2 การกระจายตัวของสปีส์ MAF บนโครโมโซมคู่ที่ 7 ถึง 12 ในประชากร โคนมหลากหลายพันธุ์ในประเทศไทย	66
3 การกระจายตัวของสปีส์ MAF บนโครโมโซมคู่ที่ 13 ถึง 18 ในประชากร โคนมหลากหลายพันธุ์ในประเทศไทย	67
4 การกระจายตัวของสปีส์ MAF บนโครโมโซมคู่ที่ 19 ถึง 24 ในประชากร โคนมหลากหลายพันธุ์ในประเทศไทย	68
5 การกระจายตัวของสปีส์ MAF บนโครโมโซมคู่ที่ 25 ถึง X ในประชากร โคนมหลากหลายพันธุ์ในประเทศไทย	69
6 การกระจายตัวของ D' บนโครโมโซมร่างกายคู่ที่ 1 ถึง 6 ในประชากร โคนมหลากหลายพันธุ์ในประเทศไทย	70
7 การกระจายตัวของ D' บนโครโมโซมร่างกายคู่ที่ 7 ถึง 12 ในประชากร โคนมหลากหลายพันธุ์ในประเทศไทย	71
8 การกระจายตัวของ D' บนโครโมโซมร่างกายคู่ที่ 13 ถึง 18 ในประชากร โคนมหลากหลายพันธุ์ในประเทศไทย	72
9 การกระจายตัวของ D' บนโครโมโซมร่างกายคู่ที่ 19 ถึง 24 ในประชากร โคนมหลากหลายพันธุ์ในประเทศไทย	73

สารบัญภาพ (ต่อ)

ภาพผนวกที่	หน้า
10 การกระจายตัวของ D' บนโครโมโซมร่างกายคู่ที่ 25 ถึง 29 ในประชากรโคนม หลากหลายพันธุ์ในประเทศไทย	74
11 การกระจายตัวของ r^2 บนโครโมโซมร่างกายคู่ที่ 1 ถึง 6 ในประชากรโคนม หลากหลายพันธุ์ในประเทศไทย	75
12 การกระจายตัวของ r^2 บนโครโมโซมร่างกายคู่ที่ 7 ถึง 12 ในประชากรโคนม หลากหลายพันธุ์ในประเทศไทย	76
13 การกระจายตัวของ r^2 บนโครโมโซมร่างกายคู่ที่ 13 ถึง 18 ในประชากรโคนม หลากหลายพันธุ์ในประเทศไทย	77
14 การกระจายตัวของ r^2 บนโครโมโซมร่างกายคู่ที่ 19 ถึง 24 ในประชากรโคนม หลากหลายพันธุ์ในประเทศไทย	78
15 การกระจายตัวของ r^2 บนโครโมโซมร่างกายคู่ที่ 25 ถึง 29 ในประชากรโคนม หลากหลายพันธุ์ในประเทศไทย	79

คำอธิบายสัญลักษณ์และคำย่อ

A	=	Adenine
C	=	Cytosine
cDNA	=	Complementary deoxyribonucleic acid
DGAT1	=	Diacylglycerol O-acyltransferase 1
DNA	=	Deoxyribonucleic acid
G	=	Guanine
GEBV	=	Genomic estimated breeding value
GWAS	=	Genome-wide association study
HWE	=	Hardy-Weinberg equilibrium
iSNP	=	Intronic single nucleotide polymorphism
kb	=	Kilo-basepair
LD	=	Linkage disequilibrium
MAF	=	Minor allele frequency
Mb	=	Mega-basepair
mRNA	=	Messenger ribonucleic acid
PCR	=	Polymerase chain reaction
QTL	=	Quantitative trait loci
RNA	=	Ribonucleic acid
rSNP	=	Regulatory single nucleotide polymorphism
SAM	=	Sentrix array matrix
SNP	=	Single nucleotide polymorphism
T	=	Thymine

ความผันแปรทางพันธุกรรมจีโนมสำหรับลักษณะการให้ผลผลิตของโคนม
ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

Genomic Variation for Dairy Production Traits in a Multibreed
Dairy Cattle Population in Thailand

คำนำ

จากการพัฒนาเทคโนโลยีชีววิทยาในระดับโมเลกุลอย่างต่อเนื่อง ตั้งแต่อดีตจนถึงปัจจุบัน ส่งผลให้สามารถจำแนกรูปแบบทางพันธุกรรมในระดับจีโนมของสิ่งมีชีวิตแต่ละชนิดได้ ข้อมูลเหล่านี้เป็นข้อมูลที่แสดงให้เห็นถึง พันธุกรรมที่รับมาจากบรรพบุรุษและถ่ายทอดไปยังสิ่งมีชีวิตรุ่นลูก ข้อมูลเหล่านี้มีความสำคัญอย่างยิ่งในการศึกษาความแตกต่างทางพันธุกรรมของสิ่งมีชีวิต ซึ่งสามารถนำไปสู่ความรู้และความเข้าใจเกี่ยวกับกระบวนการถ่ายทอดทางพันธุกรรม (Brown, 2010; Dziuda, 2010) ซึ่งเป็นประโยชน์อย่างยิ่ง สำหรับการวินิจฉัยโอกาสและความเสี่ยงในเกิดโรคทางพันธุกรรม รวมถึงการวางแผนรักษาผู้ป่วยแต่ละรายที่มีลักษณะการแสดงออกจำเพาะต่อการติดเชื้อโรคที่แตกต่างกัน (Mu *et al.*, 2011) และการเพิ่มประสิทธิภาพและความแม่นยำในการคัดเลือกและปรับปรุงพันธุ์สัตว์ ให้มีพันธุกรรมสำหรับลักษณะที่สนใจโดยตรงตามความต้องการของผู้ผลิต (Womack, 2005)

การเปลี่ยนแปลงของลำดับเบสเพียงหนึ่งตำแหน่งบนสายนิวคลีโอไทด์ของดีเอ็นเอ หรือที่เรียกว่า สนิปส์ (Single nucleotide polymorphism; SNP) อาจมีผลทำให้สิ่งมีชีวิตแต่ละชนิด รวมถึงสิ่งมีชีวิตในกลุ่มเดียวกัน มีลักษณะปรากฏที่แตกต่างกัน (กนกวรรณ และ วรัญญา, 2550) ในปัจจุบัน สนิปส์จึงถูกใช้เป็นตัวแทนในการศึกษาความผันแปรทางพันธุกรรมระดับโมเลกุล เนื่องจากสามารถพบและเกิดขึ้นด้วยความถี่สูง (Bush and Moore, 2012) และ แม้สนิปส์เพียงตำแหน่งเดียวก็อาจมีความสัมพันธ์กับลักษณะปรากฏได้แตกต่างกัน ทั้งนี้เป็นผลจากความสัมพันธ์ของสนิปส์นั้นๆกับ สนิปส์ตำแหน่งอื่นบนโครโมโซมเดียวกัน ซึ่งพิจารณาในรูป Linkage disequilibrium (Brooks, 2003) อย่างไรก็ตาม สนิปส์ที่สัมพันธ์กับลักษณะปรากฏ และมีตำแหน่งอยู่บนพื้นที่ของยีนหรือบางส่วนของยีน ไม่สามารถบ่งบอกหน้าที่ของลำดับเบสที่กำหนดรหัสกรดอะมิโนที่เปลี่ยนไปได้ แต่สนิปส์เหล่านี้สามารถบ่งบอกรูปแบบการถ่ายทอดทางพันธุกรรมที่

สามารถนำไปใช้ประโยชน์ในการพิจารณาความแตกต่างทางพันธุกรรมได้ สนิปส์ถูกใช้ประโยชน์ในการระบุพื้นที่ทางกายภาพบนโครโมโซมแต่ละคู่ เพื่อค้นหาพื้นที่ยีนบนโครโมโซมที่มีความสัมพันธ์กับลักษณะที่ปรากฏแตกต่างกัน หรือค้นหาตำแหน่งของยีนที่ควบคุมลักษณะเชิงปริมาณ (Quantitative trait loci; QTL; Fan *et al.*, 2011) ซึ่งกระบวนการดังกล่าวอาศัยความสัมพันธ์ของสนิปส์ในรูปแบบ Linkage disequilibrium (LD) เพื่อใช้ในการสร้างแผนที่ความสัมพันธ์ระหว่างตำแหน่งสนิปส์และลักษณะเชิงปริมาณบนโครโมโซมแต่ละคู่ (Kolbehdari *et al.*, 2009)

การวัดระดับ LD เป็นการประมาณค่าจากความถี่อัลลีลของสนิปส์สองตำแหน่งด้วยสองตัวแปรได้แก่ D' และ r^2 ซึ่งมีค่าอยู่ในช่วง 0 ถึง 1 การแปลความหมายทั้งสองตัวแปรมีลักษณะแตกต่างกันเพียงเล็กน้อย ในกรณีที่ D' มีค่าเท่ากับ 1 หมายถึง ตำแหน่งสนิปส์ทั้งสองมีการถ่ายทอดร่วมกัน แม้เกิดการกลายพันธุ์ (Mutation) และการแลกเปลี่ยนชิ้นส่วนโครโมโซม (Crossing over) ระหว่างการรวมตัวอีกครั้งระหว่างพันธุกรรม (Recombination) ส่วนกรณีที่ D' มีค่าน้อยกว่า 1 หมายถึง สนิปส์สองตำแหน่ง อาจเกิดการหยุดชะงักในบางช่วงของการรวมตัวอีกครั้งระหว่างพันธุกรรม (Ardlie *et al.*, 2002) อย่างไรก็ตาม ในปัจจุบันพบว่า ถึงแม้ D' มีค่าเท่ากับ 1 ก็ไม่สามารถบ่งบอกความสามารถในการถ่ายทอดร่วมกันของสนิปส์ได้ เนื่องจากการประมาณค่า D' ขึ้นอยู่กับความถี่อัลลีลของสนิปส์ และจำนวนสิ่งมีชีวิตประชากรที่ใช้ศึกษา (Ardlie *et al.*, 2002; Bohmanova *et al.*, 2010) ดังนั้น จึงได้พิจารณาค่า r^2 ซึ่งเป็นค่าที่บ่งบอกความสัมพันธ์ระหว่างสนิปส์สองตำแหน่ง โดยกรณีที่ r^2 มีค่าเท่ากับ 1 หมายถึง ตำแหน่งสนิปส์สองตำแหน่งมีการถ่ายทอดร่วมกันเมื่อเกิดการรวมตัวอีกครั้งระหว่างพันธุกรรม และมีความถี่อัลลีลของสนิปส์เหมือนกัน นอกจากนี้ r^2 เป็นตัวแปรที่สำคัญในการศึกษาความสัมพันธ์ระหว่าง QTL และ สนิปส์ เนื่องจาก r^2 เป็นความสัมพันธ์อย่างง่ายระหว่างสนิปส์ และขนาดตัวอย่างที่ใช้ในการตรวจสอบขนาดเล็ก (Wall and Pritchard, 2003) นอกจากนี้ r^2 มีส่วนช่วยในการเพิ่มความแม่นยำสำหรับการประมาณค่าความสามารถทางพันธุกรรมของสัตว์เศรษฐกิจ โดยเฉพาะในโคนม (Hayes *et al.*, 2009a) ดังนั้น การศึกษาความผันแปรทางพันธุกรรมและการประมาณค่า LD ของสนิปส์ในระดับจีโนมของโคนประชากรโคนมหลากหลายพันธุ์ในประเทศไทย อาจช่วยให้ได้ข้อมูลที่เป็นประโยชน์ต่อการเพิ่มความแม่นยำในการประเมินความสามารถทางพันธุกรรม ช่วยให้การคัดเลือกสัตว์พันธุ์มีประสิทธิภาพมากขึ้นและมีผลทำให้ศักยภาพทางพันธุกรรม และการให้ผลผลิตของโคนมไทยสูงขึ้น

วัตถุประสงค์

1. เพื่อศึกษารูปแบบความผันแปร และประมาณค่า Linkage disequilibrium ของสปีส์ในระดับจีโนม ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย
2. เพื่อศึกษาปัจจัยที่มีผลต่อการประมาณค่า Linkage disequilibrium ของสปีส์ในระดับจีโนม ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย



การตรวจเอกสาร

1. จีโนม

การศึกษาจีโนมเริ่มต้นในทศวรรษที่ 21 โดยการศึกษาชีววิทยาในระดับโมเลกุลของเซลล์สิ่งมีชีวิต และพัฒนาองค์ความรู้ด้านชีววิทยาในระดับโมเลกุลอย่างต่อเนื่อง นำไปสู่การศึกษาลำดับเบสบนสายนิวคลีโอไทด์ในระดับจีโนมของสิ่งมีชีวิต (Genome sequence) ในปี ค.ศ. 1990 และประสบความสำเร็จเป็นครั้งแรกในสิ่งมีชีวิตจำพวก Free-living โดยแบคทีเรียจำพวก *Haemophilus influenza* ในปี ค.ศ. 1995 ต่อมาได้พัฒนาการศึกษาในยีสต์ แมลงวันผลไม้ หนอนตัวกลมในตระกูล *Caenorhabditis elegans* พืช คน และสัตว์เศรษฐกิจในปัจจุบัน โดยการศึกษาลำดับเบสบนสายนิวคลีโอไทด์ในระดับจีโนมของสิ่งมีชีวิต นำไปสู่การค้นพบตำแหน่งเบสที่เปลี่ยนแปลงไปบริเวณสายนิวคลีโอไทด์ใหม่บนดีเอ็นเอได้อย่างอิสระ หรือ เรียกว่า Post-genomics หรือ Functional genomics ซึ่งศึกษาโดยใช้ฐานข้อมูลลำดับเบสบนสายนิวคลีโอไทด์ที่เคยศึกษามาก่อนเป็นพื้นฐาน และใช้ระบบปฏิบัติการคอมพิวเตอร์ในการถอดรหัสพันธุกรรมจีโนม (Genome annotation) โดยใช้ข้อมูลยีนและคุณสมบัติของยีนในการระบุรหัสทางพันธุกรรมจีโนม นอกจากนี้ Post-genomics ยังครอบคลุมไปถึงเทคนิคในการออกแบบเครื่องมือสำหรับระบุการแสดงออกของยีนที่มีความจำเพาะต่อเซลล์ หรือเนื้อเยื่อสิ่งมีชีวิต รวมถึงรูปแบบการแสดงออกของยีนบนจีโนมที่เปลี่ยนแปลงไปเมื่อเวลาผ่านไป (Brown, 2010)

จีโนม (Genome) คือ ชุดข้อมูลพันธุกรรมทั้งหมดที่เป็นองค์ประกอบภายในเซลล์สิ่งมีชีวิต ประกอบด้วยนิวเคลียสจีโนม (Nuclear genome) และไมโทคอนเดรียจีโนม (Mitochondrial genome) ในการวิเคราะห์การแสดงออกของยีนและการศึกษาในระดับจีโนมนิยมใช้นิวเคลียสจีโนมเนื่องจากมีลำดับเบสของดีเอ็นเอ (DNA sequence) ที่สมบูรณ์อยู่บนชุดโครโมโซม และนิวเคลียสเซลล์ของสัตว์เลี้ยงลูกด้วยนมอยู่ในเซลล์เม็ดเลือดขาว โดยนิวเคลียสจีโนมในสิ่งมีชีวิตมีโมเลกุลของดีเอ็นเออยู่บนโครโมโซมร่างกาย (Autosome) และ โครโมโซมเพศ (Sex chromosome) และมีลำดับเบสในนิวคลีโอไทด์ที่มีลักษณะเหมือนกัน 99.9 เปอร์เซ็นต์ (กรณีที่มีความสัมพันธ์กันทางเครือญาติ) อย่างไรก็ตาม 0.1 เปอร์เซ็นต์ของลำดับเบสในนิวคลีโอไทด์ที่แตกต่างกัน (30,000 ล้านเบส) ส่งผลให้เกิดความแตกต่างในสิ่งมีชีวิต (Dziuda, 2010) โดยหน่วยพื้นฐานที่ทำหน้าที่เป็นตัวกำหนดลักษณะของสิ่งมีชีวิต คือ กรดนิวคลีอิก (Nucleic acid) ซึ่งมีลักษณะเป็นพอลิเมอร์ของมอนอเมอร์โมเลกุลนิวคลีโอไทด์ ที่เชื่อมต่อเรียงตัวกันเป็นสายอย่างเป็นลำดับจนได้สายพอลิ-

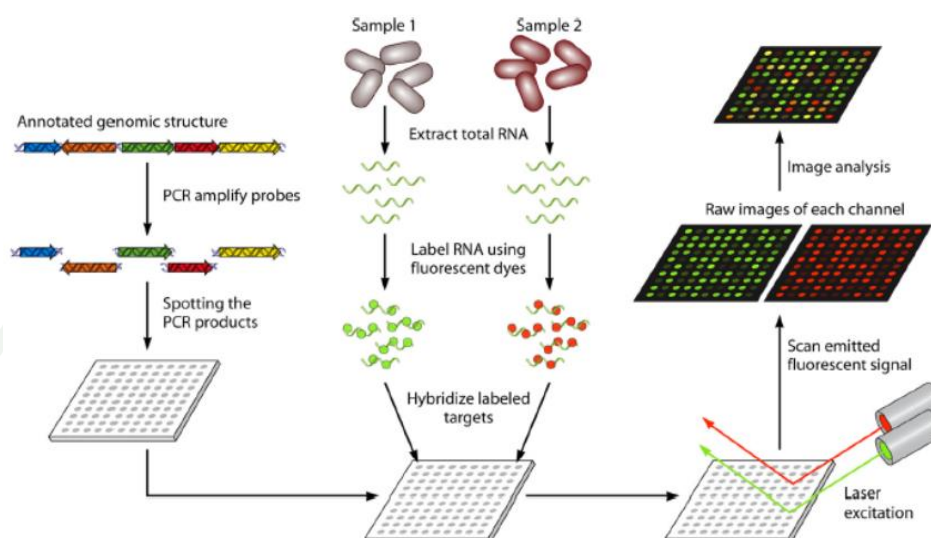
นิวคลีโอไทด์ที่มากถึง 100 ล้านนิวคลีโอไทด์ในหนึ่งโมเลกุลกรดนิวคลีอิก โดยมอนอเมอร์ของนิวคลีโอไทด์แต่ละตัวจะเชื่อมต่อกันด้วยพันธะฟอสโฟไดเอสเทอร์ (Phosphodiester bond) ซึ่งหมู่ฟอสเฟตภายในพันธะจะเชื่อมโยงระหว่างหมู่ไฮดรอกซิลที่ตำแหน่ง 5' ของนิวคลีโอไทด์กับหมู่ไฮดรอกซิลที่ตำแหน่ง 3' ของนิวคลีโอไทด์โมเลกุลถัดไป ทำให้นิวคลีโอไทด์มีโครงสร้างสันหลังของแต่ละเกลียวเป็นฟอสเฟตและน้ำตาล และมีแขนงภายในสายเกลียวเป็นเบส (Packer *et al.*, 2000)

เบสที่ถูกใช้เป็นโครงสร้างดีเอ็นเอประกอบด้วย 4 ชนิด ได้แก่ อะดีนีน (A; adenine) ไซโทซีน (C; cytosine) ไทมีน (T; thymine) และกวานีน (G; guanine) โดยเบสทั้ง 4 ชนิดมีคุณสมบัติทางเคมีที่คล้ายคลึงกัน ยกเว้น ตัวรับและตัวให้พันธะไฮโดรเจนที่อยู่โดยรอบของเบสแต่ละตัว การเข้าคู่กันของเบสมี 2 ลักษณะตามวิธีของ Watson และ Crick (1953) ได้แก่ คู่เบสระหว่างเบสอะดีนีนกับเบสไทมีน (A กับ T) และ คู่เบสระหว่างเบสไซโทซีนกับเบสกวานีน (C กับ G) โดยการเข้าคู่ในลักษณะดังกล่าวส่งผลให้ดีเอ็นเอมีโครงสร้างเป็นเกลียวคู่ และเรียงลำดับสลับไปมาบนสายดีเอ็นเอทำให้เกิดสายพอลินิวคลีโอไทด์ ซึ่งทำหน้าที่เสมือนรหัสคำสั่งที่ใช้ในกระบวนการชีวภาพของข้อมูลทางพันธุกรรมของสิ่งมีชีวิตที่ถูกถอดรหัส แปลรหัส และส่งผ่านรหัสอย่างจำเพาะจากสายเกลียวหนึ่งไปสู่อีกสายเกลียวหนึ่ง โดยการทำงานของเอนไซม์หลายชนิดซึ่งเกิดขึ้นในกระบวนการจำลองรหัสดีเอ็นเอ (DNA replication) จากนั้น รหัสข้อมูลดีเอ็นเอจะถูกส่งผ่านไปสู่อาร์เอ็นเอ (RNA) ในกระบวนการแปลรหัสดีเอ็นเอ (DNA translation) ให้อยู่ในรูปเอ็มอาร์เอ็นเอ (mRNA) ผ่านไปยังไรโบโซมที่ทำหน้าที่นำกรดอะมิโนมาสังเคราะห์ตามรหัสดีเอ็นเอเป็น โปรตีนสำหรับนำไปใช้ในกระบวนการพื้นฐานต่างๆ ในร่างกายของสิ่งมีชีวิต (Sinden, 1994)

2. เทคนิคการค้นหาลำดับเบสบนสายนิวคลีโอไทด์ในระดับจีโนม

ด้วยความรู้ด้านชีววิทยาในระดับโมเลกุลในปัจจุบัน ส่งผลให้สามารถทราบหน้าที่ของยีนทั้งนี้ ยีนแต่ละยีนมีบทบาทสำคัญในการควบคุมการทำงานของเซลล์ของสิ่งมีชีวิต และควบคุมกระบวนการสร้างโปรตีนต่างๆ เช่น โปรตีนที่ทำหน้าที่เป็นโครงสร้างของเซลล์ และเอนไซม์ เป็นต้น ในกระบวนการสร้างโปรตีนเริ่มต้นจากกระบวนการจำลองรหัสดีเอ็นเอ จากนั้นรหัสของดีเอ็นเอถูกขนส่งออกนอกนิวเคลียสในรูปเอ็มอาร์เอ็นเอ หลังจากนั้นจึงเกิดกระบวนการแปลรหัสดีเอ็นเอเพื่อสร้างโปรตีน ดังนั้น การตรวจสอบการแสดงออกของยีนในการควบคุมกระบวนการสร้างโปรตีน จึงใช้วิธีการตรวจสอบจากเอ็มอาร์เอ็นเอ (RNA) และจีโนม เช่น การกลายพันธุ์ของ

ดีเอ็นเอ และความแตกต่างของเบสบนโครโมโซมได้ด้วยเทคนิคไมโครแอเรย์ (Microarray; Miller and Tang, 2009)

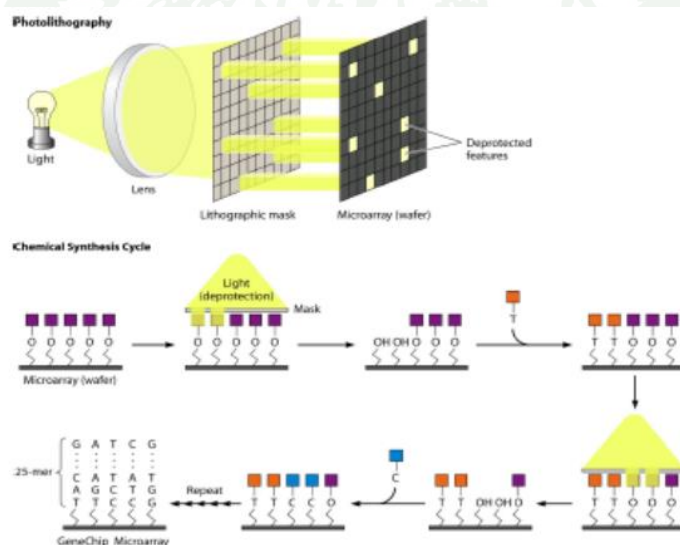


ภาพที่ 1 กระบวนการในการทำไมโครแอเรย์

ที่มา: Miller and Tang (2009)

ไมโครแอเรย์ เป็นกระบวนการที่อาศัยการจับกันอย่างจำเพาะของสายดีเอ็นเอสองสายที่เป็นคู่กันโดยใช้แผ่นไมโครแอเรย์ (Microarray slide) ซึ่งมีลักษณะเป็นแผ่นกระจกที่มีคุณสมบัติยอมให้ดีเอ็นเอสามารถเกาะติดบริเวณผิวหน้า และสามารถบรรจุชิ้นส่วนดีเอ็นเอได้เป็นจำนวนมากถึง 30,000 ชิ้น (วรพรรณ, 2549) ซึ่งเมื่อนำคอมพลีเมนต์อาร์ดีเอ็นเอ (Complementary DNA; cDNA) ที่สังเคราะห์จากอาร์เอ็นเอ หรือ โอลิโกนิวคลีโอไทด์ (Oligonucleotides) มาทำปฏิกิริยาเพื่อให้เกิดการจับกันของเบสคู่สมระหว่างดีเอ็นเออ้างอิงและดีเอ็นเอตัวอย่างที่ศึกษา (Hybridization) โดยกรณีที่ดีเอ็นเอมีลำดับนิวคลีโอไทด์คู่กัน จะสามารถจับกับชิ้นดีเอ็นเอที่ถูกติดอยู่บนแผ่นไมโครแอเรย์ ส่วนดีเอ็นเอที่มีลำดับนิวคลีโอไทด์แตกต่างกันจะไม่สามารถจับกับชิ้นดีเอ็นเอที่ถูกติดบนแผ่นไมโครแอเรย์ จากนั้นแผ่นไมโครแอเรย์ถูกติดฉลากแถบสีเรืองแสง ก่อนนำเข้าเครื่องตรวจวัดสีที่เกิดจากเรืองแสงของดีเอ็นเอ และวัดความเข้มของแถบสี ซึ่งเป็นตัวบ่งบอกความแตกต่างของรูปแบบเบสบนสายนิวคลีโอไทด์ของดีเอ็นเอ และวัดได้โดยใช้โปรแกรมด้านชีวสารสนเทศศาสตร์ (Fadiel and Naftolin, 2003; ภาพที่ 1)

การทำไมโครแอเรย์ ประกอบด้วย 2 ขั้นตอน คือ การเตรียมแผ่นไมโครแอเรย์ และโพรบที่ใช้ในการตรวจสอบดีเอ็นเอ โดยการเตรียมดีเอ็นเอบนแผ่นไมโครแอเรย์เป็นกระบวนการที่ใช้การโคลน (Clone) ยีน หรือ ดีเอ็นเอ ด้วยเทคนิคการเพิ่มชิ้นส่วนดีเอ็นเอ (Polymerase chain reaction; PCR) ก่อนพิมพ์ลงแผ่นกระจก โดยใช้ Photolithography ซึ่งเป็นการใช้แสงช่วยให้ชิ้นส่วนดีเอ็นเอติดบนแผ่นกระจกด้วยพันธะโคเวเลนต์ (Covalent bond; ภาพที่ 2) ส่วนการออกแบบโพรบ ซึ่งมีลักษณะจำเพาะกับการนำไปใช้ประโยชน์ เช่น การออกแบบโพรบเพื่อใช้ตรวจสอบระดับการแสดงออกของยีนซึ่งใช้ลำดับเบสในบริเวณปลาย 3' ก่อนถึงตำแหน่ง Poly (A) หรือการออกแบบเพื่อวิเคราะห์ Splicing ของกระบวนการจำลองรหัสดีเอ็นเอซึ่งใช้ลำดับเบสเฉพาะในบริเวณเอ็กซอน (Exon) ที่เรียกว่า เอ็กซอนแอเรย์ (Exon array) หรือ การออกแบบโพรบเพื่อใช้ตรวจสอบดีเอ็นเอที่สามารถครอบคลุมพื้นที่ทั้งจีโนม (Miller and Tang, 2009)

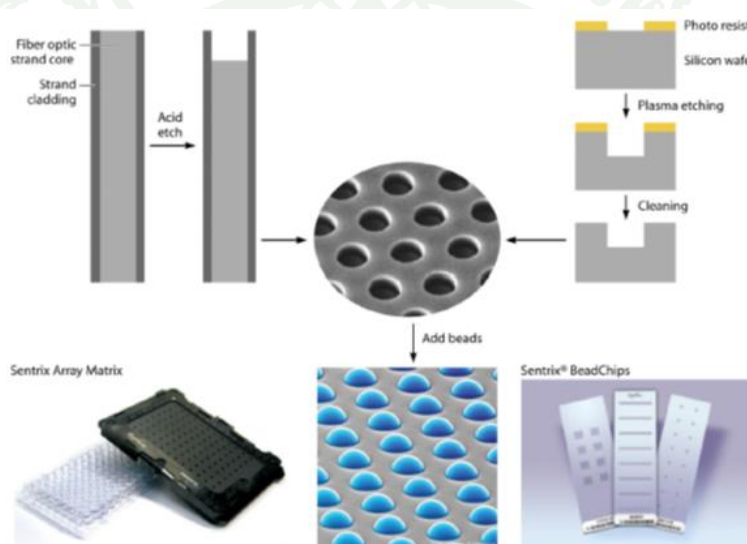


ภาพที่ 2 การใช้ Photolithography ในการทำให้ดีเอ็นเอติดบนแผ่นกระจก

ที่มา: Miller and Tang (2009)

Bead arrays มีลักษณะคล้ายคลึงกับการพิมพ์ไมโครแอเรย์ และ Situ-hybridized ไมโครแอเรย์ โดยมีสารตั้งต้น (Substrate) ที่จำเพาะกับการตรวจสอบความหนาแน่นของกรดนิวคลีอิก และเปลี่ยนการใช้แผ่นกระจก (Glass slides) มาใช้ซิลิคอนเวฟเฟอร์ (Silicon wafers) สำหรับพื้นผิวในการยึดชิ้นส่วนดีเอ็นเออ้างอิง โดย Bead arrays ใช้เม็ดซิลิกาที่มีขนาด 3 ไมโครเมตร และสุ่มประกอบลงบนพื้นผิวของ Sentrrix array matrix (SAM) หรือ Sentrrix BeadChip (ภาพที่ 3) โดย

SAM ประกอบด้วยหลุมจำนวน 96 หลุม แต่ละหลุมมีขนาด 1.4 มิลลิเมตร และกลุ่มใยแก้วนำแสง โดยแต่ละกลุ่มประกอบด้วยแอเรย์ที่มีลักษณะเฉพาะประมาณ 50,000 Light-conducting fibers ขนาด 5 มิลลิเมตร ซึ่งจะถูกกักตรอนโดยใช้สารเคมีในการสร้าง Microwell สำหรับ Single bead นอกจากนี้ Bead array ยังประกอบด้วยเม็ด Bead มากกว่า 1,536 ชนิด โดยแต่ละชนิดจะมีคุณสมบัติในการจับที่ไม่เหมือนกัน ดังนั้น การใช้ Beadchips จึงเหมาะสำหรับการวิเคราะห์ตัวอย่างที่มีความหนาแน่นสูง เช่น การศึกษาจีโนมไทป์ทั้งจีโนม ซึ่งต้องใช้ลำดับเบสบนสายนิวคลีโอไทด์ยาวประมาณ 105 ถึง 106 คู่เบสสำหรับการพิจารณาสนิปส์ทั้งจีโนม (Miller and Tang, 2009)



ภาพที่ 3 เทคนิค Bead arrays โดยใช้เม็ดซิลิกาขนาด 3 ไมโครเมตร ที่สุมประกอบลงบนพื้นผิวของ Sentry array matrix (SAM) หรือ Sentry BeadChip

ที่มา: Miller and Tang (2009)

อย่างไรก็ตาม ในปัจจุบันเทคนิคดังกล่าวถูกพัฒนาจนสามารถประยุกต์ใช้ในสัตว์เศรษฐกิจ เช่นบริษัท Illumina® ได้พัฒนาผลิตภัณฑ์ในการกำหนดสนิปส์จีโนมไทป์ของโคนม เช่น BovineHD Genotyping Beadchip ซึ่งสามารถตรวจสนิปส์ได้มากกว่า 777,000 ตำแหน่ง และสัตว์เศรษฐกิจชนิดอื่นๆ ข้อมูลที่ได้จากเทคโนโลยีดังกล่าวเหมาะสำหรับการสร้างประชากรอ้างอิง เพื่อพิจารณาความสัมพันธ์ในระดับจีโนม ศึกษาความผันแปรทางพันธุกรรม จัดทำแผนที่ยีน และอื่นๆ สำหรับใช้ประโยชน์ในการทำนายความสามารถทางพันธุกรรม และพัฒนาเครื่องมือสำหรับการปรับปรุงพันธุ์ (Illumina, 2013)

3. ความผันแปรทางพันธุกรรมในระดับจีโนม

ความผันแปรทางพันธุกรรมในระดับจีโนม มีผลทำให้เกิดความแตกต่างระหว่างสิ่งมีชีวิต ซึ่งเป็นผลมาจากลำดับเบสที่เปลี่ยนแปลงไปบนสายนิวคลีโอไทด์ (Schrooten, 2004) ซึ่งเกิดขึ้นในระหว่างกระบวนการสร้างเซลล์สืบพันธุ์ (Pryce *et al.*, 2010; Conrad *et al.*, 2012) ด้วยเหตุนี้ ความแตกต่างของลำดับเบสดังกล่าว จึงถูกนำมาใช้ประโยชน์ในศึกษาความผันแปรทางพันธุกรรมของประชากรสิ่งมีชีวิต โดยพิจารณาผ่านข้อมูลสปีส์ซึ่งเป็นรูปแบบความแตกต่างทางพันธุกรรมที่สามารถพบและเกิดขึ้นด้วยความถี่สูง (Frazer *et al.*, 2009; Bush and Moore, 2012)

3.1 สนิปส์ (Single nucleotide polymorphism; SNP)

สนิปส์ คือ ความแตกต่างของเบสเพียงหนึ่งตำแหน่งบนสายนิวคลีโอไทด์ของดีเอ็นเอ ซึ่งมีความสัมพันธ์กับการทำงานและการควบคุมการแสดงออกของโปรตีน (กนกวรรณ และ วรรณญา, 2550) โดยเฉพาะ Synonymous SNP ซึ่งเป็นตำแหน่งสปีส์ที่ทำให้เกิดการเปลี่ยนแปลงลำดับเบสของกรดอะมิโนที่ถูกถอดรหัสเป็นโปรตีน นอกจากนี้ สนิปส์ยังมีความสัมพันธ์กับ สนิปส์ ในตำแหน่งอื่นๆ เนื่องจากสปีส์มีความสัมพันธ์ในรูป LD ซึ่งมีผลทำให้สปีส์ที่พบบริเวณ เอ็กซอน อินทรอน และบริเวณสายนิวคลีโอไทด์ที่ไม่มีการถอดรหัส มีความสัมพันธ์กับลักษณะปรากฏแม้เพียงตำแหน่งเดียวก็อาจมีผลโดยตรงต่อลักษณะที่ปรากฏได้ (Brooks, 2003) โดยสปีส์สามารถจำแนกได้ 2 ประเภทตามบริเวณที่พบสปีส์คือ

3.1.1 Linked SNPs เป็นสปีส์ที่เกิดขึ้นบนดีเอ็นเอบริเวณที่ไม่ใช่ยีนจึงไม่ส่งผลต่อการแสดงออกของโปรตีน

3.1.2 Causative SNP เป็นสปีส์ที่เกิดขึ้นบนดีเอ็นเอบริเวณยีนและบริเวณที่ทำหน้าที่ควบคุมการแสดงออกของยีน โดยสปีส์ในกลุ่มนี้มีผลต่อการแสดงออกของโปรตีน ซึ่งสามารถแบ่งกลุ่มสปีส์ในกลุ่มนี้ได้ 2 ประเภท คือ

ก) สนิปส์ที่เกิดขึ้นบริเวณที่ไม่มีการถอดรหัส (Non-coding SNP) บนสายนิวคลีโอไทด์ โดยสปีส์กลุ่มนี้ประกอบด้วย 1) Regulatory SNP (rSNP) เป็นสปีส์ที่เกิดขึ้นบริเวณโปรโมเตอร์ (Promoter region) มีความสำคัญในการควบคุมกระบวนการแสดงออกของยีนในการ

สร้างโปรตีน 2) สนิปส์ที่เกิดขึ้นบริเวณรอยต่อระหว่าง เอ็กซอน และ อินทรอน หรือที่เรียกว่า สไปลซิงไซต์ มีผลทำให้การตัดต่ออาร์เอ็นเอผิดปกติ ซึ่งทำให้เกิดความผิดปกติของจำนวน กรดอะมิโนบนสายพอลิเปปไทด์ และ 3) Intronic SNP (iSNP) เป็นสนิปส์ที่เกิดขึ้นบริเวณอินทรอน ซึ่งไม่ส่งผลทำให้เกิดการเปลี่ยนแปลงข้อมูลทางพันธุกรรม

ข) สนิปส์ที่เกิดขึ้นบริเวณที่มีการถอดรหัส (Coding SNP) บนสาย นิวคลีโอไทด์ โดยสนิปส์ในกลุ่มนี้เกี่ยวข้องกับลำดับเบสที่ใช้ในกระบวนการแปลรหัสกรดอะมิโน (Triplet codon) ซึ่งมีผลทำให้เกิดการเปลี่ยนแปลงชนิดของกรดอะมิโน (Synonymous SNP) และ ไม่ก่อให้เกิดการเปลี่ยนแปลงชนิดของกรดอะมิโน (Non-synonymous SNP)

3.2 การใช้ประโยชน์จากความผันแปรทางพันธุกรรมจีโนม

การศึกษาความแตกต่างทางพันธุกรรมในระดับจีโนม นำไปสู่การค้นพบความ หลากหลายทางพันธุกรรมของสิ่งมีชีวิต โดยเฉพาะสนิปส์ ซึ่งปัจจุบันถูกนำมาประยุกต์ใช้ด้าน การแพทย์สำหรับการวินิจฉัยโรคของมนุษย์ อาทิ โครงการจีโนมมนุษย์ (Mu *et al.*, 2011) โดยมี เป้าหมายในการใช้ประโยชน์จากสนิปส์ ในการค้นหาพื้นที่ที่เกี่ยวข้องกับการเกิดโรคทางพันธุกรรม เพื่อให้เกิดความรู้และเข้าใจในกระบวนการเกิดโรค รวมทั้งความเสี่ยงต่อการเกิดโรคทางพันธุกรรม ที่จะนำไปสู่การค้นหาวิธีในการป้องกัน และการรักษาโรคที่ถูกต้องและเหมาะสมกับผู้ป่วยในแต่ละ ราย โดยใช้หลักการความผันแปรทางพันธุกรรมของยีนที่มีหลากหลายในกลุ่มประชากร มาทำนาย อัตราการเกิดโรคทางพันธุกรรม เช่น กลุ่มประชากรที่มีความผันแปรทางพันธุกรรมเหมือนกันจะมี แนวโน้มหรือความเสี่ยงต่อการเกิดโรคที่คล้ายคลึงกัน

ด้านการผลิตปศุสัตว์ พบว่า การศึกษาความผันแปรทางพันธุกรรมในระดับจีโนม นำไปสู่การทำแผนที่ยีน (Gene mapping) ยีนเครื่องหมาย (Gene marker) และ Candidate gene ของ ลักษณะที่สำคัญทางเศรษฐกิจ ผ่านการศึกษาความสัมพันธ์ระหว่างข้อมูลทางพันธุกรรมกับลักษณะ การให้ผลผลิต (Genome-wide association study; GWAS) ซึ่งเป็นการตรวจสอบข้อมูลทาง พันธุกรรมในระดับจีโนม เพื่อค้นหาตำแหน่งของความแปรปรวนทางพันธุกรรมที่มีความสัมพันธ์ กับการแสดงออกลักษณะสำคัญทางเศรษฐกิจ โดยใช้เครื่องหมายทางพันธุกรรมสนิปส์ (Cole *et al.*, 2011; Chamberlain *et al.*, 2012; Meredith *et al.*, 2012) ซึ่งความสัมพันธ์เหล่านี้เป็นตัวกำหนด ตำแหน่งที่ควบคุมลักษณะเชิงปริมาณ (Quantitative trait loci; QTL) บนโครโมโซมแต่ละคู่

(Chamberlain *et al.*, 2012) และในปัจจุบันความสัมพันธ์ดังกล่าวได้ถูกนำมาประยุกต์ใช้ในการปรับปรุงพันธุ์สัตว์โดยเฉพาะในโคนม ซึ่งเป็นการใช้ประโยชน์จากเครื่องหมายทางพันธุกรรม สนิปส์เพื่อเพิ่มความแม่นยำในการทำนายค่าการผสมพันธุ์ รวมถึงการทำความเข้าใจในกระบวนการควบคุมทางพันธุกรรมสำหรับลักษณะการให้ผลผลิต ที่นำไปสู่การระบุยีนหรือตำแหน่งยีนที่มีผลต่อลักษณะสำคัญทางเศรษฐกิจ เช่น การให้ผลผลิตน้ำนม และคุณภาพน้ำนม เช่น ไขมัน โปรตีน เปอร์เซ็นต์โปรตีน เปอร์เซ็นต์ไขมัน และจำนวนเซลล์โซมาติกในน้ำนม เป็นต้น (Kolbehdari *et al.*, 2009)

4. การศึกษาความผันแปรทางพันธุกรรมในระดับจีโนมของโคนม

การศึกษาความผันแปรทางพันธุกรรมในระดับจีโนมของโคนม เป็นข้อมูลพื้นฐานที่สำคัญในการทำเข้าใจในกระบวนการถ่ายทอดข้อมูลทางพันธุกรรมของโคนม ที่จะนำไปสู่การทำแผนที่ความสัมพันธ์ระหว่างเครื่องหมายพันธุกรรมหรือสนิปส์กับยีนที่ควบคุมลักษณะปริมาณ ซึ่งเป็นลักษณะปรากฏที่ถูกควบคุมด้วย QTL โดยการกระจายตัวของตำแหน่งยีน QTL บนโครโมโซมแต่ละคู่ พบว่า มีตำแหน่งยีนจำนวนน้อยที่มีอิทธิพลต่อลักษณะปรากฏมาก และมีตำแหน่งยีนจำนวนมากที่มีอิทธิพลต่อลักษณะปรากฏน้อย (Hayes *et al.*, 1999) ดังนั้น การค้นหา QTL ในตำแหน่งยีนที่มีอิทธิพลปานกลางถึงมากจึงเป็นสิ่งสำคัญ เนื่องจากเป็นข้อมูลที่สามารถช่วยเพิ่มความแม่นยำในการคัดเลือกสัตว์ที่มีพันธุกรรมดีสำหรับลักษณะที่พึงประสงค์ได้

Quantitative trait loci เป็นวิธีการระบุพื้นที่บนโครโมโซมแต่ละคู่ที่มีความสัมพันธ์กับลักษณะที่ปรากฏ ซึ่งมีสมมุติฐาน เพื่อการค้นหาตำแหน่งยีนที่มีอิทธิพลต่อลักษณะการให้ผลผลิตที่ไม่ทราบพื้นที่บนจีโนมโคนม โดยใช้สนิปส์ค้นหาความสัมพันธ์ระหว่างความผันแปรของอัลลีล ณ ตำแหน่งยีนและความผันแปรของลักษณะการให้ผลผลิต ซึ่งสนิปส์เป็นตัวระบุตำแหน่งทางกายภาพบนโครโมโซมที่มีการถ่ายทอดข้อมูลทางพันธุกรรมจากบรรพบุรุษจนถึงปัจจุบัน (Kolbehdari *et al.*, 2009; Fan *et al.*, 2011) อย่างไรก็ตาม สนิปส์ที่เกิดขึ้นบนสายนิวคลีโอไทด์ของดีเอ็นเอในโคนมแต่ละตำแหน่ง เป็นเพียงหนึ่งในสามของเบสที่ทำหน้าที่เป็นรหัสของกรดอะมิโนที่ถูกแปลรหัสไปเป็นโปรตีน ด้วยเหตุนี้ สนิปส์ที่เกิดขึ้นในบริเวณพื้นที่ยีนหรือพื้นที่บางส่วนของยีน จึงไม่สามารถบ่งบอกหน้าที่ของลำดับเบสที่กำหนดรหัสกรดอะมิโนที่เปลี่ยนไป แต่เราสามารถใช้อรรถประโยชน์จากสนิปส์เหล่านั้นมาพิจารณารูปแบบการถ่ายทอดและความแตกต่างทางพันธุกรรมได้ (Hayes, 2011)

ในปัจจุบัน สนิปส์จึงถูกใช้ประโยชน์ในการตรวจสอบการเปลี่ยนแปลงทางพันธุกรรม และความสัมพันธ์ระหว่างสนิปส์กับลักษณะการให้ผลผลิตของโคนม เนื่องจากสนิปส์มีความสัมพันธ์กับสนิปส์ตำแหน่งอื่นบนแต่ละโครโมโซมหรือสัมพันธ์กับตำแหน่งสนิปส์บน QTL ซึ่งความผันแปรของสนิปส์เหล่านี้ มีผลทำให้เกิดความแตกต่างของลักษณะการให้ผลผลิตในโคนม (Bush and Moore, 2012) นอกจากนี้ ผลจากการศึกษาของ Cole *et al.* (2011) พบว่า สนิปส์ที่สัมพันธ์กับลักษณะการให้ผลผลิตของโคนมมีอยู่ประมาณ 1,586 ตำแหน่ง ซึ่งกระจายตัวอยู่บนโครโมโซมร่างกายและโครโมโซมเพศในลักษณะที่ไม่สม่ำเสมอ โดยโครโมโซมบางคู่อาจมีจำนวนสนิปส์มากหรือน้อยในจำนวนที่ไม่เท่ากัน อย่างไรก็ตาม การศึกษาความแตกต่างทางพันธุกรรมสนิปส์ในระดับจีโนมของโคนมในปัจจุบัน ยังพบข้อจำกัดด้านต้นทุนการวิเคราะห์ความแตกต่างพันธุกรรมสนิปส์ที่สูง ดังนั้น การวางแผนการทดลองก่อนศึกษาจึงเป็นสิ่งสำคัญโดยเฉพาะการตรวจสอบแผนที่ QTL ในประชากรโคนมที่ศึกษา

การทำแผนที่ QTL เป็นการใช้ประโยชน์จากความสัมพันธ์ระหว่างสนิปส์และ QTL ซึ่งความสัมพันธ์เหล่านี้เกิดขึ้นเมื่อมีชิ้นส่วนโครโมโซมในประชากรปัจจุบันได้รับการถ่ายทอดจากบรรพบุรุษร่วมกัน โดยชิ้นส่วนโครโมโซมเหล่านี้ถูกนำมาตรวจสอบย้อนกลับไปยังบรรพบุรุษที่มีร่วมกัน ซึ่งเป็นการตรวจสอบโดยใช้อัลลีล หรือ Haplotypes ของเครื่องหมายพันธุกรรม ซึ่งกรณีพบตำแหน่ง QTL บนชิ้นส่วนของโครโมโซมในประชากรโคนมก็จะพบรูปแบบอัลลีลของยีนที่ควบคุมลักษณะเชิงปริมาณเหมือนกัน โดยทั่วไปการทำแผนที่ QTL ในระดับจีโนม เป็นการใช้ประโยชน์จากระดับ LD ระหว่างตำแหน่งสนิปส์บนโครโมโซมแต่ละคู่ (Kolbehdari *et al.*, 2009) ดังนั้น การศึกษาระดับ LD จึงเป็นพื้นฐานสำคัญในการระบุความสัมพันธ์ของตำแหน่งสนิปส์ในระดับจีโนมโคนม

5. การศึกษา linkage disequilibrium ในระดับจีโนม

Linkage disequilibrium เป็นความสัมพันธ์ที่ไม่ได้เกิดจากการสุ่ม (Non-random association) ระหว่างอัลลีลของสนิปส์สองตำแหน่ง ซึ่งมีความสำคัญในการเพิ่มความแม่นยำในการประมาณค่าทางพันธุกรรม (Genomic estimation breeding value; GEBV; Hayes *et al.*, 2009a; Hayes *et al.*, 2009b; Zhou *et al.*, 2013) และช่วยเพิ่มอำนาจในการตรวจสอบความสัมพันธ์ระหว่างสนิปส์และลักษณะการให้ผลผลิต (Walters *et al.*, 2012; Yang *et al.*, 2014) รวมถึงการพิจารณาความผันแปรของยีนที่มีอิทธิพลต่อลักษณะที่สำคัญทางเศรษฐกิจ (Kolbehdari *et al.*, 2009;

Zapata, 2013) ด้วยเหตุนี้ การศึกษาการกระจายตัวและโครงสร้างของ LD จึงเป็นสิ่งสำคัญในการใช้ประโยชน์จากข้อมูลความแตกต่างทางพันธุกรรมสนิปส์ในระดับจีโนมของโคนมในปัจจุบัน

5.1 การประมาณค่า Linkage disequilibrium

การประมาณค่า Linkage disequilibrium (LD) เป็นการประมาณค่าจากความถี่อัลลีลของสนิปส์สองตำแหน่งบนแต่ละโครโมโซมของโคนม โดยใช้สองตัวแปร คือ D' และ r^2 (Ardlie *et al.*, 2002; Khatkar *et al.*, 2008; Bohmanova *et al.*, 2010; Espigolan *et al.*, 2013) โดยค่า D' เป็นค่ามาตรฐานที่ใช้บ่งบอกระดับ LD ที่สัมพันธ์กับค่าสูงสุดของความถี่อัลลีลสนิปส์ ซึ่งคำนวณจากสมการที่ 1

$$D' = \begin{cases} \frac{D}{\min [f(A) \times f(b), f(a) \times f(B)]} & \text{if } D > 0 \\ \frac{D}{\min [f(A) \times f(B), f(a) \times f(b)]} & \text{if } D < 0 \end{cases}, \quad \dots [1]$$

เมื่อ D คำนวณจาก $D = f(AB) \times f(ab) - f(Ab) \times f(aB)$ โดย $f(A)$, $f(a)$, $f(B)$ และ $f(b)$ เป็นความถี่ของอัลลีล A , a , B และ b ตามลำดับ และ $f(AB)$, $f(ab)$, $f(Ab)$ และ $f(aB)$ คือ ความถี่ของ Haplotype AB , ab , Ab และ aB ในประชากรที่ศึกษา (Lewontin, 1964)

อย่างไรก็ตาม การประมาณค่า D' ขึ้นอยู่กับความถี่อัลลีลของสนิปส์และจำนวนประชากรที่ใช้ศึกษา ดังนั้นค่า D' เท่ากับ 1 จึงไม่สามารถนำมาใช้ประโยชน์ในการเปรียบเทียบระดับ LD ระหว่างอัลลีลสนิปส์ได้ และเพื่อลดข้อจำกัดของการประมาณค่า LD Hill and Robertson (1968) จึงได้นำเสนอวิธีมาตรฐานในการประมาณค่า LD โดยใช้ที่เหลื่อมความสัมพันธ์ระหว่างสนิปส์สองตำแหน่งโดยใช้สัญลักษณ์แทนด้วย r^2 ซึ่งเป็นวิธีการประมาณค่าจาก D และความถี่อัลลีลสนิปส์สองตำแหน่งดังสมการที่ 2

$$r^2 = \frac{(D)^2}{f(A) \times f(a) \times f(B) \times f(b)} \quad \dots [2]$$

โดย LD ที่ประมาณได้จากค่า D' และ r^2 จะมีค่าอยู่ในช่วง 0 (Incomplete LD) ถึง 1 (Complete LD) โดยกรณีค่า D' และ r^2 เท่ากับ 1 หมายถึง สนิปส์สองตำแหน่งมีการถ่ายทอดร่วมกัน

แม้ว่าเกิดการกลายพันธุ์ (Mutation) หรือ การเปลี่ยนแปลงยีนโดยการเกิดการแลกเปลี่ยนชิ้นส่วนโครโมโซม (Crossing over) ขึ้นระหว่างการแบ่งเซลล์สืบพันธุ์ขณะเกิดการรวมตัวอีกครั้งระหว่างข้อมูลพันธุกรรม (Recombination) และกรณีที่มีค่าน้อยกว่า 1 หมายถึง สนิปส์สองตำแหน่งมีการหยุดชะงักในถ่ายทอดร่วมกัน ขณะเกิดการรวมตัวอีกครั้งระหว่างข้อมูลพันธุกรรม (Ardlie *et al.*, 2002)

5.2 ปัจจัยที่มีอิทธิพลต่อระดับ Linkage disequilibrium

ระดับ LD ในระดับจีโนมของโคโนมแต่ละประชากรมักผันแปรและแตกต่างกันไปตามปัจจัยต่างๆ เช่น ระยะห่างระหว่างคู่สนิปส์ ค่า Minor allele frequency (MAF) ขนาดประชากร และพันธุ์โค เป็นต้น ซึ่งมีรายละเอียดดังนี้

5.2.1 ระยะห่างระหว่างคู่สนิปส์บนโครโมโซม

ระยะห่างระหว่างคู่สนิปส์บนโครโมโซมแต่ละคู่ มีความสัมพันธ์กับการแลกเปลี่ยนชิ้นส่วนโครโมโซม โดยคู่สนิปส์ที่มีระยะห่างมากจะมีโอกาสเกิดการแลกเปลี่ยนชิ้นส่วนโครโมโซมได้สูงกว่าคู่สนิปส์ที่มีระยะห่างน้อย (Weinstein, 1936) นอกจากนี้ ระยะห่างระหว่างคู่สนิปส์บนแต่ละโครโมโซม ยังมีความสัมพันธ์กับอัตราการรวมตัวอีกครั้งระหว่างข้อมูลพันธุกรรม โดยโครโมโซมที่มีระยะห่างระหว่างคู่สนิปส์มากจะมีอัตราการรวมตัวอีกครั้งระหว่างข้อมูลพันธุกรรมอีกครั้งลดลง (Farre *et al.*, 2013) ซึ่งมีผลทำให้เกิดการสูญเสียระดับ LD ระหว่างคู่สนิปส์บนโครโมโซมแต่ละคู่ ด้วยเหตุนี้ ระดับ LD ที่ถูกใช้ประโยชน์ในการศึกษาความสัมพันธ์ระหว่างสนิปส์และลักษณะการให้ผลผลิต ($r^2 > 0.3$; Ardlie *et al.*, 2002) และการประมาณค่าทางพันธุกรรม ($r^2 > 0.2$; Meuwissen *et al.*, 2001) จึงพบสำหรับคู่สนิปส์ที่มีระยะห่างน้อยบนแต่ละโครโมโซม

สอดคล้องกับ Sargolzaei *et al.* (2008) ที่ศึกษาอิทธิพลของระยะห่างระหว่างคู่สนิปส์ต่อระดับ LD (r^2) ในระดับจีโนมของประชากรโคโนมพันธุ์โฮลส์ไตนในประเทศสหรัฐอเมริกา โดยใช้ข้อมูลสนิปส์บนโครโมโซมร่างกายจำนวน 8,624 ตำแหน่ง จากโคโนมพ่อพันธุ์จำนวน 521 ตัว คู่สนิปส์ทั้งหมด 580,517 คู่ถูกจำแนกตามระยะห่างระหว่างคู่สนิปส์ออกเป็น 10 กลุ่ม โดยแต่ละกลุ่มมีช่วงห่างระหว่างคู่สนิปส์ที่เพิ่มขึ้นทุก 0.1 Mb ผลจากการศึกษา พบว่า

ระยะห่างระหว่างคู่สนิปส์ที่เพิ่มขึ้นจาก 0 ถึง 1 Mb มีผลทำให้ระดับ LD ลดลงจาก 0.58 ถึง 0.08 นอกจากนี้ คู่สนิปส์บนช่วงห่างระหว่าง 0 ถึง 0.1 Mb มีสัดส่วนคู่สนิปส์ที่มีระดับ LD สูง ($r^2 > 0.3$) มากที่สุด (ร้อยละ 68.34)

ในทำนองเดียวกับ Bohmanova *et al.* (2010) ที่ศึกษาระดับการลดลงของ LD (D' และ r^2) เมื่อระยะห่างระหว่างคู่สนิปส์เพิ่มขึ้น โดยใช้ข้อมูลสนิปส์จำนวน 38,590 ตำแหน่ง จาก โคนมพพันธุ์โฮลสไตน์ที่ถูกเลี้ยงในประเทศสหรัฐอเมริกาจำนวน 887 ตัว โดยคู่สนิปส์ทั้งหมดถูกจำแนกเป็น 13 กลุ่มตามระยะห่างระหว่างคู่สนิปส์ที่เพิ่มขึ้นทุก 40 kb, 100 kb, 200 kb, 500 kb, 1 Mb, 3 Mb, 5 Mb, 30 Mb, 50 Mb และ 100 Mb ผลจากการศึกษาพบว่า ระดับ LD ลดลงเมื่อระยะห่างระหว่างคู่สนิปส์เพิ่มขึ้น โดยค่าเฉลี่ย r^2 มีค่าลดลงจาก 0.2 ถึง 0 และค่าเฉลี่ย D' มีค่าลดลงจาก 0.76 ถึง 0.09 เมื่อระยะห่างระหว่างคู่สนิปส์เพิ่มขึ้นจาก 10 kb ถึง 100 Mb อีกทั้งระดับ LD ที่สามารถใช้ประโยชน์ได้ ($D' = 0.72$; $r^2 = 0.20$) อยู่บนช่วงห่างระหว่างคู่สนิปส์ 40 ถึง 60 kb และเมื่อระยะห่างระหว่างคู่สนิปส์มากกว่า 5 Mb พบว่า มีคู่สนิปส์ที่มีระดับ LD ($r^2 > 0.3$) น้อยมาก (ร้อยละ 0) สอดคล้องกับ Espigolan *et al.* (2013) พบว่า ระดับ LD ลดลงเมื่อระยะห่างระหว่างคู่สนิปส์เพิ่มขึ้น โดย r^2 ลดลงจาก 0.34 (1 kb) ถึง 0.11 (100 kb) ซึ่งแตกต่างกับค่า D' ที่มีระดับการลดลงที่ต่ำกว่าค่า r^2

5.2.2 Minor allele frequency (MAF)

Khatkar *et al.* (2008) ศึกษาอิทธิพลของ MAF ต่อระดับ LD ในประชากรโคนมพันธุ์โฮลสไตน์ที่ถูกเลี้ยงในประเทศออสเตรเลีย โดยจำแนกสนิปส์ออกเป็น 3 กลุ่ม ตามค่าต่ำสุดของ MAF คือ 0.05, 0.10 และ 0.20 ตามลำดับ ผลจากการศึกษา พบว่า MAF มีอิทธิพลต่อระดับ LD โดยเฉพาะคู่สนิปส์ที่มีระยะห่างไม่เกิน 40 kb ซึ่งกรณีที่คู่สนิปส์มีระยะห่างระหว่าง 1 ถึง 10 kb สำหรับกลุ่มสนิปส์ที่มีค่า MAF มากกว่าหรือเท่ากับ 0.05 มีค่าเฉลี่ย r^2 เท่ากับ 0.59 ซึ่งแตกต่างกับกลุ่มสนิปส์ที่มีค่า MAF มากกว่าหรือเท่ากับ 0.20 มีค่าเฉลี่ย r^2 เท่ากับ 0.70 อย่างไรก็ตาม เมื่อเปรียบเทียบอิทธิพลของ MAF ต่อการประมาณค่า r^2 และ D' พบว่า การประมาณค่า D' จะได้รับอิทธิพลจากค่า MAF สำหรับคู่สนิปส์ที่มีระยะห่างไม่เกิน 40 kb น้อยกว่าค่า r^2

นอกจากนี้ Bohmarova *et al.* (2010) รายงานว่า ค่าเฉลี่ย D' ที่มีค่าสูงเมื่อค่า MAF ที่ใช้ในการประมาณค่า LD ต่ำ เป็นผลจากสมการประมาณค่า D' มีตัวหารเท่ากับค่าต่ำสุด

ของอัลลีลสปีส์ ด้วยเหตุนี้ กรณีที่คู่สปีส์มีความถี่ต่ำ D ที่อยู่ในสมการประมาณค่า D' ถูกหารด้วยตัวหารขนาดเล็ก ซึ่งมีผลทำให้ D' มีค่าสูงเกินจริง สอดคล้องกับ Espigolan *et al.* (2013) ที่พบว่า D' มีสูงกว่าค่าจริง เมื่อ MAF มีค่าต่ำ ซึ่งเป็นผลจากสมการในการพิจารณา D' ที่มีตัวหารเท่ากับค่าต่ำสุดของอัลลีลสปีส์ ดังนั้นการประมาณค่า LD โดยใช้ D' จึงได้รับอิทธิพลมาจากค่า MAF มากกว่า r^2

5.2.3 ขนาดประชากร

โครงสร้างประชากรที่ใช้ในการศึกษามีผลทำให้ค่า LD ในแต่ละประชากรแตกต่างกัน ซึ่งเป็นผลมาจากความหลากหลายทางพันธุกรรมของโคนมที่ใช้ในการศึกษา ดังนั้น Khatar *et al.* (2008) จึงได้ศึกษาอิทธิพลของขนาดประชากรต่อระดับ LD โดยจำแนกกลุ่มการศึกษาเป็น 10 กลุ่ม ประกอบด้วย กลุ่มที่ 1 ประกอบด้วยโคนมจำนวน 25 ตัว กลุ่มที่ 2 ประกอบด้วยโคนมจำนวน 50 ตัว กลุ่มที่ 3 ประกอบด้วยโคนมจำนวน 75 ตัว กลุ่มที่ 4 ประกอบด้วยโคนมจำนวน 100 ตัว กลุ่มที่ 5 ประกอบด้วยโคนมจำนวน 150 ตัว กลุ่มที่ 6 ประกอบด้วยโคนมจำนวน 200 ตัว กลุ่มที่ 7 ประกอบด้วยโคนมจำนวน 300 ตัว กลุ่มที่ 8 ประกอบด้วยโคนมจำนวน 400 ตัว กลุ่มที่ 9 ประกอบด้วยโคนมจำนวน 500 ตัว และ กลุ่มที่ 10 ประกอบด้วยโคนมจำนวน 1,000 ตัว (ประชากรอ้างอิง) ผลจากการศึกษา พบว่า การประมาณค่า LD จากข้อมูลสปีส์ในประชากรโคนมขนาดเล็ก มีผลทำให้ระดับ LD มีแนวโน้มสูงกว่าประชากรขนาดใหญ่ โดยเฉพาะระยะห่างระหว่างคู่สปีส์น้อยกว่า 40 kb

อย่างไรก็ตาม เมื่อพิจารณาความสัมพันธ์ระหว่างระดับ LD และขนาดประชากรกลุ่มทดสอบและกลุ่มประชากรอ้างอิง พบว่า การประมาณค่า r^2 ควรใช้จำนวนประชากรมากกว่าหรือเท่ากับ 75 ตัว จึงสามารถประมาณค่า r^2 ที่มีค่าเบี่ยงเบนไปจากกลุ่มประชากรอ้างอิงน้อยที่สุด (ค่าสหสัมพันธ์เท่ากับ 0.91) ซึ่งแตกต่างกับการประมาณค่า D' ที่จำเป็นต้องใช้จำนวนประชากรมากกว่าหรือเท่ากับ 400 ตัว (ค่าสหสัมพันธ์เท่ากับ 0.89) สอดคล้องกับการศึกษาอิทธิพลของขนาดประชากรต่อระดับ LD ในประชากรโคนมพันธุ์โฮลสไตน์ในประเทศสหรัฐอเมริกาที่รายงานโดย Bohmanova *et al.* (2010) ซึ่งพบว่า ประชากรขนาดเล็กไม่ส่งผลต่อการประมาณค่า r^2 ซึ่งแตกต่างกับการประมาณค่า D' ที่มีค่าสูงเกินจริงเมื่อประชากรขนาดเล็กและระยะห่างระหว่างคู่สปีส์เพิ่มสูงขึ้น ลักษณะดังกล่าวแสดงให้เห็นว่าการประมาณค่า D' นั้นขึ้นอยู่กับขนาดประชากรมากกว่า r^2 โดยขนาดประชากรน้อยสุดในการประมาณค่า r^2 คือ 55 ตัวและ D' คือ 444 ตัว

5.2.4 พันธุ์โค

Pérez O'Brien *et al.* (2014) ศึกษาในระดับ LD ของโคในตระกูลอินเดีย (*Bos indicus*) และยุโรป (*Bos taurus*) จำนวน 453 ตัว ประกอบด้วยโคพันธุ์แอ่งก้นจำนวน 31 ตัว โฮลสไตน์จำนวน 59 ตัว บราวน์สวิส จำนวน 79 ตัว Fleckvieh จำนวน 96 ตัว บราห์มัน จำนวน 35 ตัว Gir จำนวน 53 ตัว และ Nelore จำนวน 100 ตัว โดยโคเหล่านี้ถูกตรวจสอบความแตกต่างทางพันธุกรรมสนิปส์ด้วย Illuminar BovineHD Genotyping BeadChip assay ประกอบด้วยสนิปส์จำนวน 735,239 ตำแหน่ง ผลจากการศึกษาพบว่า ระดับ LD มีค่าที่แตกต่างกันระหว่างโคตระกูลอินเดียและยุโรป โดยเฉพาะระยะห่างระหว่างคู่สนิปส์ที่มากกว่า 100 kb และระดับ LD ของโคในตระกูลอินเดียมีค่าต่ำกว่าโคตระกูลยุโรป สอดคล้องกับ Lu *et al.* (2012) ที่พบว่า พันธุ์โคมีผลต่อระดับ LD ที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($P < 0.001$) ด้วยเหตุนี้ การศึกษาระดับ LD จึงมีลักษณะจำเพาะกับประชากรที่ใช้ในการศึกษา (Pritchard and Przeworski, 2001)

อุปกรณ์และวิธีการ

1. ลักษณะข้อมูลและประชากรที่ใช้ในการศึกษา

1.1 ประชากรโคนม

ประชากรโคนมที่ใช้ในการศึกษาครั้งนี้ ประกอบด้วย โคนมเพศเมียที่มีข้อมูลพันธุ์ประวัติและข้อมูลการให้ผลผลิตน้ำนมครั้งแรกครบสมบูรณ์ จำนวน 1,413 ตัว โคนมเหล่านี้ถูกเลี้ยงดูในฟาร์มของเกษตรกร ในเขตพื้นที่ความดูแลขององค์การส่งเสริมกิจการโคนมแห่งประเทศไทย (อ.ส.ค.) จำนวน 195 ราย โดยจำแนกเป็น เกษตรกรในพื้นที่ภาคกลาง จำนวน 78 ราย (901 ตัว) ภาคใต้ จำนวน 67 ราย (298 ตัว) และภาคตะวันออกเฉียงเหนือ จำนวน 50 ราย (214 ตัว) ซึ่งเป็นส่วนหนึ่งในประชากรโคนมที่เข้าร่วมโครงการ “การพัฒนาระบบการประเมินความสามารถทางพันธุกรรมโคนมในประเทศไทย” ที่ได้รับทุนสนับสนุนการศึกษาวิจัยจาก สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ มหาวิทยาลัยเกษตรศาสตร์ และองค์การส่งเสริมกิจการโคนมแห่งประเทศไทย

โคนมที่ปรากฏในชุดข้อมูลมีลักษณะเป็นโคนมหลากหลายพันธุ์ ที่เกิดจากการผสมพันธุ์ระหว่างโคตระกูลยุโรป (*Bos taurus*) เช่น โคพันธุ์โฮลสไตน์ เฮอร์ชี บราวน์สวิส และเรดเดน เป็นต้น และโคตระกูลอินเดีย (*Bos indicus*) เช่น โคพันธุ์ชาฮิวาล บราห์มัน และพื้นเมือง เป็นต้น โดยมีพันธุ์โคหลัก คือ พันธุ์โฮลสไตน์ซึ่งมีองค์ประกอบของระดับสายเลือดตั้งแต่ร้อยละ 28 ถึง 100 นอกจากนี้ โคนมบางตัวมีองค์ประกอบทางพันธุกรรมของพันธุ์โคที่แตกต่างกันมากกว่า 6 พันธุ์ ซึ่งโคนมเหล่านี้เป็นลูกของโคพ่อพันธุ์จำนวน 314 ตัว และโคแม่พันธุ์จำนวน 1,284 ตัว

1.2 การจัดเก็บตัวอย่างเลือด

โคนมที่ปรากฏในชุดข้อมูลและยังคงมีชีวิตอยู่ในขณะที่ศึกษา ถูกพิจารณาจัดเก็บตัวอย่างเลือดบนพื้นฐานของความสมบูรณ์ของข้อมูลพันธุ์ประวัติและการให้ผลผลิตน้ำนม ทั้งนี้ โคนมทุกตัวต้องมีข้อมูลพันธุ์ประวัติและการให้ผลผลิตน้ำนมครั้งแรกครบสมบูรณ์ ข้อมูลดังกล่าวถูกรวบรวมตั้งแต่ พ.ศ. 2546 ถึง 2555 โดยข้อมูลพันธุ์ประวัติ ประกอบด้วย

- 1) หมายเลขประจำตัว (ตัวสัตว์ พ่อ และแม่ของสัตว์)
- 2) สัดส่วนพันธุกรรม (ตัวสัตว์ พ่อ และแม่ของสัตว์)
- 3) วันเดือนปีที่สัตว์เกิด

ส่วนข้อมูลการให้ผลผลิตน้ำนม ได้แก่ วันเดือนปีที่สัตว์คลอดลูก และข้อมูลการให้ผลผลิตน้ำนมรายตัวประกอบด้วย 1) ปริมาณการให้ผลผลิตน้ำนม 2) ปริมาณโปรตีนนม และ 3) ปริมาณไขมันนม ทั้งนี้ข้อมูลโคนมที่มีข้อมูลพันธุ์ประวัติไม่ถูกต้องหรือมีการจดบันทึกข้อมูลผิดพลาด และขาดรายละเอียดในส่วนใดส่วนหนึ่งจะไม่ถูกนำมาพิจารณาในการเก็บตัวอย่างเลือด

หลังจากนั้น โคนมที่ผ่านการพิจารณาถูกจัดกลุ่มตามพื้นที่ตั้งฟาร์ม โดยฟาร์มที่อยู่ติดกันหรือมีระยะทางใกล้เคียงกันถูกเก็บตัวอย่างในวันเดียวกัน และฟาร์มที่อยู่ในเขตพื้นที่เดียวกันก็ถูกพิจารณาเก็บตัวอย่างเลือดในช่วงเดียวกัน สำหรับขั้นตอนการเก็บตัวอย่างเลือดใช้วิธีการเจาะเส้นเลือดดำบริเวณโคนหางหรือคอ โดยโคนมแต่ละตัวถูกเก็บตัวอย่างเลือดประมาณ 9 มิลลิลิตรในหลอดเก็บตัวอย่างแบบสูญญากาศที่มีสารป้องกันการแข็งตัวของเลือด (Ethylenediaminetetraacetic acid; EDTA) และตัวอย่างเลือดที่เก็บได้ถูกเก็บรักษาที่อุณหภูมิ 4 องศาเซลเซียสเพื่อรักษาคุณภาพของเลือดจนกระทั่งส่งถึงห้องปฏิบัติการกลางคณะเกษตร มหาวิทยาลัยเกษตรศาสตร์

1.3 การสกัดสารพันธุกรรมดีเอ็นเอ

ตัวอย่างเลือดที่จัดเก็บได้ถูกนำมาสกัดสารพันธุกรรมภายใน 72 ชั่วโมง ภายหลังการเก็บตัวอย่างเลือดโดยใช้ชุดสกัดเชิงการค้า Master Pure™ DNA Purification Kit (Epicentre®, USA) ซึ่งตัวอย่างเลือดจากโคนมตัวอย่างละ 200 ไมโครลิตร ถูกนำมาใช้ในการสกัดสารพันธุกรรมโดยนำตัวอย่างเลือดใส่หลอดทดลองขนาด 1.5 มิลลิลิตร เติมสารละลาย Red cell lysis ปริมาตร 600 ไมโครลิตรเพื่อย่อยเซลล์เม็ดเลือดแดง จากนั้นบ่มที่อุณหภูมิห้องเป็นเวลา 5 นาที และ Vortex เพื่อให้สารละลายรวมตัวกับตัวอย่างเลือด ก่อนบ่มที่อุณหภูมิห้องอีกครั้งเป็นเวลา 5 นาที แล้วนำไปปั่นตกตะกอนด้วยเครื่อง Micro centrifuge ที่ความเร็ว 10,000xg เป็นเวลา 0.25 นาที แยกสารละลายส่วนใสออก เติมสารละลาย Tissue and cells lysis ลงในตะกอนตัวอย่าง เพื่อย่อยองค์ประกอบของเซลล์และเนื้อเยื่อ และเติม RNase A ปริมาตร 1 ไมโครลิตร เพื่อย่อยสารพันธุกรรมอาร์เอ็นเอ จากนั้นนำตัวอย่างไปบ่มที่อุณหภูมิ 37 องศาเซลเซียส เป็นเวลา 30 นาที เมื่อครบกำหนดระยะเวลา นำตัวอย่างทั้งหมดออกมาวางบนน้ำแข็งเป็นเวลา 5 นาที เพื่อหยุดการทำงานของเอนไซม์ เติม

สารละลาย MPC protein precipitation ปริมาตร 175 ไมโครลิตร แล้วนำไปปั่นตกตะกอน จากนั้นเทสารละลายส่วนใสในหลอดทดลองใหม่แล้วเติม Isopropanol ปริมาตร 500 ไมโครลิตร และนำตัวอย่างไปปั่นตกตะกอนด้วยเครื่อง Micro centrifuge ที่อุณหภูมิ 4 องศาเซลเซียสเป็นเวลา 10 นาที เทส่วนใสทิ้งแล้วล้างตะกอนดีเอ็นเอด้วย 70% Ethanol จำนวน 2 ครั้ง จากนั้นระเหย Ethanol ด้วยตู้อบลมร้อนที่อุณหภูมิ 40 องศาเซลเซียส แล้วเติมสารละลาย TE buffer เพื่อรักษาสภาพดีเอ็นเอ

จากนั้นตัวอย่างดีเอ็นเอทั้งหมดถูกตรวจสอบคุณภาพโดยใช้เครื่อง Spectrophotometer รุ่น NanoDrop 2000 (Thermo Fisher Scientific, Wilmington, DE, USA) และดีเอ็นเอที่มีคุณภาพอยู่ในเกณฑ์บริสุทธิ์ สังเกตจากสัดส่วนค่าดูดกลืนแสงที่ 260/280 มีค่าอยู่ในช่วง 1.8 ถึง 2.0 หลังจากนั้นตัวอย่างสารละลายดีเอ็นเอถูกนำมาปรับความเข้มข้นด้วยสารละลาย TE buffer ให้มีความเข้มข้นเท่ากับ 50 ng/μl แล้วย้ายลงใน PCR plate ที่เติม Nuclease-Free water จากนั้นนำ PCR plate ไประเหยน้ำออกให้แห้งด้วยเครื่อง Virtis Bench Top Manifold Freeze Dryer รุ่น Bench Top 2K Freeze Dryer (SP industries, INC.) และปิดฝา PCR plate ให้สนิท ก่อนเก็บรักษาภายใต้การควบคุมอุณหภูมิที่ -80 องศาเซลเซียส

1.4 การจีโนไทป์ดีเอ็นเอด้วย GeneSeek Genomic Profiler low density BeadChip (GGP 9K)

ตัวอย่างดีเอ็นเอของโคนมจำนวน 1,413 ตัว ถูกจีโนไทป์ด้วยชิปเชิงการค้า GGP 9K ของบริษัท GeneSeek โดยหลักการจีโนไทป์ด้วย GGP 9K เป็นกระบวนการที่อาศัยการจับกันอย่างจำเพาะของสายดีเอ็นเอสองสายที่เป็นคู่กัน โดยใช้แผ่นชิปที่ประกอบด้วยเม็ด Beads จำนวน 9,000 เม็ด ที่ทราบตำแหน่งบนโครโมโซมและระยะห่างระหว่างสนิปส์บนโครโมโซมแต่ละคู่ และเมื่อนำตัวอย่างดีเอ็นเอที่ผ่านกระบวนการเพิ่มชิ้นส่วนดีเอ็นเอ ด้วยเอนไซม์ตัดจำเพาะมาทำปฏิกิริยาเพื่อให้จับคู่กัน (Hybridization) ซึ่งพบว่า ดีเอ็นเอที่มีลำดับนิวคลีโอไทด์คู่กันสามารถจับกับชิ้นดีเอ็นเอที่ถูกติดอยู่บนเม็ด Beads ส่วนดีเอ็นเอที่มีลำดับนิวคลีโอไทด์ที่แตกต่างกันไม่สามารถยึดจับกับชิ้นดีเอ็นเอที่ถูกติดบนเม็ด Beads ได้ จากนั้น ดิจิตอลแกบส์เรืองแสงบนเม็ด Beads ก่อนนำไปเข้าเครื่อง Iscan เพื่อตรวจวัดค่าที่เรืองแสง และวัดความเข้มของแถบสี ซึ่งเป็นตัวบ่งบอกชนิดของเบสที่ถูกวัดได้

ข้อมูลสนิปส์ที่ได้จากการจีโนไทป์ด้วย GGP 9K ประกอบด้วยข้อมูลสนิปส์ แผนที่สนิปส์ และสนิปส์จีโนไทป์ของโคนมรายตัว ซึ่งมีรายละเอียดดังนี้

ก. ข้อมูลสnpที่ได้แก่

- 1) ชื่อสnp (SNPs name)
- 2) จำนวนคอนมที่ไม่พบตำแหน่งสnpในแต่ละตำแหน่ง (No calls)
- 3) สัดส่วนการปรากฏของสnpในแต่ละตำแหน่งในประชากรคอนม (Call frequency)
- 4) ความถี่ของสnpจีโนไทป์แบบ AA ในแต่ละตำแหน่ง (AA frequency)
- 5) ความถี่ของสnpจีโนไทป์แบบ AB ในแต่ละตำแหน่ง (AB frequency)
- 6) ความถี่ของสnpจีโนไทป์แบบ BB ในแต่ละตำแหน่ง (BB frequency)
- 7) ค่า Minor allele frequency (MAF) ของสnpในแต่ละตำแหน่ง

ข. ข้อมูลแผนที่สnpที่ได้แก่

- 1) ชื่อสnp (SNPs name)
- 2) โครโมโซม (Chromosome)
- 3) ตำแหน่งทางกายภาพบนโครโมโซมคอนม (Position of SNPs)
- 4) รูปแบบอัลลีลสnp (SNPs) คือ ลำดับเบสของสnpแต่ละตำแหน่งที่สามารถเกิดขึ้นได้ ประกอบด้วย 8 รูปแบบ คือ A/C, A/G, A/T, C/G, G/C, T/A, T/C และ T/G

ค. ข้อมูลสnpจีโนไทป์ของคอนมรายตัว ได้แก่

- 1) หมายเลขประจำตัวสัตว์
- 2) รูปแบบสnpจีโนไทป์ในแต่ละตำแหน่งของคอนมรายตัว

ข้อมูลสnpทั้งหมด 8,810 ตำแหน่งของคอนมจำนวน 1,413 ตัว สามารถพบตำแหน่งสnpของคอนมรายตัวมากที่สุดจำนวน 8,790 ตำแหน่ง (ตารางที่ 1) และสnpในแต่ละตำแหน่งมีค่าเฉลี่ยของรูปแบบสnpจีโนไทป์ AB มากที่สุด (0.443; ตารางที่ 2)

ตารางที่ 1 จำนวนสนิปส์ที่พบในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

ภูมิภาค	จำนวนโคนม (ตัว)	จำนวนสนิปส์ (ตำแหน่ง)		
		ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย $\pm$ SD
กลาง	901	5,342	8,790	8,697.410 $\pm$ 358.720
ใต้	298	4,865	8,789	8,725.570 $\pm$ 262.492
ตะวันออกเฉียงเหนือ	214	6,462	8,787	8,669.700 $\pm$ 467.909
ทั้งหมด	1,413	4,865	8,790	8,695.830 $\pm$ 372.485

หมายเหตุ: SD คือ ส่วนเบี่ยงเบนมาตรฐาน

ตารางที่ 2 รูปแบบสนิปส์จีโนไทป์ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

ภูมิภาค	รูปแบบสนิปส์จีโนไทป์ ¹			ค่าเฉลี่ย
	AA $\pm$ SD	AB $\pm$ SD	BB $\pm$ SD	MAF $\pm$ SD
กลาง	0.262 $\pm$ 0.022	0.443 $\pm$ 0.022	0.295 $\pm$ 0.019	0.483 $\pm$ 0.007
ใต้	0.258 $\pm$ 0.023	0.444 $\pm$ 0.040	0.298 $\pm$ 0.020	0.480 $\pm$ 0.007
ตะวันออกเฉียงเหนือ	0.260 $\pm$ 0.017	0.443 $\pm$ 0.032	0.297 $\pm$ 0.017	0.482 $\pm$ 0.006
ทั้งหมด	0.260 $\pm$ 0.021	0.443 $\pm$ 0.038	0.296 $\pm$ 0.019	0.482 $\pm$ 0.007

หมายเหตุ: ¹คือ ค่าเฉลี่ยของรูปแบบสนิปส์จีโนไทป์ในประชากรโคนมที่ศึกษา

SD คือ ส่วนเบี่ยงเบนมาตรฐาน

1.5 การตรวจคุณภาพสนิปส์

ข้อมูลสนิปส์จากโคนมจำนวน 8,810 ตำแหน่ง ประกอบด้วย 1) ข้อมูลสนิปส์ ได้แก่ ชื่อสนิปส์ จำนวนโคนมที่ไม่พบตำแหน่งสนิปส์ในแต่ละตำแหน่ง และสัดส่วนการปรากฏของสนิปส์ในแต่ละตำแหน่งในประชากรโคนม และ 2) ข้อมูลแผนที่สนิปส์ ได้แก่ ชื่อสนิปส์ โครโมโซม และตำแหน่งทางกายภาพบนโครโมโซมของโคนม ถูกรวมเข้าด้วยกันโดยใช้โปรแกรมสำเร็จรูป Microsoft Excel (Microsoft Office, 2010) เพื่อใช้ในการตรวจคุณภาพสนิปส์ โดยพิจารณาจาก

สัดส่วนการปรากฏของสปีส์ในแต่ละตำแหน่งในประชากรโคนม (ต้องมากกว่าร้อยละ 90) ซึ่งเป็นตัวบ่งชี้ประสิทธิภาพการจีโนไทป์ด้วยชิปเชิงการค้า เนื่องจากสปีส์ทุกตำแหน่งบนแผ่นชิปสามารถพบในประชากรโคนมที่ศึกษาได้ทั้งหมด ดังนั้น จำนวนโคนมที่ไม่พบตำแหน่งสปีส์ในแต่ละตำแหน่งจึงเป็นผลมาจากกระบวนการจีโนไทป์ จากนั้นสปีส์ที่ผ่านเกณฑ์กำหนดและทราบตำแหน่งทางกายภาพบนโครโมโซมแต่ละคู่ถูกนำมาคำนวณค่า Minor allele frequency เพื่อพิจารณารูปแบบอัลลีลของสปีส์แต่ละตำแหน่งดังสมการที่ 3 และ 4

$$\text{ความถี่ของอัลลีล } A = \frac{1}{2} AB + AA \quad \dots\dots[3]$$

$$\text{ความถี่ของอัลลีล } B = \frac{1}{2} AB + BB \quad \dots\dots[4]$$

เมื่อ AA , AB และ BB คือ ความถี่ของรูปแบบสปีส์จีโนไทป์ในประชากรโคนมที่ศึกษา และกรณีที่ความถี่อัลลีล A มากกว่า B แสดงว่า B คือ Minor allele frequency ซึ่งมีค่าอยู่ในช่วง 0 ถึง 0.5

2. วิธีการศึกษา

การวิจัยในครั้งนี้ เป็นการศึกษาความแตกต่างทางพันธุกรรมสปีส์ในระดับจีโนมของโคนมในประเทศไทย ซึ่งเป็นแหล่งพันธุกรรมในประชากรโคนมหลากหลายพันธุ์ (Multibreed) ที่ประกอบด้วยโคนมพันธุ์แท้และลูกผสม ซึ่งถูกจำแนกเป็น 2 ส่วนการศึกษาตามวัตถุประสงค์ของงานวิจัย ได้แก่

2.1 ศึกษารูปแบบความผันแปรและประมาณค่า Linkage disequilibrium ของสปีส์ในระดับจีโนมของประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

2.1.1 ศึกษารูปแบบความผันแปรของสปีส์ในระดับจีโนมของโคนม

ข้อมูลสปีส์ที่ผ่านการตรวจสอบคุณภาพ ถูกจำแนกตามตำแหน่งทางกายภาพบนโครโมโซมแต่ละคู่ เพื่อพิจารณาการกระจายตัวของสปีส์ ตำแหน่งสปีส์บนโครโมโซม ความหนาแน่นสปีส์ (พิจารณาจากจำนวนสปีส์ที่พบบนแต่ละโครโมโซมต่อระยะห่างระหว่าง

ตำแหน่งสปีส์เริ่มต้นและสุดท้ายบนแต่ละโครโมโซม) และค่า Minor allele frequency (MAF) ด้วยสถิติเชิงพรรณนา (Descriptive statistics) ได้แก่ ค่าความถี่ (Frequency) ค่าต่ำสุด (Min) ค่าสูงสุด (Max) ค่ามัธยฐาน (Median) ค่าเฉลี่ย (Mean) และส่วนเบี่ยงเบนมาตรฐาน (Standard deviation) โดยใช้โปรแกรมสำเร็จรูป SAS (SAS, 2003) เพื่อพิจารณารูปแบบความแตกต่างของสปีส์บนโครโมโซมแต่ละคู่ และสปีส์ในระดับจีโนมของโคนมทั้งหมด จากนั้นสร้างกราฟแสดงแนวโน้มการเปลี่ยนแปลงสัดส่วนสปีส์ในระดับจีโนมด้วยโปรแกรมสำเร็จรูป Microsoft Excel (Microsoft Office, 2010) เพื่อใช้ประกอบการพิจารณาถึงแนวโน้มการเปลี่ยนแปลงโดยภาพรวม

2.1.2 การประมาณค่า Linkage disequilibrium

ข้อมูลพันธุประวัติของโคนมรายตัว ประกอบด้วย หมายเลขประจำตัวสัตว์ พ่อและแม่ของสัตว์ ถูกนำมาสร้างความสัมพันธ์ระหว่างโคนมที่ปรากฏในชุดข้อมูล โดยโคนมที่มีพ่อแม่ตัวเดียวกันถูกกำหนดตัวเลขเหมือนกัน ส่วนโคนมที่มีพ่อหรือแม่ตัวเดียวกันถูกกำหนดตัวเลขที่แตกต่างกัน เช่นเดียวกับโคนมที่เกิดจากพ่อและแม่ที่แตกต่างกัน จากนั้น โคนมแต่ละตัวถูกจำแนกตามเพศเป็น 2 กลุ่ม และกำหนดตัวเลขให้กับโคนมเพศผู้เท่ากับ 1 และ เพศเมียเท่ากับ 2 จากนั้นข้อมูลดังกล่าวถูกรวมเข้ากับข้อมูลรูปแบบสปีส์จีโนไทป์ในแต่ละตำแหน่งของโคนมรายตัว โดยรูปแบบสปีส์จีโนไทป์ในประชากรที่ศึกษา ประกอบด้วย AA AB และ BB ด้วยรูปแบบอัลลีลสปีส์ที่แตกต่างกันจำนวน 8 ชุด คือ A/C, A/G, A/T, C/G, G/C, T/A, T/C และ T/G ซึ่งรูปแบบอัลลีลสปีส์เหล่านี้ ถูกนำมากำหนดรหัสใหม่ให้อยู่ในรูป A = 1, C = 2, G = 3 และ T = 4 โดยข้อมูลรูปแบบสปีส์จีโนไทป์ที่ใช้ในการประมาณค่า LD ครั้งนี้จะอยู่ในรูปดังนี้

ก) รูปแบบอัลลีลสปีส์ A/C กรณีรูปแบบสปีส์จีโนไทป์ AA = 1 1, AB = 1 2 และ BB = 2 2

ข) รูปแบบอัลลีลสปีส์ A/G กรณีรูปแบบสปีส์จีโนไทป์ AA = 1 1, AB = 1 3 และ BB = 3 3

ค) รูปแบบอัลลีลสปีส์ A/T กรณีรูปแบบสปีส์จีโนไทป์ AA = 1 1, AB = 1 4 และ BB = 4 4

ง) รูปแบบอัลลีลสปีส์ C/G กรณีรูปแบบสปีส์จีโนไทป์ AA = 2 2, AB = 2 3 และ BB = 3 3

จ) รูปแบบอัลลีลสปีส์ G/C กรณีรูปแบบสปีส์จีโนไทป์ AA = 3 3, AB = 3 2 และ BB = 2 2

ก) รูปแบบอัลลีลสปีส์ T/A กรณีรูปแบบสปีส์จีโนไทป์ AA = 4 4, AB = 4 1 และ BB = 1 1

ข) รูปแบบอัลลีลสปีส์ T/C กรณีรูปแบบสปีส์จีโนไทป์ AA = 4 4, AB = 4 2 และ BB = 2 2

ค) รูปแบบอัลลีลสปีส์ T/G กรณีรูปแบบสปีส์จีโนไทป์ AA = 4 4, AB = 4 3 และ BB = 3 3

กรณีที่รูปแบบสปีส์จีโนไทป์ UU = 0 0 ในรูปแบบอัลลีลสปีส์ทุกรูปแบบ เนื่องจากเป็นตำแหน่งสปีส์ที่ไม่ถูกจีโนไทป์ (Missing genotype)

จากนั้นสปีส์ที่ทราบตำแหน่งทางกายภาพบนโครโมโซมร่างกายแต่ละคู่ ถูกเชื่อมโยงเข้ากับข้อมูลแผนที่สปีส์ด้วย Haploview ซอฟต์แวร์ (Barrett *et al.*, 2005) เพื่อประมาณค่าพารามิเตอร์ D' และ r^2 ซึ่งเป็นการประมาณค่าจากตำแหน่งสปีส์ที่มีค่า $MAF \geq 0.05$ มีค่า P-value ของ Hardy-Weinberg equilibrium (HWE) ≥ 0.0001 และมีระยะห่างระหว่างคู่สปีส์ไม่เกิน 5 Mb

โดยค่า D' และ r^2 ที่ได้จากการประมาณ ถูกพิจารณาด้วยสถิติเชิงพรรณนา ได้แก่ ค่าความถี่ ค่ามัธยฐาน ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน เป็นต้น โดยใช้โปรแกรมสำเร็จรูป SAS (SAS, 2003) เพื่อพิจารณาการกระจายตัวของค่า LD บนโครโมโซมคอนมแต่ละคู่ จากนั้นสร้างกราฟแสดงแนวโน้มการเปลี่ยนแปลงค่า D' และ r^2 ในระดับจีโนมด้วยโปรแกรมสำเร็จรูป Microsoft Excel (Microsoft Office, 2010) เพื่อใช้ประกอบการพิจารณาถึงแนวโน้มการเปลี่ยนแปลงโดยภาพรวม

2.2 ศึกษาปัจจัยที่มีผลต่อระดับ Linkage disequilibrium ในระดับจีโนมของประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

2.2.1 ศึกษาอิทธิพลของระยะห่างระหว่างคู่สนิปส์ต่อระดับ Linkage disequilibrium

ค่า D' และ r^2 ที่ได้จากการประมาณการศึกษาที่ 1.2 ถูกจำแนกตามช่วงห่างระหว่างคู่สนิปส์ออกเป็น 3 กลุ่ม ได้แก่ กลุ่มสนิปส์ที่มีระยะห่างระหว่างคู่สนิปส์ที่เพิ่มขึ้นทุก 10 kb, 100 kb และ 1 Mb ซึ่งมีรายละเอียดในแต่ละกลุ่มดังนี้

ก) กลุ่มที่ 1 คือ คู่สนิปส์ที่มีระยะห่างระหว่างคู่สนิปส์ที่เพิ่มขึ้นทุก 10 kb ประกอบด้วยคู่สนิปส์ 10 กลุ่ม จากระยะห่างระหว่างคู่สนิปส์ 0 ถึง 100 kb

ข) กลุ่มที่ 2 คือ คู่สนิปส์ที่มีระยะห่างระหว่างคู่สนิปส์ที่เพิ่มขึ้นทุก 100 kb ประกอบด้วยคู่สนิปส์ 9 กลุ่ม จากระยะห่างระหว่างคู่สนิปส์ 100 ถึง 1,000 kb

ค) กลุ่มที่ 3 คือ คู่สนิปส์ที่มีระยะห่างระหว่างคู่สนิปส์ที่เพิ่มขึ้นทุก 1 Mb ประกอบด้วยคู่สนิปส์ 4 กลุ่ม จากระยะห่างระหว่างคู่สนิปส์ 1 ถึง 5 Mb

สนิปส์ในแต่ละกลุ่ม ถูกนำมาพิจารณาด้วยสถิติเชิงพรรณนา และเปรียบเทียบค่าเฉลี่ยสี่เหลี่ยม (Least square mean) ของค่า D' และ r^2 บนระยะห่างระหว่างคู่สนิปส์ที่เพิ่มขึ้น ในแต่ละกลุ่ม ด้วยวิธี Bonferroni ที่ระดับนัยสำคัญทางสถิติ 0.05 โดยใช้ชุดคำสั่งในโปรแกรมสำเร็จรูป SAS (SAS, 2003) ได้แก่ PROC UNIVARIATE, PROC FREQ และ PROC GLM

2.2.2 ศึกษาอิทธิพลของค่า MAF ต่อระดับ Linkage disequilibrium

ข้อมูลสนิปส์จีโนไทป์ของโคนมรายตัว ในประชากรโคนมที่ศึกษาจำนวน 8,207 ตำแหน่ง ถูกจำแนกเป็น 4 กลุ่มตามค่าต่ำสุดของ MAF คือ 0.05, 0.10, 0.15 และ 0.20 ตามลำดับ เพื่อพิจารณาอิทธิพลระดับค่า MAF ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย โดยสนิปส์จีโนไทป์ของโคนมรายตัวในแต่ละกลุ่ม ถูกนำมาเชื่อมโยงเข้ากับข้อมูลแผนที่สนิปส์ด้วย Haploview ซอฟต์แวร์ (Barrett *et al.*, 2005) เพื่อประมาณค่าพารามิเตอร์ D' และ r^2 โดยสนิปส์ที่ใช้ในการศึกษาในแต่ละกลุ่มที่มีค่า P-value ของ Hardy-Weinberg equilibrium (HWE) ≥ 0.0001 และมีระยะห่างระหว่างคู่สนิปส์บนโครโมโซมร่างกายแต่ละคู่ไม่เกิน 5 Mb (ตารางที่ 3)

ตารางที่ 3 จำนวนสนิปส์ และคู่สนิปส์ที่ผ่านเกณฑ์พิจารณาในแต่ละกลุ่มของค่า MAF

เกณฑ์พิจารณาดำเน่งสนิปส์	กลุ่มสนิปส์			
	MAF ≥ 0.05	MAF ≥ 0.10	MAF ≥ 0.15	MAF ≥ 0.20
จำนวนสนิปส์ (ตำแหน่ง)	8,207	8,093	7,918	7,613
จำนวนสนิปส์ที่มี P-value ของ HWE มากกว่า 0.0001 (ตำแหน่ง)	8,161	7,863	7,837	7,522
จำนวนคู่สนิปส์ที่มีระยะไม่เกิน 5 Mb	129,558	125,150	118,575	108,458

จากนั้นค่า D' และ r^2 ที่ได้จากการประมาณถูกพิจารณาด้วยสถิติเชิงพรรณนา จำแนกตามระยะห่างระหว่างคู่สนิปส์ตามการศึกษาที่ 2.1 โดยใช้โปรแกรมสำเร็จรูป SAS (SAS, 2003) เพื่อพิจารณารูปแบบความแตกต่างของ D' และ r^2 บนโครโมโซมแต่ละคู่ และในระดับจีโนมของโคนมทั้งหมด ผลที่ได้จากการวิเคราะห์ทางสถิติถูกสร้างกราฟแสดงแนวโน้มการเปลี่ยนแปลงระดับ LD ในระดับจีโนมด้วยโปรแกรมสำเร็จรูป Microsoft Excel (Microsoft Office, 2010) เพื่อใช้ประกอบการพิจารณาถึงแนวโน้มการเปลี่ยนแปลงโดยภาพรวม

2.2.3 ศึกษาอิทธิพลของขนาดประชากรที่ใช้ศึกษาต่อระดับ Linkage disequilibrium

ข้อมูลสนิปส์จีโนมไทป์ทั้งหมด 8,207 ตำแหน่งของโคนมรายตัว จำนวน 1,413 ตัว ที่ปรากฏในชุดข้อมูลทั้งหมด ถูกสุ่มเป็น 6 กลุ่ม โดยใช้หลักการสุ่มแบบ Bootstrap ของ Teare *et al.* (2002) ซึ่งเป็นการสุ่มจำนวนโคนมจากข้อมูลโคนมทั้งหมดที่ปรากฏในชุดข้อมูล สำหรับประมาณค่า LD โดย กลุ่มที่ 1 มีจำนวนโคนม 1,059 ตัว (3/4; ร้อยละ 75) กลุ่มที่ 2 มีจำนวนโคนม 707 ตัว (1/2; ร้อยละ 50) กลุ่มที่ 3 มีจำนวนโคนม 354 ตัว (1/4; ร้อยละ 25) กลุ่มที่ 4 มีจำนวนโคนม 177 ตัว (1/8; ร้อยละ 12.5) กลุ่มที่ 5 มีจำนวนโคนม 89 ตัว (1/16; ร้อยละ 6.25) และกลุ่มที่ 6 มีจำนวนโคนม 45 ตัว (1/32; ร้อยละ 3.125) และการประมาณค่า LD สำหรับขนาดประชากรที่แตกต่างกันในแต่ละกลุ่มใช้หลักเกณฑ์การคัดเลือกสนิปส์และวิธีการประมาณค่า LD เช่นเดียวกับการศึกษาที่ 1.2

โดยค่า D' และ r^2 ของคู่สนิปส์บนโครโมโซมแต่ละคู่ของโคนมในแต่ละกลุ่มที่มีระยะห่างระหว่างคู่สนิปส์น้อยกว่า 5 Mb ถูกนำมาเชื่อมโยงกับค่า D' และ r^2 ที่ได้จากการศึกษาที่ 1.2 ด้วยโปรแกรมสำเร็จรูป Microsoft Excel (Microsoft Office, 2010) และคัดเลือกคู่สนิปส์ที่

สามารถประมาณค่า D' และ r^2 ได้ในทุกกลุ่มของขนาดประชากร มาพิจารณาด้วยสถิติเชิงพรรณนา ได้แก่ ค่าความถี่ ค่าต่ำสุด ค่าสูงสุด ค่ามัธยฐาน ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน โดยใช้โปรแกรมสำเร็จรูป SAS (SAS, 2003) และใช้ชุดคำสั่ง PROC CORR เพื่อพิจารณาความสัมพันธ์ของระดับ LD ที่ได้จากการประมาณในแต่ละขนาดประชากรที่ศึกษากับข้อมูลประชากรโคนมทั้งหมดที่ปรากฏในชุดข้อมูลจากนั้นค่า D' และ r^2 ถูกสร้างกราฟแสดงความสัมพันธ์ระหว่างค่า D' และ r^2 ที่ประมาณได้ ด้วยโปรแกรมสำเร็จรูป Microsoft Excel (Microsoft Office, 2010) เพื่อใช้ประกอบการพิจารณาขนาดประชากรโคนมต่ำสุด สำหรับเป็นตัวแทนข้อมูลโคนมทั้งหมดที่ปรากฏในชุดข้อมูลที่ศึกษา



ผลและวิจารณ์

1. ศึกษารูปแบบความผันแปรและประมาณค่า Linkage disequilibrium ของสปีส์ในระดับจีโนมของประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

การพิจารณาข้อมูลทางพันธุกรรมในประชากรโคนมไทยหลากหลายพันธุ์ เป็นการพิจารณาบนพื้นฐานข้อมูลพันธุกรรมที่นำไปสู่ความเข้าใจของลักษณะปรากฏที่แตกต่างกัน โดยใช้สปีส์ 8,810 ตำแหน่งที่มีความจำเพาะบนโครโมโซม และรูปแบบเบสที่หลากหลายบนสายนิวคลีโอไทด์ของดีเอ็นเอ ดังนั้น ในการศึกษาครั้งนี้จึงเป็นการพิจารณาความแตกต่างของสปีส์ โดยพิจารณาจากสัดส่วนและเปอร์เซ็นต์ของสปีส์ในระดับจีโนมของโคนมจำนวน 1,413 ตัว ประกอบด้วย สปีส์ที่สามารถใช้ประโยชน์ได้ในการศึกษาครั้งนี้จำนวน 8,573 ตำแหน่ง คิดเป็นร้อยละ 97 ของสปีส์ทั้งหมด โดยสปีส์เหล่านี้ปรากฏในประชากรโคนมที่ศึกษามากที่สุดจำนวน 1,413 ตัว และน้อยที่สุดจำนวน 1,275 ตัว นอกจากนี้ประชากรโคนมที่ศึกษาครั้งนี้ มีค่าเฉลี่ยสปีส์ที่ปรากฏอยู่ที่ $1,398 \pm 11$ ตัว

1.1 การกระจายตัวของสปีส์บนโครโมโซม

สปีส์ในระดับจีโนมของโคนมจำนวน 8,573 ตำแหน่ง ครอบคลุมพื้นที่จีโนมของประชากรโคนมที่ศึกษาทั้งหมด 2,656.205 Mb ด้วยตำแหน่งสปีส์บนโครโมโซมร่างกายเริ่มต้น 0.037 (โครโมโซมคู่ที่ 27) ถึง 1.176 Mb (โครโมโซมคู่ที่ 19) และตำแหน่งสปีส์บนโครโมโซมร่างกายสุดท้าย 42.851 Mb (โครโมโซมคู่ที่ 25) ถึง 158.162 Mb (โครโมโซมคู่ที่ 1) ส่วนโครโมโซมเพศมีตำแหน่งสปีส์เริ่มต้นคือ 0.0485 Mb และมีตำแหน่งสปีส์สุดท้ายเท่ากับ 148.482 Mb นอกจากนี้ โครโมโซมร่างกายคู่ที่ 1 มีค่าเฉลี่ยตำแหน่งสปีส์บนโครโมโซมมากที่สุด (85.595 Mb) และ โครโมโซมคู่ที่ 25 มีค่าเฉลี่ยตำแหน่งสปีส์บนโครโมโซมน้อยที่สุด (21.22 Mb) และโครโมโซมเพศมีค่าเฉลี่ยตำแหน่งสปีส์บนโครโมโซมเท่ากับ 72.132 Mb (ตารางที่ 3)

ตารางที่ 4 จำนวนสปีส์ ความหนาแน่นสปีส์ ค่าเฉลี่ยตำแหน่งสปีส์บนโครโมโซม ส่วนเบี่ยงเบนมาตรฐาน และค่า Minor allele frequency (MAF) บนโครโมโซมร่างกายและโครโมโซมเพศ ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

โครโมโซม	ตำแหน่งสปีส์บนโครโมโซม (Mb)			จำนวนสปีส์ (ตำแหน่ง)	ความหนาแน่น สปีส์ ¹	ค่า Minor allele frequency (MAF)		
	เริ่มต้น	สุดท้าย	ค่าเฉลี่ย $\pm$ SD			ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย $\pm$ SD
1	0.135	158.162	85.595 $\pm$ 48.660	532	3.367	0.000	0.500	0.381 $\pm$ 0.076
2	0.190	136.484	70.479 $\pm$ 42.828	438	3.214	0.025	0.500	0.376 $\pm$ 0.084
3	0.306	121.750	56.837 $\pm$ 35.310	410	3.376	0.075	0.499	0.376 $\pm$ 0.077
4	0.254	120.615	60.704 $\pm$ 35.442	361	2.999	0.027	0.500	0.382 $\pm$ 0.079
5	0.097	120.784	60.077 $\pm$ 35.712	386	3.198	0.026	0.497	0.372 $\pm$ 0.084
6	0.203	121.357	60.111 $\pm$ 35.697	441	3.640	0.107	0.500	0.376 $\pm$ 0.077
7	0.245	112.610	56.433 $\pm$ 34.405	346	3.079	0.107	0.499	0.376 $\pm$ 0.075
8	0.459	113.321	57.829 $\pm$ 34.218	350	3.101	0.029	0.500	0.384 $\pm$ 0.077
9	0.821	105.463	49.420 $\pm$ 32.558	330	3.154	0.072	0.499	0.375 $\pm$ 0.079
10	1.127	104.215	52.086 $\pm$ 31.690	322	3.124	0.001	0.500	0.380 $\pm$ 0.079
11	0.109	107.043	53.376 $\pm$ 32.531	343	3.208	0.001	0.499	0.374 $\pm$ 0.078
12	0.622	91.092	44.323 $\pm$ 28.110	270	2.984	0.038	0.499	0.370 $\pm$ 0.083
13	0.313	84.149	42.955 $\pm$ 25.565	272	3.244	0.063	0.498	0.376 $\pm$ 0.083
14	1.464	84.616	41.863 $\pm$ 25.807	280	3.367	0.001	0.499	0.377 $\pm$ 0.080
15	1.243	85.012	43.502 $\pm$ 24.824	275	3.283	0.020	0.500	0.380 $\pm$ 0.076
16	0.556	80.925	38.641 $\pm$ 23.948	301	3.745	0.000	0.499	0.368 $\pm$ 0.098
17	0.111	74.966	38.960 $\pm$ 23.131	223	2.979	0.043	0.497	0.372 $\pm$ 0.082

ตารางที่ 4 (ต่อ)

โครโมโซม	ตำแหน่งสปีส์บนโครโมโซม (Mb)			จำนวนสปีส์ (ตำแหน่ง)	ความหนาแน่น สปีส์ ¹	ค่า Minor allele frequency (MAF)		
	เริ่มต้น	สุดท้าย	ค่าเฉลี่ย $\pm$ SD			ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย $\pm$ SD
18	0.084	65.979	37.237 $\pm$ 21.875	239	3.627	0.012	0.500	0.374 $\pm$ 0.083
19	1.176	64.007	35.413 $\pm$ 19.294	220	3.501	0.031	0.498	0.380 $\pm$ 0.078
20	0.247	71.794	32.030 $\pm$ 21.127	284	3.969	0.131	0.500	0.373 $\pm$ 0.085
21	0.803	70.608	39.815 $\pm$ 20.991	250	3.581	0.000	0.499	0.380 $\pm$ 0.086
22	0.499	60.931	30.738 $\pm$ 19.105	199	3.293	0.103	0.500	0.379 $\pm$ 0.075
23	0.403	52.129	29.225 $\pm$ 15.403	206	3.983	0.136	0.499	0.381 $\pm$ 0.070
24	0.542	62.644	30.840 $\pm$ 18.829	213	3.430	0.013	0.499	0.380 $\pm$ 0.083
25	0.138	42.851	21.228 $\pm$ 13.889	159	3.723	0.121	0.497	0.377 $\pm$ 0.080
26	0.727	51.680	27.715 $\pm$ 16.268	175	3.435	0.000	0.500	0.374 $\pm$ 0.095
27	0.037	45.369	22.741 $\pm$ 14.819	155	3.419	0.070	0.498	0.372 $\pm$ 0.085
28	0.041	46.195	23.691 $\pm$ 14.895	149	3.228	0.004	0.499	0.372 $\pm$ 0.089
29	0.630	50.972	25.602 $\pm$ 15.987	159	3.158	0.147	0.498	0.378 $\pm$ 0.073
X	0.485	148.482	72.132 $\pm$ 47.343	285	1.926	0.105	0.500	0.357 $\pm$ 0.095

หมายเหตุ: ¹ คือ จำนวนตำแหน่งสปีส์ต่อความยาวโครโมโซม 1 Mb; SD คือ ส่วนเบี่ยงเบนมาตรฐาน

การกระจายตัวของสปีส์บนโครโมโซมแต่ละคู่ในประชากรโคนม (ตารางที่ 3) พบว่า โครโมโซมร่างกายคู่ที่ 28 มีจำนวนสปีส์น้อยที่สุด (149 ตำแหน่ง) และโครโมโซมคู่ที่ 1 มีจำนวนสปีส์มากที่สุด (532 ตำแหน่ง) และโครโมโซมเพศมีจำนวนสปีส์ทั้งหมด 285 ตำแหน่ง เมื่อพิจารณาความหนาแน่นของสปีส์บนโครโมโซมแต่ละคู่ในประชากร พบว่า มีความหนาแน่นของสปีส์บนโครโมโซมร่างกายมากกว่าโครโมโซมเพศ โดยโครโมโซมร่างกายคู่ที่ 17 มีความหนาแน่นของสปีส์น้อยที่สุด (2.979 ตำแหน่งต่อความยาวโครโมโซมที่เพิ่มขึ้น 1 Mb) และโครโมโซมคู่ที่ 20 มีความหนาแน่นของสปีส์มากที่สุด (3.969 ตำแหน่งต่อความยาวโครโมโซมที่เพิ่มขึ้น 1 Mb) ส่วนโครโมโซมเพศมีความหนาแน่นของสปีส์เท่ากับ 1.926 ตำแหน่งต่อความยาวโครโมโซมที่เพิ่มขึ้น 1 Mb อย่างไรก็ตาม การกระจายตัวของสปีส์บนโครโมโซมแต่ละคู่ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทยมีการกระจายตัวไม่สม่ำเสมอ โดยมีการกระจายตัวเป็นกลุ่มบนพื้นที่บางส่วนของโครโมโซม ลักษณะความผันแปรทางพันธุกรรมเช่นนี้ สอดคล้องกับ ทวี และคณะ (2557) สำหรับโคนมไทยหลากหลายพันธุ์ และที่ปรากฏในประชากรโคนมโฮลสไตน์ในประเทศแคนาดา (Sargolzaei *et al.*, 2008) ประเทศเยอรมนี (Qanbari *et al.*, 2010) และประเทศสหรัฐอเมริกา (Cole *et al.*, 2011)

1.2 การกระจายตัว Minor allele frequency

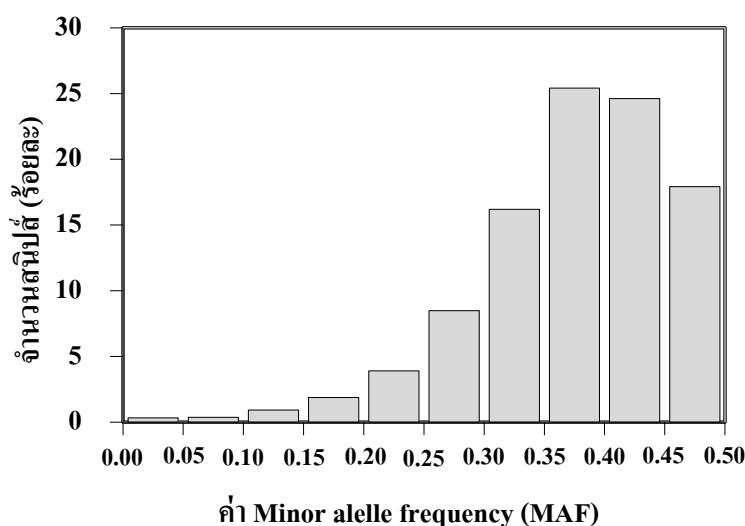
การกระจายตัวของ Minor allele frequency (MAF) บนโครโมโซมโคนมแต่ละคู่ ในประชากรที่ศึกษา พบว่า มีลักษณะแตกต่างกันทั้งโครโมโซมร่างกายและโครโมโซมเพศ (ตารางที่ 3) โดยโครโมโซมร่างกายมีการกระจายตัวของค่า MAF น้อยที่สุดอยู่ในช่วง 0 (โครโมโซมคู่ที่ 1, 16 และ 26) ถึง 0.147 (โครโมโซมคู่ที่ 29) และมีค่า MAF มากที่สุดอยู่ในช่วง 0.497 (โครโมโซมคู่ที่ 5, 17 และ 25) ถึง 0.5 (โครโมโซมคู่ที่ 1, 2, 4, 6, 8, 10, 15, 18, 20, 22 และ 26) และบนโครโมโซมเพศมีค่า MAF น้อยที่สุดเท่ากับ 0.105 และมีค่า MAF มากที่สุดเท่ากับ 0.5 อย่างไรก็ตาม เมื่อพิจารณาการกระจายตัวของ MAF บนโครโมโซมร่างกาย พบว่า มีค่าเฉลี่ยมากที่สุด (0.384) บนโครโมโซมคู่ที่ 8 และมีค่าเฉลี่ยน้อยที่สุด (0.368) บนโครโมโซมคู่ที่ 16 ส่วนโครโมโซมเพศมีค่าเฉลี่ยเท่ากับ 0.357 นอกจากนี้ การกระจายตัวของ MAF ในระดับจีโนมในประชากรโคนมที่ศึกษามีค่าเฉลี่ยเท่ากับ 0.376 โดยค่าเฉลี่ยดังกล่าวมีค่ามากกว่าโคนมที่รายงานโดย The Bovine HapMap Consortium (2009) เช่นพันธุ์โฮลสไตน์ (0.248) บราวน์สวิส (0.239) และเจอร์ซี่ (0.233) เป็นต้น ในทำนองเดียวกับโคนมพันธุ์โฮลสไตน์ (0.32) เจอร์ซี่ (0.28) และ บราวน์สวิส (0.29) ที่ศึกษาโดย Wiggans *et al.* (2012) และโคนมพันธุ์โฮลสไตน์ที่เลี้ยงในประเทศอเมริกาเหนือ (Bohmanova *et al.*, 2010) 0.29 ออสเตรเลีย 0.28 (Khatkar *et al.*, 2008) และเยอรมนี 0.28 (Qanbari *et al.*, 2010) ตามลำดับ

ลักษณะความผันแปรนี้ อาจเป็นผลมาจากโคนมที่เลี้ยงดูภายใต้ฟาร์มเกษตรกรในประเทศไทย มีลักษณะเป็นโคนมพันธุ์แท้และลูกผสมที่เกิดจากการผสมพันธุ์ระหว่างพันธุ์โคในตระกูลโคยุโรป (*Bos taurus*) เช่น โฮลสไตน์ เจอร์ซี่ บราวน์สวิส และเรดเดน เป็นต้น และในตระกูลโคอินเดีย (*Bos indicus*) เช่น ซาฮิวาล บราห์มัน และพื้นเมือง เป็นต้น นอกจากนี้ โคนมบางตัวยังมีองค์ประกอบพันธุ์โคนมที่แตกต่างกันได้มากกว่า 3 พันธุ์ (พจน์ และคณะ, 2557) โดยความแตกต่างของพันธุ์โคดังกล่าว มีผลทำให้โคนมในประเทศไทยมีลักษณะเป็นโคนมหลากหลายพันธุ์ และมีลักษณะที่แตกต่างจากโคนมที่เลี้ยงในประเทศอื่นๆ ซึ่งมีลักษณะเป็นโคนมพันธุ์แท้หรือลูกผสมชั่วแรก (Koonawootrittriron *et al.*, 2009) รวมทั้งอาจมีผลต่อความผันแปรของความถี่อัลลีลสปีส์ที่สูงขึ้น (Wiener *et al.*, 2004) และจากการศึกษาของ Pérez O'Brien *et al.* (2014) พบว่า การกระจายตัวของ MAF ในระดับจีโนมของโคในตระกูลยุโรปและอินเดียมีรูปแบบที่แตกต่างกัน โดยสัดส่วนสปีส์ของโคตระกูลยุโรปจะมีค่าเพิ่มสูงขึ้นเมื่อค่า MAF สูงขึ้น ซึ่งตรงข้ามกับโคตระกูลอินเดียที่มีสัดส่วนสปีส์ลดลง

อย่างไรก็ตาม เมื่อพิจารณาการกระจายตัวของสัดส่วนสปีส์บนโครโมโซมของโคนมในประชากรที่ศึกษา พบว่า มีลักษณะเหมือนกันทั้งโครโมโซมร่างกายและโครโมโซมเพศ (สัดส่วนสปีส์เพิ่มขึ้นเมื่อค่า MAF เพิ่มขึ้น; ภาพผนวก 1-5) นอกจากนี้ ยังพบว่ามีสัดส่วนสปีส์มากกว่าร้อยละ 96 ที่มีการกระจายตัวบนช่วงค่า MAF มากกว่า 0.2 (ภาพที่ 4) ซึ่งมีลักษณะคล้ายกับการศึกษาของ Sargoizaei *et al.* (2008) และ Wiggans *et al.* (2012) ที่พบว่า สัดส่วนสปีส์ในระดับจีโนมของโคนมพันธุ์โฮลสไตน์มีค่าเพิ่มสูงขึ้นเมื่อค่า MAF มากกว่า 0.2 ทั้งนี้ เนื่องจากการใช้ประโยชน์สปีส์ที่มีค่า MAF ต่ำกว่า 0.05 ส่งผลทำให้อำนาจในการตรวจสอบอัตราส่วนของความเสี่ยงทางพันธุกรรมของการกำหนดตำแหน่งผิดที่ (False position) สูงกว่าสปีส์ที่มีค่า MAF สูงกว่า 0.2 อย่างมีนัยสำคัญทางสถิติ (Ardlie *et al.*, 2002)

ด้วยเหตุนี้ ในปัจจุบันการคัดกรองสปีส์ก่อนนำมาใช้ประโยชน์ในการศึกษา LD ความสัมพันธ์ระหว่างข้อมูลสปีส์และลักษณะการให้ผลผลิต และการประเมินค่าทางพันธุกรรมจึงได้พิจารณาค่า MAF ของสปีส์ร่วมด้วย อย่างไรก็ตาม จากการศึกษาของ Tabangin *et al.* (2009) แสดงให้เห็นว่า การเลือกใช้สปีส์ที่มีค่า MAF มากกว่าหรือเท่ากับ 0.005 และ 0.025 ไม่แตกต่างกับการใช้สปีส์ที่มีค่า MAF มากกว่าหรือเท่ากับ 0.05, 0.125 และ 0.250 ในการศึกษาความสัมพันธ์ระหว่างสปีส์และลักษณะการให้ผลผลิต ซึ่งการไม่นำสปีส์ที่มีค่า MAF น้อยกว่า 0.05 มาใช้ในการวิเคราะห์ทำให้มีสปีส์จำนวนมากที่สูญเสียไป ซึ่งอาจมีผลต่อการตรวจสอบสปีส์ที่มีอิทธิพลต่อลักษณะการให้ผลผลิต ดังนั้น การคัดเลือกสปีส์จึงขึ้นอยู่กับผู้วิจัยและในปัจจุบันการศึกษา LD พบว่า มีการคัดเลือกสปีส์ที่มีค่า MAF มากกว่าหรือเท่ากับ 0.05 มาใช้ในการศึกษา (Qanbari *et al.*, 2010; Espigolan *et al.*, 2013) ส่วนการศึกษาความสัมพันธ์ระหว่างข้อมูลจีโนมและลักษณะการ

ให้ผลผลิตในโคนม มีการเลือกใช้สปีส์ที่มีค่า MAF มากกว่า 0.005 (Pryce *et al.*, 2010) 0.015 (Guo *et al.*, 2012) และ 0.025 (Meredith *et al.*, 2012) และการประมาณค่าทางพันธุกรรมจีโนมมีการเลือกใช้สปีส์ที่มีค่า MAF มากกว่า 0.01 (Hayes *et al.*, 2009; Ogawa *et al.*, 2014)



ภาพที่ 4 การกระจายตัวของสปีส์ Minor allele frequency ในระดับจีโนมของโคนม

1.3 ระดับค่า Linkage disequilibrium บนโครโมโซมร่างกายในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

ระดับ LD เป็นการประมาณจากค่า D' และ r^2 สำหรับคู่สปีส์บนโครโมโซมร่างกาย แต่ละคู่ที่มีระยะห่างระหว่างคู่สปีส์ไม่เกิน 5 Mb ซึ่งในประชากรโคนมที่ศึกษา ประกอบด้วย สปีส์ทั้งหมด 8,207 ตำแหน่ง ที่มีค่า MAF มากกว่า 0.05 และทราบตำแหน่งทางกายภาพบนโครโมโซมร่างกาย โดยสามารถสร้างคู่สปีส์ที่ใช้ในการประมาณค่า LD ด้วยระยะห่างระหว่างคู่สปีส์ไม่เกิน 5 Mb จำนวน 129,588 คู่ (ตารางที่ 4) สปีส์เหล่านี้มีการกระจายอยู่บนโครโมโซมร่างกายที่แตกต่างกัน ส่งผลให้จำนวนคู่สปีส์บนโครโมโซมร่างกายแต่ละคู่มีจำนวนที่แตกต่างกัน ซึ่งในประชากรที่ศึกษา พบว่า โครโมโซมคู่ที่ 28 มีจำนวนคู่สปีส์น้อยที่สุด (2,018 คู่) และโครโมโซมคู่ที่ 1 มีจำนวนคู่สปีส์มากที่สุด (8,750 คู่) อย่างไรก็ตาม เมื่อพิจารณาระยะห่างระหว่างคู่สปีส์บนโครโมโซมร่างกายแต่ละคู่ พบว่า ระยะห่างระหว่างคู่สปีส์น้อยที่สุดอยู่ในช่วง 0 (โครโมโซมคู่ที่ 2 และ 3) ถึง 0.020 Mb (โครโมโซมคู่ที่ 4, 12 และ 14) และระยะห่างระหว่างคู่สปีส์มากที่สุดอยู่ในช่วง 4.995 (โครโมโซมคู่ที่ 17) ถึง 0.500 Mb (โครโมโซมคู่ที่ 1, 2, 3, 6, 9, 10, 11 และ 20) และระยะห่างระหว่างคู่สปีส์บนโครโมโซมคู่ที่ 16 มีค่าเฉลี่ยน้อยที่สุด (1.987 Mb) และโครโมโซมคู่ที่ 4 มีค่าเฉลี่ยมากที่สุด (2.532 Mb)

การกระจายตัวของค่า LD (D' และ r^2) มีลักษณะแตกต่างกัน บนโครโมโซมร่างกายแต่ละคู่ โดยค่า D' สำหรับคู่สนิปส์ที่มีระยะห่างไม่เกิน 5 Mb มีค่าน้อยที่สุดเท่ากับ 0 และมีค่ามากที่สุดเท่ากับ 1 บนโครโมโซมร่างกายทุกคู่ และการกระจายตัวของค่า D' ในประชากรโคนมที่ศึกษามีค่าเฉลี่ยน้อยที่สุดบนโครโมโซมคู่ที่ 28 (0.207) และมีค่าเฉลี่ยมากที่สุดบนโครโมโซมคู่ที่ 20 (0.368) อย่างไรก็ตาม การกระจายตัวของ D' บนโครโมโซมในประชากรที่ศึกษา พบว่า มีลักษณะที่แตกต่างกับการกระจายตัวของค่า r^2 ซึ่งมีค่าน้อยที่สุดเท่ากับ 0 (โครโมโซมทุกคู่) และมีค่ามากที่สุดอยู่ในช่วง 0.569 (โครโมโซมคู่ที่ 4) ถึง 1 (โครโมโซมคู่ที่ 3, 5, 6, 15, 19, 20, 21 และ 23) และค่า r^2 มีค่าเฉลี่ยน้อยที่สุดบนโครโมโซมคู่ที่ 22 และ 27 (0.030) และมีค่าเฉลี่ยมากที่สุดบนโครโมโซมคู่ที่ 16 (0.090) ซึ่งระดับ LD ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทยนี้ คล้ายคลึงกับ โคนมพันธุ์โฮลสไตน์ที่รายงานโดย Khatkar *et al.* (2008) สำหรับคู่สนิปส์ที่มีระยะห่างไม่เกิน 5 Mb อย่างไรก็ตาม ระดับ LD ดังกล่าวเป็นระดับที่ต่ำเมื่อเทียบกับการศึกษาในปัจจุบัน เช่น โคนมพันธุ์โฮลสไตน์ที่รายงานโดย Sargolzaei *et al.* (2008) และ Bohmanova *et al.* (2010) นอกจากนี้ ในประชากรโคนมที่ศึกษา พบว่า มีโครโมโซมร่างกายบางคู่ที่มีระดับค่า LD สูงกว่าโครโมโซมร่างกายคู่อื่นๆ สอดคล้องกับการศึกษาของ Bohmanova *et al.* (2010) และ Espigolan *et al.* (2013) ซึ่งเป็นผลมาจากอัตราการรวมอีกครั้งระหว่างข้อมูลพันธุกรรมสนิปส์มีความสัมพันธ์กับความยาวโครโมโซม (Farre *et al.*, 2013)

นอกจากนี้ ระดับ LD (D' และ r^2) บนโครโมโซมร่างกายแต่ละคู่ มีแนวโน้มลดลงเมื่อระยะห่างระหว่างคู่สนิปส์เพิ่มสูงขึ้น โดย D' มีแนวโน้มลดลงต่ำและมีรูปแบบการลดลงที่ไม่ชัดเจน (ภาพผนวก 6-10) ซึ่งแตกต่างกับ r^2 ที่มีแนวโน้มลดลงชัดเจนเมื่อระยะห่างระหว่างสนิปส์เพิ่มสูงขึ้น (ภาพผนวก 11-15) นอกจากนี้ ระดับ LD บนจีโนมโคนมทั้งหมดมีค่าลดลงเมื่อระยะห่างระหว่างคู่สนิปส์เพิ่มขึ้น (ภาพที่ 5 และ 6) ซึ่งสอดคล้องกับผลการศึกษาของ Khatkar *et al.* (2008) Sargolzaei *et al.* (2008) และ Bohmanova *et al.* (2010) สำหรับโคนมพันธุ์โฮลสไตน์ รวมทั้งโคนมพันธุ์เจอร์ซี่ ที่รายงานโดย de Roos *et al.* (2008) ซึ่งรูปแบบการลดลงของระดับ LD ได้รับอิทธิพลมาจากหลายปัจจัย เช่น Genetic drift, Admixture และ Inbreeding ซึ่งมีลักษณะจำเพาะกับประชากร รวมทั้งอัตราการรวมตัวอีกครั้งระหว่างข้อมูลพันธุกรรม (Recombination rate) การเปลี่ยนแปลงของยีน (Gene conversion) เช่น การแลกเปลี่ยนชิ้นส่วนโครโมโซม (Crossing over) เกิดขึ้นในกระบวนการสร้างเซลล์สืบพันธุ์ (Meiosis) ของสิ่งมีชีวิตโดยเฉพาะในระยะ Prophase I และการคัดเลือก (Selection) ที่จำเพาะกับพื้นที่จีโนม (Ardlie *et al.*, 2002) ดังนั้น ค่า LD จึงเป็นค่าที่จำเพาะกับประชากรที่ศึกษา เนื่องจากการประมาณค่า LD เป็นการประมาณค่าจากคู่สนิปส์บนโครโมโซมแต่ละคู่ ซึ่งสนิปส์เหล่านั้นมีความถี่อัลลีลที่ต่างกัน นอกจากนี้ เกณฑ์การพิจารณาสนิปส์และความหนาแน่นของสนิปส์ที่ใช้ในการศึกษา รวมถึงซอฟต์แวร์ที่ใช้ในการประมาณค่า LD ที่แตกต่างกัน (Pritchard and Przeworski, 2001)

ตารางที่ 5 จำนวนคู่สนิปส์ ค่าต่ำสุด ค่าสูงสุด ค่าเฉลี่ยของระยะห่างระหว่างสนิปส์ และค่า LD (D' และ r^2) บนโครโมโซมร่างกายที่ระยะห่างระหว่างสนิปส์ไม่เกิน 5 Mb ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

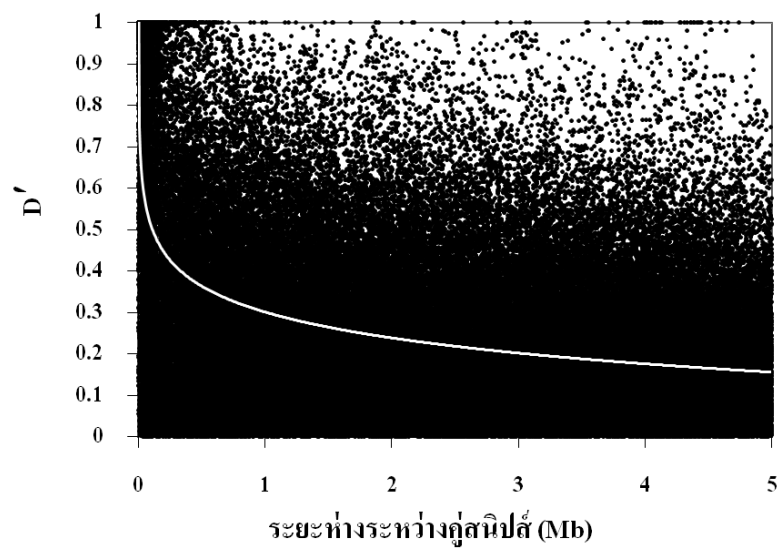
โครโมโซม คู่ที่	จำนวนสนิปส์ (ตำแหน่ง)	จำนวน คู่สนิปส์	ระยะห่างระหว่างคู่สนิปส์ (Mb)			D'			r^2		
			ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย $\pm$ SD	ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย $\pm$ SD	ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย $\pm$ SD
1	530	8,750	0.001	5.000	2.350 $\pm$ 1.496	0.000	1.000	0.278 $\pm$ 0.247	0.000	0.998	0.057 $\pm$ 0.112
2	432	6,477	0.000	5.000	2.435 $\pm$ 1.458	0.000	1.000	0.270 $\pm$ 0.220	0.000	0.995	0.046 $\pm$ 0.078
3	405	7,104	0.000	5.000	2.219 $\pm$ 1.550	0.000	1.000	0.297 $\pm$ 0.257	0.000	1.000	0.065 $\pm$ 0.135
4	359	4,992	0.020	4.999	2.532 $\pm$ 1.413	0.000	1.000	0.228 $\pm$ 0.180	0.000	0.569	0.035 $\pm$ 0.051
5	380	5,911	0.001	4.999	2.339 $\pm$ 1.506	0.000	1.000	0.305 $\pm$ 0.280	0.000	1.000	0.051 $\pm$ 0.107
6	441	8,699	0.001	5.000	2.424 $\pm$ 1.433	0.000	1.000	0.248 $\pm$ 0.207	0.000	1.000	0.043 $\pm$ 0.083
7	346	4,865	0.002	4.998	2.480 $\pm$ 1.427	0.000	1.000	0.236 $\pm$ 0.194	0.000	0.998	0.040 $\pm$ 0.072
8	347	4,870	0.019	4.997	2.521 $\pm$ 1.440	0.000	1.000	0.243 $\pm$ 0.196	0.000	0.899	0.043 $\pm$ 0.070
9	326	4,826	0.001	5.000	2.393 $\pm$ 1.477	0.000	1.000	0.281 $\pm$ 0.239	0.000	0.997	0.048 $\pm$ 0.089
10	319	4,605	0.005	5.000	2.493 $\pm$ 1.419	0.000	1.000	0.244 $\pm$ 0.191	0.000	0.939	0.043 $\pm$ 0.066
11	340	4,937	0.018	5.000	2.479 $\pm$ 1.422	0.000	1.000	0.230 $\pm$ 0.185	0.000	0.987	0.035 $\pm$ 0.059
12	266	3,576	0.020	4.999	2.444 $\pm$ 1.438	0.000	1.000	0.232 $\pm$ 0.197	0.000	0.984	0.037 $\pm$ 0.071
13	270	3,930	0.010	4.999	2.434 $\pm$ 1.439	0.000	1.000	0.254 $\pm$ 0.208	0.000	0.994	0.042 $\pm$ 0.070
14	274	4,090	0.020	4.999	2.485 $\pm$ 1.421	0.000	1.000	0.266 $\pm$ 0.205	0.000	0.912	0.047 $\pm$ 0.074
15	270	4,079	0.001	4.996	2.428 $\pm$ 1.462	0.000	1.000	0.222 $\pm$ 0.198	0.000	1.000	0.039 $\pm$ 0.083
16	298	5,969	0.001	4.999	1.987 $\pm$ 1.584	0.000	1.000	0.360 $\pm$ 0.306	0.000	1.000	0.090 $\pm$ 0.179
17	222	2,903	0.010	4.995	2.484 $\pm$ 1.419	0.000	1.000	0.234 $\pm$ 0.189	0.000	0.997	0.038 $\pm$ 0.066

ตารางที่ 5 (ต่อ)

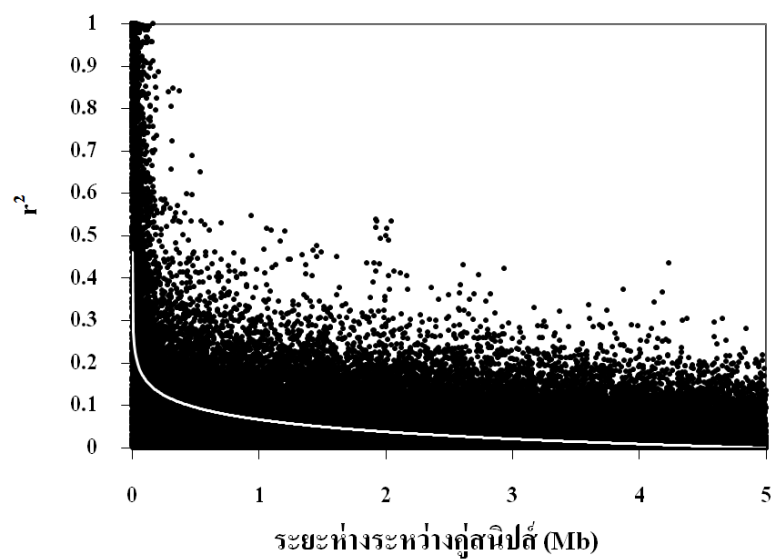
โครโมโซม	จำนวนสปีส์ (ตำแหน่ง)	จำนวน คู่สปีส์	ระยะห่างระหว่างคู่สปีส์ (Mb)			D'			r ²		
			ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย ± SD	ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย ± SD	ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย ± SD
18	231	3,666	0.001	4.999	2.094 ± 1.547	0.000	1.000	0.242 ± 0.229	0.000	0.998	0.036 ± 0.075
19	217	3,437	0.001	4.999	2.396 ± 1.476	0.000	1.000	0.224 ± 0.211	0.000	1.000	0.044 ± 0.111
20	284	5,875	0.001	5.000	2.062 ± 1.625	0.000	1.000	0.368 ± 0.305	0.000	1.000	0.088 ± 0.180
21	247	4,353	0.001	4.999	2.138 ± 1.566	0.000	1.000	0.307 ± 0.268	0.000	1.000	0.061 ± 0.127
22	199	2,933	0.017	4.998	2.449 ± 1.430	0.000	1.000	0.211 ± 0.175	0.000	0.799	0.030 ± 0.051
23	203	4,025	0.001	4.999	2.175 ± 1.562	0.000	1.000	0.268 ± 0.252	0.000	1.000	0.067 ± 0.152
24	211	3,214	0.000	4.999	2.497 ± 1.421	0.000	1.000	0.224 ± 0.182	0.000	0.999	0.036 ± 0.065
25	159	2,472	0.009	4.998	2.378 ± 1.448	0.000	1.000	0.235 ± 0.210	0.000	0.990	0.034 ± 0.061
26	172	2,489	0.010	4.999	2.428 ± 1.452	0.000	1.000	0.230 ± 0.197	0.000	0.650	0.031 ± 0.048
27	153	2,301	0.019	4.996	2.437 ± 1.441	0.000	1.000	0.221 ± 0.196	0.000	0.993	0.030 ± 0.053
28	147	2,018	0.001	4.989	2.407 ± 1.434	0.000	1.000	0.207 ± 0.179	0.000	0.877	0.031 ± 0.061
29	159	2,192	0.004	4.997	2.443 ± 1.419	0.000	1.000	0.217 ± 0.183	0.000	0.996	0.036 ± 0.069

หมายเหตุ: SD คือ ส่วนเบี่ยงเบนมาตรฐาน

ระดับ LD เป็นการศึกษาข้อมูลพื้นฐานของสนิปส์ในระดับจีโนม ซึ่งนำไปสู่ความรู้และความเข้าใจเกี่ยวกับกระบวนการถ่ายทอดทางพันธุกรรมของสิ่งมีชีวิต เนื่องจากค่า D' และ r^2 เป็นค่าที่ใช้บ่งบอกความสัมพันธ์ระหว่างตำแหน่งสนิปส์บนโครโมโซมแต่ละคู่ ที่มีการถ่ายทอดข้อมูลพันธุกรรมมาจากบรรพบุรุษ ด้วยเหตุนี้ ความสัมพันธ์ของสนิปส์ในรูป LD จึงถูกนำมาใช้สร้างแผนที่ QTL ระหว่างคู่สนิปส์ที่อยู่บนพื้นที่ยีนบนโครโมโซม เพื่อระบุพื้นที่โครโมโซมที่มีความสัมพันธ์กับลักษณะการให้ผลผลิตของโคนม (Goddard and Hayes, 2009) และในปัจจุบันระดับ LD ระหว่างคู่สนิปส์ถูกใช้เป็นเกณฑ์การคัดกรองสนิปส์ในการศึกษาความสัมพันธ์ระหว่างสนิปส์กับลักษณะการให้ผลผลิต โดยจากการศึกษา Dunning *et al.* (2000) พบว่า ระดับ LD ของคู่สนิปส์ที่มากกว่า 0.3 สามารถเพิ่มความสามารถในการทดสอบความสัมพันธ์ระหว่างสนิปส์และปริมาณการให้ผลผลิตของโคนมได้ อย่างไรก็ตาม ผลจากการศึกษา Bohmanova *et al.* (2010) พบว่า สัดส่วนคู่สนิปส์ในประชากรโคนมพันธุ์โฮลสไตน์มีเพียงร้อยละ 30 ที่มีระดับ LD มากกว่า 0.3 และเป็นคู่สนิปส์ที่มีระยะห่างน้อยกว่า 40 kb สำหรับโคนมพันธุ์โฮลสไตน์ที่เลี้ยงในประเทศออสเตรเลีย (Khatkar *et al.*, 2008) 75 kb สำหรับโคนมพันธุ์โฮลสไตน์ที่เลี้ยงในประเทศเยอรมนี (Qanbari *et al.*, 2010) และ 100 kb สำหรับโคนมพันธุ์โฮลสไตน์ที่เลี้ยงในประเทศอเมริกาเหนือ (Sargolzaei *et al.*, 2008) นอกจากนี้ Meuwissen *et al.* (2001) รายงานระดับ LD ที่มีค่ามากกว่า 0.2 สามารถช่วยเพิ่มความแม่นยำในการประมาณค่าความสามารถทางพันธุกรรมในโคนม



ภาพที่ 5 การกระจายของค่า D' บนระยะห่างระหว่างคู่สnp ไม่เกิน 5 Mb บนจีโนมในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย



ภาพที่ 6 การกระจายของค่า r^2 บนระยะห่างระหว่างคู่สnp ไม่เกิน 5 Mb บนจีโนมในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

2. ศึกษาปัจจัยที่มีผลต่อระดับ Linkage disequilibrium ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

2.1 อิทธิพลของระยะห่างระหว่างคู่สนิปส์ต่อระดับ LD ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

จากผลการศึกษาที่ 1 พบว่า ประชากรโคนมหลากหลายพันธุ์ในประเทศไทย มีระดับ LD (D' และ r^2) ลดลงเมื่อระยะห่างระหว่างคู่สนิปส์เพิ่มสูงขึ้น ดังนั้น เพื่อทดสอบการลดลงของระดับ LD ในจีโนมในประชากรโคนมที่ศึกษา คู่สนิปส์ทั้งหมดบนโครโมโซมร่างกายถูกจำแนกเป็น 3 กลุ่ม โดยกลุ่มที่ 1 ประกอบด้วย คู่สนิปส์จำนวน 10 กลุ่ม ที่มีระยะห่างระหว่างคู่สนิปส์เพิ่มขึ้นทุก 10 kb (0 ถึง 100 kb) กลุ่มที่ 2 ประกอบด้วย คู่สนิปส์จำนวน 9 กลุ่ม ที่มีระยะห่างระหว่างคู่สนิปส์เพิ่มขึ้นทุก 100 kb (100 ถึง 1,000 kb) และกลุ่มที่ 3 ประกอบด้วย คู่สนิปส์จำนวน 4 กลุ่ม ที่มีระยะห่างระหว่างคู่สนิปส์เพิ่มขึ้นทุก 1 Mb (1 ถึง 5 Mb; ตารางที่ 5)

ผลการศึกษา พบว่า ระยะห่างระหว่างคู่สนิปส์ที่เพิ่มขึ้นในแต่ละกลุ่ม มีระดับ LD (D' และ r^2) แตกต่างกัน ($P < 0.05$) โดยกลุ่มที่ 1 มีค่า D' น้อยที่สุดในช่วง 0 (30 ถึง 40 kb และ 60 ถึง 70 kb) ถึง 0.038 (0 ถึง 10 kb) และมีค่า D' มากที่สุดเท่ากับ 1 (บนระยะห่างระหว่างคู่สนิปส์ทุกกลุ่ม) และการกระจายตัวของคู่สนิปส์ในแต่ละกลุ่มมีค่าเฉลี่ย D' ลดลงในระดับที่แตกต่างกัน ($P < 0.05$) เมื่อระยะห่างระหว่างคู่สนิปส์เพิ่มสูงขึ้น โดยระยะห่างระหว่างคู่สนิปส์ 0 ถึง 10 kb มีค่าเฉลี่ย D' มากที่สุด (0.904) รองลงมาคือ ระยะห่างระหว่างคู่สนิปส์ 10 ถึง 20 kb (0.805) และ 20 ถึง 30 kb (0.768) ตามลำดับ และค่าเฉลี่ย D' น้อยที่สุด (0.511) พบในกลุ่มคู่สนิปส์ที่มีระยะห่างระหว่าง 90 ถึง 100 kb อย่างไรก็ตาม เมื่อพิจารณาค่า r^2 พบว่า มีค่าเฉลี่ยที่แตกต่างกันเมื่อระยะห่างระหว่างคู่สนิปส์เพิ่มสูงขึ้น ($P < 0.05$) ยกเว้น คู่สนิปส์ที่มีระยะห่างระหว่าง 40 ถึง 50 kb และ 50 ถึง 60 kb อีกทั้งคู่สนิปส์ที่มีระยะห่างระหว่าง 70 ถึง 80 kb ไม่พบความแตกต่างกับคู่สนิปส์ที่มีระยะห่างระหว่าง 60 ถึง 70 kb และ 80 ถึง 90 kb โดย r^2 มีค่าน้อยที่สุดเท่ากับ 0 (บนระยะห่างระหว่างคู่สนิปส์ทุกกลุ่ม) และมีค่ามากที่สุดในช่วง 0.886 (20 ถึง 30 kb) ถึง 1 (0 ถึง 10 kb, 10 ถึง 20 kb, 30 ถึง 40 kb และ 40 ถึง 50 kb) และการกระจายตัวค่า r^2 ของคู่สนิปส์ที่มีระยะห่างระหว่าง 0 ถึง 10 kb มีค่ามากที่สุด (0.515) รองลงมา คือ คู่สนิปส์ที่มีระยะห่างระหว่าง 10 ถึง 20 kb (0.321) และ 20 ถึง 30 kb (0.278) ตามลำดับ และค่าเฉลี่ย r^2 น้อยที่สุด (0.127) พบในคู่สนิปส์ที่มีระยะห่างระหว่าง 90 ถึง 100 kb

อย่างไรก็ตาม เมื่อพิจารณาระดับ LD บนช่วงระยะห่างระหว่างคู่สนิปส์ที่เพิ่มขึ้นในกลุ่มที่ 2 พบว่า ระดับ LD มีค่าเฉลี่ยที่แตกต่างกัน ($P < 0.05$) เมื่อระยะห่างระหว่างคู่สนิปส์เพิ่มสูงขึ้น และการกระจายตัวของคู่สนิปส์ในแต่ละกลุ่มมีค่า D' และ r^2 ที่แตกต่างกัน ($P < 0.05$) ในคู่สนิปส์ที่มีระยะห่างระหว่าง 100 ถึง 200 kb และ 200 ถึง 300 kb (ตารางที่ 5) โดยค่า D' มีค่าในช่วง 0 ถึง 1 บนระยะห่างระหว่างคู่สนิปส์ทุกกลุ่ม และคู่สนิปส์ที่มีระยะห่างระหว่าง 100 ถึง 200 kb (0.472) มีค่าเฉลี่ยมากที่สุด รองลงมาคือ คู่สนิปส์ที่มีระยะห่างระหว่าง 200 ถึง 300 kb (0.351) และค่าเฉลี่ยน้อยที่สุด (0.260) พบในคู่สนิปส์ที่มีระยะห่างระหว่าง 900 ถึง 1,000 kb และการกระจายตัวของค่า r^2 ในแต่ละกลุ่มมีค่าน้อยที่สุดเท่ากับ 0 (บนระยะห่างระหว่างคู่สนิปส์ทุกกลุ่ม) และมีค่ามากที่สุดในช่วง 0.459 (800 ถึง 900 kb) ถึง 1 (100 ถึง 200 kb) ส่วนคู่สนิปส์ที่มีระยะห่างระหว่าง 100 ถึง 200 kb มีค่าเฉลี่ย r^2 มากที่สุด (0.116) รองลงมาคือ คู่สนิปส์ที่มีระยะห่างระหว่าง 200 ถึง 300 kb (0.070) และมีค่าเฉลี่ย r^2 น้อยที่สุด (0.031) บนระยะห่างระหว่างคู่สนิปส์ 900 ถึง 1,000 kb

นอกจากนี้ ระดับค่า LD บนช่วงระยะห่างระหว่างคู่สนิปส์ในกลุ่มที่ 3 พบว่า คู่สนิปส์ในแต่ละกลุ่มมีระดับ LD (D' และ r^2) ที่แตกต่างกัน ($P < 0.05$) โดย ค่า D' มีค่าในช่วง 0 ถึง 1 บนระยะห่างระหว่างคู่สนิปส์ทุกกลุ่ม และมีค่าเฉลี่ย D' มากที่สุด (0.247) บนระยะห่างระหว่างคู่สนิปส์ 1 ถึง 2 Mb และมีค่าเฉลี่ยน้อยที่สุด (0.191) บนระยะห่างระหว่างคู่สนิปส์ 4 ถึง 5 Mb ส่วนค่า r^2 พบว่า มีค่าน้อยที่สุดเท่ากับ 0 (บนระยะห่างระหว่างคู่สนิปส์ทุกกลุ่ม) และมีค่ามากที่สุดในช่วง 0.023 (4 ถึง 5 Mb) ถึง 0.039 (1 ถึง 2 Mb)

ตารางที่ 6 จำนวนคู่สนิปส์ ค่าต่ำสุด ค่าสูงสุด ค่าเฉลี่ยลิสแควร์ ความคลาดเคลื่อนมาตรฐานของ LD (D' และ r^2) ที่ระยะห่างระหว่างสนิปส์บนจีโนมที่แตกต่างกันในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

ระยะห่างระหว่าง คู่สนิปส์	จำนวน คู่สนิปส์	D'			ร้อยละคู่สนิปส์ $D' \geq 0.8$	r^2			ร้อยละคู่สนิปส์ $r^2 \geq 0.3$
		ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย $\pm$ SE		ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย $\pm$ SE	
0 ถึง 10 kb	997	0.038	1.000	0.904 ± 0.006^a	84.65	0.000	1.000	0.515 ± 0.003^a	61.08
10 ถึง 20 kb	841	0.005	1.000	0.805 ± 0.007^b	69.56	0.000	1.000	0.321 ± 0.003^b	38.41
20 ถึง 30 kb	987	0.004	1.000	0.768 ± 0.006^c	64.13	0.000	0.886	0.278 ± 0.003^c	33.64
30 ถึง 40 kb	832	0.000	1.000	0.730 ± 0.007^d	57.21	0.000	1.000	0.246 ± 0.003^d	29.33
40 ถึง 50 kb	749	0.012	1.000	0.694 ± 0.007^e	50.47	0.000	1.000	0.202 ± 0.003^e	23.77
50 ถึง 60 kb	699	0.001	1.000	0.644 ± 0.007^e	48.21	0.000	0.998	0.188 ± 0.003^e	22.60
60 ถึง 70 kb	641	0.000	1.000	0.572 ± 0.008^f	37.44	0.000	0.998	0.167 ± 0.003^f	17.94
70 ถึง 80 kb	565	0.003	1.000	0.548 ± 0.008^{fg}	29.73	0.000	0.995	0.164 ± 0.003^f	19.65
80 ถึง 90 kb	510	0.001	1.000	0.529 ± 0.009^{fg}	31.76	0.000	0.995	0.150 ± 0.004^f	16.86
90 ถึง 100 kb	501	0.001	1.000	0.511 ± 0.009^g	26.55	0.000	0.995	0.127 ± 0.004^g	11.38
100 ถึง 200 kb	3,247	0.000	1.000	0.472 ± 0.003^a	23.53	0.000	1.000	0.116 ± 0.001^a	10.22
200 ถึง 300 kb	2,245	0.000	1.000	0.351 ± 0.004^b	8.20	0.000	0.848	0.070 ± 0.002^b	3.16
300 ถึง 400 kb	2,444	0.000	1.000	0.315 ± 0.004^c	5.16	0.000	0.689	0.062 ± 0.002^{bc}	2.09
400 ถึง 500 kb	2,625	0.000	1.000	0.306 ± 0.004^{cd}	4.11	0.000	0.650	0.057 ± 0.002^{cd}	2.02
500 ถึง 600 kb	2,490	0.000	1.000	0.293 ± 0.004^{de}	3.65	0.000	0.513	0.053 ± 0.002^{de}	1.24
600 ถึง 700 kb	2,680	0.000	1.000	0.274 ± 0.004^{ef}	1.98	0.000	0.530	0.046 ± 0.002^{ef}	0.93

ตารางที่ 6 (ต่อ)

ระยะห่างระหว่าง คู่สนิปส์	จำนวน คู่สนิปส์	D'			ร้อยละคู่สนิปส์ D' ≥ 0.8	r ²			ร้อยละคู่สนิปส์ r ² ≥ 0.3
		ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย ± SE		ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย ± SE	
700 ถึง 800 kb	2,711	0.000	1.000	0.278 ± 0.004 ^{ef}	2.73	0.000	0.459	0.047 ± 0.002 ^{ef}	0.77
800 ถึง 900 kb	2,632	0.000	1.000	0.266 ± 0.004 ^f	1.90	0.000	0.546	0.046 ± 0.002 ^{ef}	0.84
900 ถึง 1,000 kb	2,667	0.000	1.000	0.260 ± 0.004 ^f	1.87	0.000	0.538	0.043 ± 0.002 ^f	0.56
1 ถึง 2 Mb	25,464	0.000	1.000	0.247 ± 0.001 ^a	1.25	0.000	0.535	0.039 ± 0.001 ^a	0.35
2 ถึง 3 Mb	24,895	0.000	1.000	0.224 ± 0.001 ^b	0.76	0.000	0.374	0.032 ± 0.001 ^b	0.15
3 ถึง 4 Mb	24,354	0.000	1.000	0.208 ± 0.001 ^c	0.62	0.000	0.436	0.027 ± 0.001 ^c	0.02
4 ถึง 5 Mb	23,782	0.000	1.000	0.191 ± 0.001 ^d	0.48	0.000	1.000	0.023 ± 0.001 ^d	0.03

หมายเหตุ: ค่าเฉลี่ยลีสแควร์ที่มีตัวกแตกต่างกันในแต่ละแถวแนวนตั้ง แสดงความแตกต่างอย่างมีนัยสำคัญทางสถิติ (P<0.05)

SE คือ ความคลาดเคลื่อนมาตรฐาน

ผลจากการศึกษานี้ชี้ให้เห็นว่า ระยะห่างระหว่างคู่สนิปส์ในระดับจีโนมของประชากรโคนมที่ศึกษามีอิทธิพลทำให้ระดับ LD ลดลง ซึ่งสอดคล้องกับการรายงานของ Sargolzaei *et al.* (2008) ที่พบว่า การลดลงของระดับ LD ในระดับจีโนมของโคนมเป็นผลจากระยะห่างระหว่างคู่สนิปส์ที่เพิ่มสูงขึ้น ด้วยเหตุนี้ ระดับ LD ที่มีค่าสูง ($r^2 \geq 0.3$) ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทยจึงพบเฉพาะคู่สนิปส์ที่มีระยะห่างน้อยกว่า 30 kb ซึ่งลักษณะความผันแปรเช่นนี้ เป็นผลมาจากระยะห่างระหว่างคู่สนิปส์ที่เพิ่มขึ้นมีความสัมพันธ์กับอัตราการรวมตัวอีกครั้งระหว่างข้อมูลพันธุกรรมที่ลดลง (Farre *et al.*, 2013) นอกจากนี้ การเกิดการเปลี่ยนแปลงของยีนและการกลายพันธุ์ก็มีผลทำให้โอกาสในการถ่ายทอดร่วมกันของคู่สนิปส์ เมื่อเกิดกระบวนการรวมตัวอีกครั้งระหว่างข้อมูลพันธุกรรมลดลง โดยเฉพาะคู่สนิปส์ที่มีระยะห่างมาก (Ardlie *et al.*, 2002)

อย่างไรก็ตาม การกระจายตัวของระดับ LD ที่มีค่าสูง ซึ่งพิจารณาจากค่า $D' \geq 0.8$ และ $r^2 \geq 0.3$ พบว่า ในประชากรโคนมที่ศึกษา มีสัดส่วนคู่สนิปส์ที่มีระดับ LD สูงแตกต่างกันในแต่ละช่วงห่างระหว่างสนิปส์ โดยคู่สนิปส์ที่มีระยะห่างระหว่าง 0 ถึง 10 kb มีสัดส่วนคู่สนิปส์ที่มีค่า $D' \geq 0.8$ สูงสุด (ร้อยละ 84.65) นอกจากนี้ การกระจายตัวของสัดส่วนคู่สนิปส์ที่มีค่า $D' \geq 0.8$ ลดลงเมื่อระยะห่างระหว่างคู่สนิปส์เพิ่มขึ้นจากร้อยละ 84.65 (0 ถึง 10 kb) ถึง 0.48 (4 ถึง 5 Mb)

ในทำนองเดียวกับสัดส่วนคู่สนิปส์ที่มีค่า $r^2 \geq 0.3$ พบว่า คู่สนิปส์ที่มีระยะห่างระหว่าง 0 ถึง 10 kb มีสัดส่วนคู่สนิปส์มากที่สุด (ร้อยละ 61.08) และมีสัดส่วนคู่สนิปส์ลดลงจากร้อยละ 61.08 (0 ถึง 10 kb) ถึง 0.02 (3 ถึง 4 Mb) โดยการกระจายตัวของคู่สนิปส์ที่มีค่า $D' \geq 0.8$ มีสัดส่วนคู่สนิปส์ที่ลดลงเมื่อระยะห่างระหว่างคู่สนิปส์เพิ่มสูงขึ้นในอัตราที่ต่ำ เมื่อเทียบกับการลดลงของค่า r^2 ซึ่งมีสัดส่วนคู่สนิปส์ลดลงอย่างรวดเร็วเมื่อระยะห่างระหว่างคู่สนิปส์เพิ่มขึ้น สอดคล้องกับผลการศึกษาของ Khatkar *et al.* (2008) ที่พบว่า r^2 มีค่าลดลงอย่างรวดเร็วเมื่อระยะห่างระหว่างคู่สนิปส์เพิ่มขึ้นเมื่อเทียบกับค่า D'

นอกจากนี้ ในประชากรโคนมที่ศึกษา พบว่า มีคู่สนิปส์บางส่วนที่มีระดับ LD สูง ซึ่งเห็นได้จากคู่สนิปส์ที่มีระยะห่างระหว่าง 10 ถึง 20 kb ที่มีคู่สนิปส์ร้อยละ 69.56 ($D' \geq 0.8$) และ 38.41 ($r^2 \geq 0.3$) จากข้อมูลสนิปส์ทั้งหมด ลักษณะความผันแปรนี้ คล้ายคลึงกับโคนมพันธุ์โฮลสไตน์ที่รายงานโดย Sargolzaei *et al.* (2008) และ Bohmanova *et al.* (2010) โดยลักษณะดังกล่าว เป็นผลเนื่องจากการกระจายตัวของสนิปส์บนโครโมโซมของโคนม ซึ่งมีลักษณะการ

กระจายตัวไม่สม่ำเสมอและมีลักษณะอยู่รวมกันเป็นกลุ่มบนพื้นที่โครโมโซมบางส่วน มีผลทำให้การกระจายตัวของคู่สนิปส์มีอัตราการรวมตัวอีกครั้งระหว่างข้อมูลพันธุกรรมสูง (Farre *et al.*, 2013) นอกจากนี้ ประชากรโคนมหลากหลายพันธุ์ในประเทศไทย ได้รับการคัดเลือกโดยเกษตรกรผู้เลี้ยงโคนมมาโดยตลอด โดยมีเป้าหมายหลักของการคัดเลือกคือลักษณะการให้ผลผลิตน้ำนม ซึ่งอาจมีผลทำให้เกิดการเปลี่ยนแปลงความถี่อัลลีลของสนิปส์บางตำแหน่งบริเวณพื้นที่ยีนที่ควบคุมลักษณะการให้ผลผลิตน้ำนมมีระดับ LD สูง (Ardlie *et al.*, 2002)

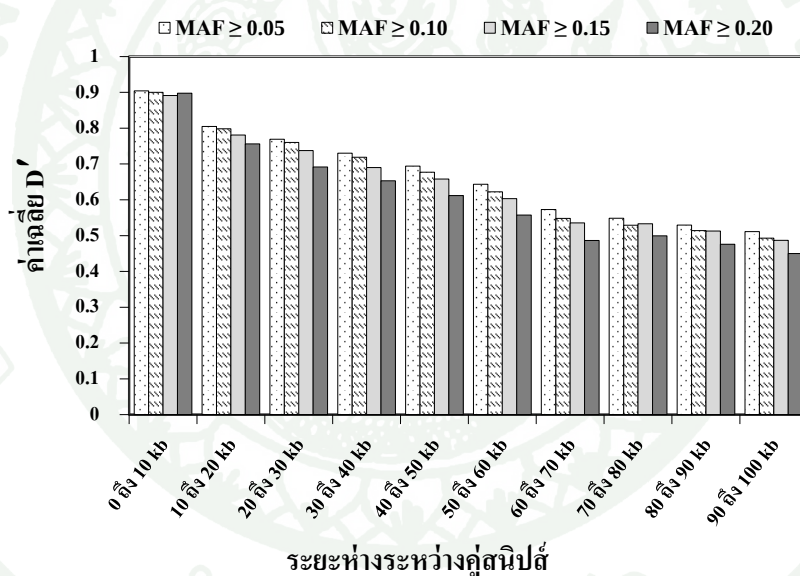
2.2 อิทธิพลของระดับค่า MAF ต่อระดับ LD ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

จากการประมาณค่า LD ในระดับจีโนมของโคนมในประชากรที่ศึกษา ด้วยระดับค่า MAF ที่แตกต่างกัน 4 กลุ่ม คือ $MAF \geq 0.05$, 0.10, 0.15 และ 0.20 พบว่า ค่าเฉลี่ย D' แตกต่างกันบนช่วงห่างระหว่างคู่สนิปส์ที่เพิ่มขึ้นทุก 10 kb โดยเฉพาะคู่สนิปส์ที่มีระยะห่างมากกว่า 10 kb ซึ่งพบว่า ค่าเฉลี่ย D' ในกลุ่ม $MAF \geq 0.05$ มีค่ามากที่สุด รองลงมาได้แก่ กลุ่ม $MAF \geq 0.10$, 0.15 และ 0.20 ตามลำดับ (ภาพที่ 7) เช่นเดียวกับค่าเฉลี่ย D' บนช่วงห่างระหว่างคู่สนิปส์ที่เพิ่มขึ้นทุก 100 kb และ 1 Mb (ภาพที่ 9) ซึ่งสอดคล้องกับการศึกษา Bohmanova *et al.* (2010) สำหรับโคนมพันธุ์โฮลสไตน์ในประเทศสหรัฐอเมริกา และในประเทศออสเตรเลียที่รายงานโดย Khatkar *et al.* (2008) โดยลักษณะความผันแปรนี้ เป็นผลเนื่องจากการประมาณค่า D' ขึ้นอยู่กับความถี่อัลลีลของสนิปส์ ดังนั้นกรณีที่ใช้สนิปส์ที่มีความถี่อัลลีลต่ำในการประมาณค่า D' จึงมีผลทำให้ค่าที่ได้มีค่ามากกว่าค่าแท้จริง (Ardlie *et al.*, 2002)

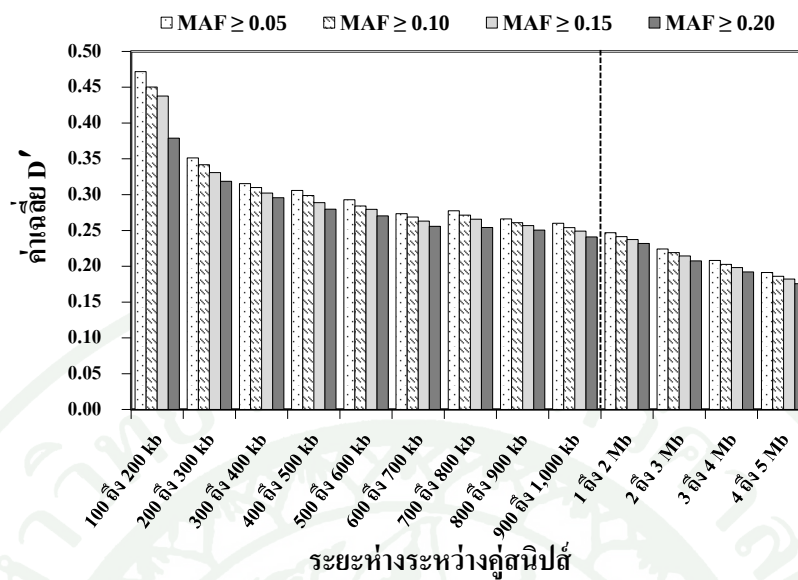
อย่างไรก็ตาม เมื่อเปรียบเทียบอิทธิพลของค่า MAF ระหว่าง D' และ r^2 พบว่า มีความแตกต่างกัน โดยค่าเฉลี่ยของ r^2 บนช่วงห่างระหว่างคู่สนิปส์ที่เพิ่มขึ้นทุก 10 kb และมีระยะห่างไม่เกิน 30 kb กลุ่มสนิปส์ที่มีค่า $MAF \geq 0.20$ มีค่าเฉลี่ย r^2 มากที่สุด รองลงมาได้แก่ กลุ่มสนิปส์ที่มีค่า $MAF \geq 0.15$, 0.10 และ 0.05 ตามลำดับ (ภาพที่ 9) นอกจากนี้ ค่าเฉลี่ย r^2 ในแต่ละกลุ่ม MAF บนช่วงห่างระหว่างคู่สนิปส์ที่เพิ่มขึ้นทุก 100 kb และ 1 Mb ไม่มีความแตกต่างกัน (ภาพที่ 10) ซึ่งความแตกต่างระหว่างค่าเฉลี่ย D' และ r^2 ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทยที่ศึกษานี้ อาจชี้ให้เห็นว่าการประมาณค่า LD จากค่า D' ได้รับอิทธิพลจากความถี่อัลลีลของสนิปส์มากกว่าค่า r^2 และ D' ที่มีค่าสูงอาจไม่สามารถบ่งบอกความสามารถในถ่ายทอดรวมกันได้ ซึ่ง

คล้ายคลึงกับ Khatkar *et al.* (2008) สำหรับประชากรโคนมพันธุ์โฮลสไตน์ และ Espigolan *et al.* (2013) สำหรับประชากรโคพันธุ์ Nellore

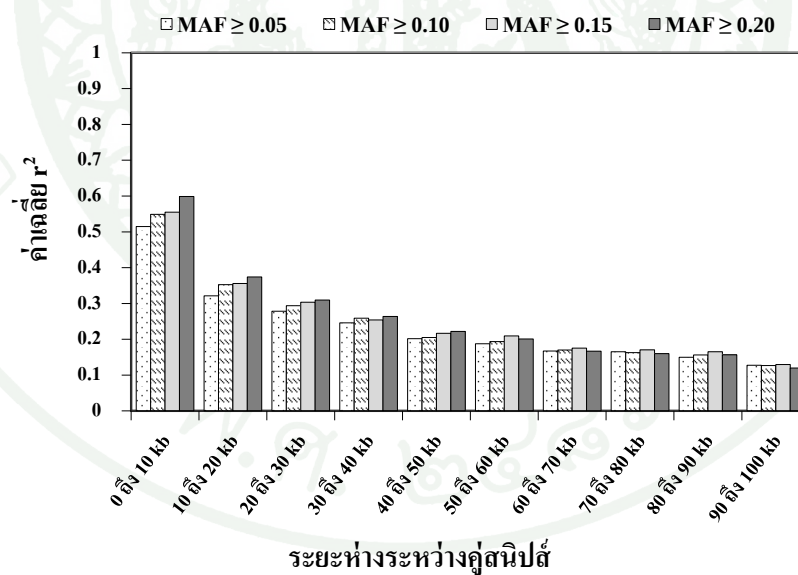
นอกจากนี้ Bohmanova *et al.* (2010) รายงานว่า การประมาณค่า D' ที่มีค่าสูงเมื่อความถี่อัลลีลของสปีส์ต่ำ เนื่องจากตัวหารในสมการประมาณค่า D' มีค่าเท่ากับค่าน้อยสุดของความถี่อัลลีลของสปีส์ ด้วยเหตุนี้ กรณีที่สปีส์ที่ใช้ศึกษามีความถี่ของอัลลีลต่ำ ทำให้ค่า D' ที่อยู่ในสมการประมาณค่าถูกหารด้วยตัวเลขจำนวนน้อยส่งผลให้ค่า D' ที่ประมาณได้มีค่าสูง ดังนั้น ในปัจจุบันการศึกษา LD จึงใช้การประมาณ D' และ r^2 รวมกัน (Khatkar *et al.*, 2008; Bohmanova *et al.*, 2010; Espigolan *et al.*, 2013) หรือการประมาณจากค่า r^2 เพียงค่าเดียว (Sargolzaei *et al.*, 2008; Qanbari *et al.*, 2010)



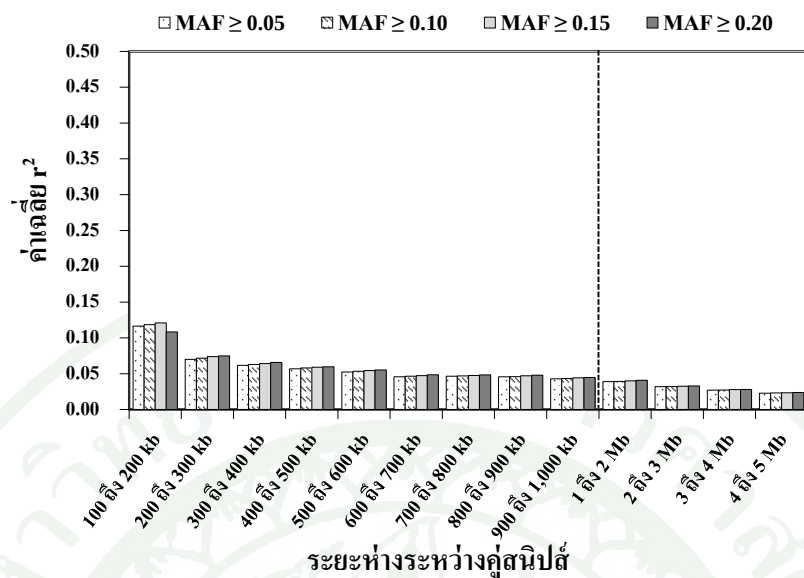
ภาพที่ 7 ค่าเฉลี่ย D' ในกลุ่มสปีส์ที่มีระดับ MAF แตกต่างกัน บนช่วงห่างระหว่างคู่สปีส์ไม่เกิน 100 kb



ภาพที่ 8 ค่าเฉลี่ย D' ในกลุ่มสนิปส์ที่มีระดับ MAF แตกต่างกัน บนช่วงห่างระหว่างคู่สนิปส์ไม่เกิน 1,000 kb และ 5 Mb



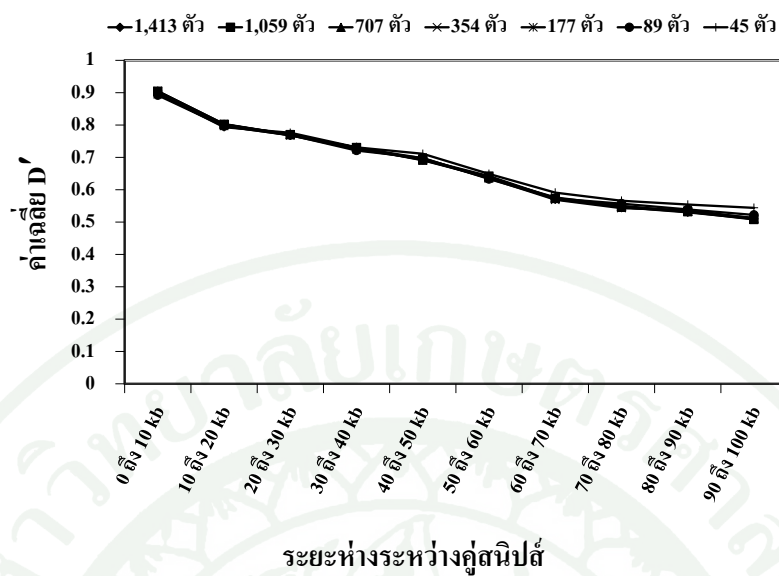
ภาพที่ 9 ค่าเฉลี่ย r^2 ในกลุ่มสนิปส์ที่มีระดับ MAF แตกต่างกัน บนช่วงห่างระหว่างคู่สนิปส์ไม่เกิน 100 kb



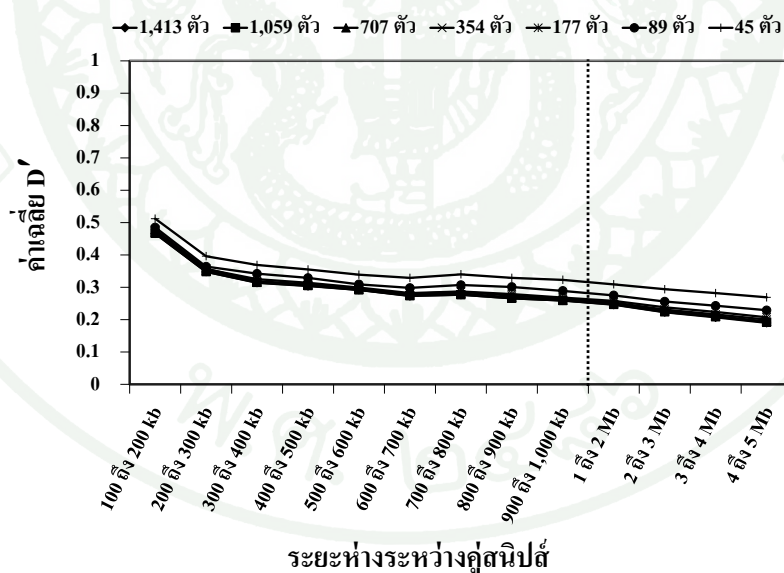
ภาพที่ 10 ค่าเฉลี่ย r^2 ในกลุ่มสนิปส์ที่มีระดับ MAF แตกต่างกัน บนช่วงห่างระหว่างคู่สนิปส์ไม่เกิน 1,000 kb และ 5 Mb

2.3 อิทธิพลของขนาดประชากรต่อระดับ LD ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

จากการประมาณค่า LD ด้วยข้อมูลสนิปส์จีโนไทป์ในประชากรที่มีขนาดแตกต่างกัน 7 กลุ่ม คือ กลุ่มที่ 1 ประกอบด้วยโคนมจำนวน 1,413 ตัว (ประชากรอ้างอิง) กลุ่มที่ 2 ประกอบด้วยโคนมจำนวน 1,059 ตัว กลุ่มที่ 3 ประกอบด้วยโคนมจำนวน 707 ตัว กลุ่มที่ 4 ประกอบด้วยโคนมจำนวน 354 ตัว กลุ่มที่ 5 ประกอบด้วยโคนมจำนวน 177 ตัว กลุ่มที่ 6 ประกอบด้วยโคนมจำนวน 89 ตัว และกลุ่มที่ 7 ประกอบด้วยโคนมจำนวน 45 ตัว ผลจากการศึกษาครั้งนี้ พบว่า ขนาดประชากรมีผลต่อระดับ LD ที่แตกต่างกัน โดยเฉพาะค่าเฉลี่ย D' ในกลุ่มที่ 7 (จำนวน 45 ตัว) มีค่าเฉลี่ยมากกว่าทุกกลุ่มสำหรับระยะห่างระหว่างคู่สนิปส์มากกว่า 90 kb ซึ่งแตกต่างกับค่าเฉลี่ย r^2 ที่ได้รับอิทธิพลจากขนาดประชากรที่ใช้ในการศึกษาน้อยกว่า (ภาพที่ 11, 12, 13 และ 14)



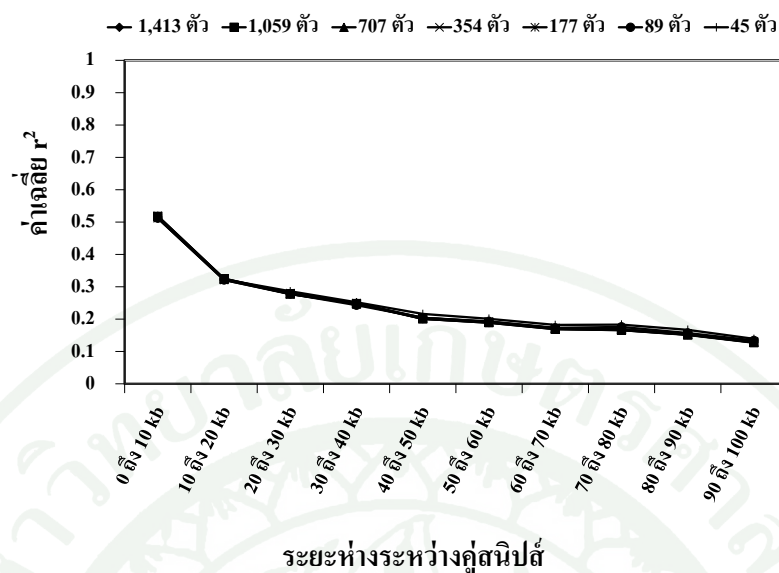
ภาพที่ 11 ความแตกต่างของระดับ D' ในกลุ่มประชากรที่มีขนาด แตกต่างกัน บนช่วงห่างระหว่าง คู่สนิปส์ไม่เกิน 100 kb



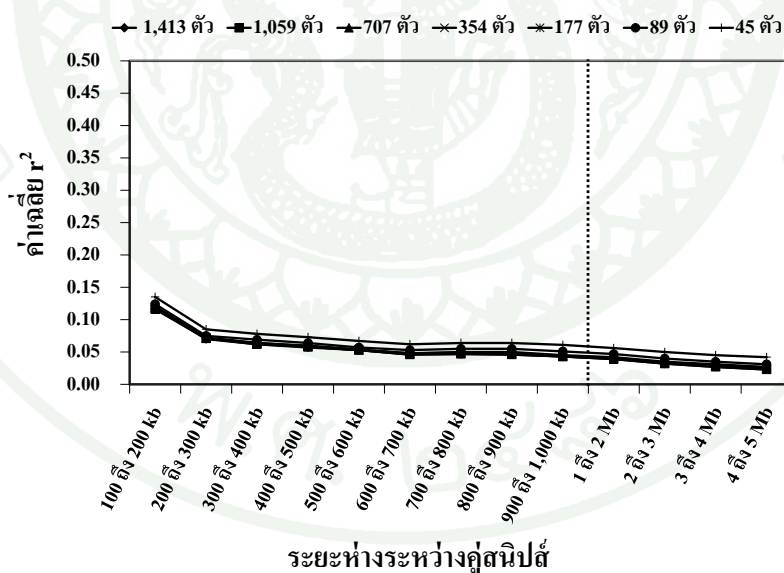
ภาพที่ 12 ความแตกต่างของระดับ D' ในกลุ่มประชากรที่มีขนาด แตกต่างกัน บนช่วงห่างระหว่าง คู่สนิปส์ไม่เกิน 1,000 kb และ 5 Mb

เมื่อพิจารณาความสัมพันธ์ระหว่างค่า D' และ r^2 ที่ประมาณค่าได้ในแต่ละกลุ่มกับประชากรอ้างอิงที่ระดับความสัมพันธ์ของสองตัวแปรมากกว่า 0.9 (ภาพที่ 15 และ 16) พบว่า การประมาณค่า LD จากค่า D' จำเป็นต้องใช้ขนาดประชากรน้อยสุดจำนวน 177 ตัว ในการประมาณค่า D' ที่มีค่าเบี่ยงเบนจากประชากรอ้างอิงในระดับที่ต่ำ โดยจำนวนประชากรนี้แตกต่างกับการประมาณค่า r^2 ด้วยจำนวนประชากรน้อยสุดที่มีค่าเบี่ยงเบนจากประชากรอ้างอิงในระดับที่ต่ำ คือ 89 ตัว ผลจากการศึกษาในประชากรโคนมหลากหลายพันธุ์ในประเทศไทยนี้ สอดคล้องกับผลจากการศึกษาของ Teare *et al.* (2002) ที่รายงานว่า การประมาณค่า D' ได้รับอิทธิพลมาจากขนาดประชากรมากกว่า r^2 อย่างไรก็ตาม ขนาดประชากรน้อยสุดที่ใช้ประมาณค่า LD ในประชากรโคนมพันธุ์โฮลสไตน์ในประเทศสหรัฐอเมริกา คือ 444 ตัว (D') และ 55 ตัว (r^2 ; Bohmanova *et al.*, 2010) สอดคล้องกับประชากรโคนมพันธุ์โฮลสไตน์ในประเทศออสเตรเลียที่รายงานโดย Khatkar *et al.* (2008) ซึ่งพบว่า ขนาดประชากรน้อยสุดในการประมาณค่า LD คือ 400 ตัว (D') และ 75 ตัว (r^2)

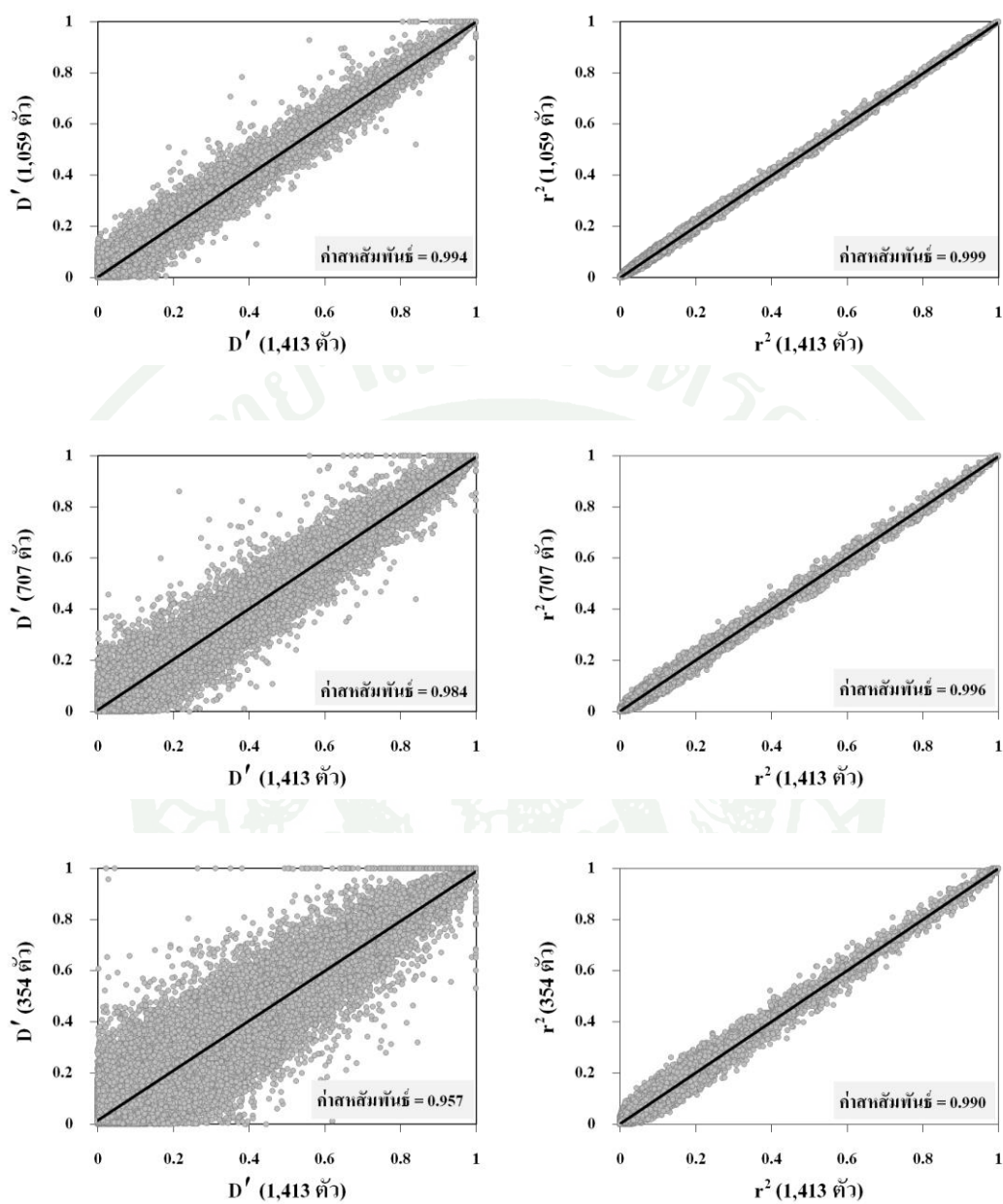
ผลจากการศึกษาครั้งนี้ ชี้ให้เห็นว่าประชากรโคนมหลากหลายพันธุ์ในประเทศไทย จำเป็นต้องใช้ประชากรโคนมที่มีขนาดใหญ่กว่าการรายงานของ Khatkar *et al.* (2008) และ Bohmanova *et al.* (2010) สำหรับการประมาณค่า r^2 อาจเป็นผลเนื่องมาจากความหลากหลายทางพันธุกรรมของโคนมในประเทศไทย ที่ประกอบด้วยพันธุกรรมของโคตระกูลยุโรปและอินเดีย ภายในโคนมบางตัว (พจน์ และคณะ, 2557) ด้วยเหตุนี้ จึงอาจส่งผลทำให้การประมาณค่า r^2 จำเป็นต้องใช้ประชากรขนาดใหญ่ขึ้น เนื่องจากความแตกต่างทางพันธุกรรมระหว่างโคพันธุ์แท้และโคนมลูกผสมมีผลต่อระดับ LD ที่แตกต่างกัน (Lu *et al.*, 2012) รวมทั้งความแตกต่างของโคนมตระกูลยุโรปและอินเดีย มีผลทำให้ระดับ LD แตกต่างกัน (R Porto-Neto *et al.*, 2014)



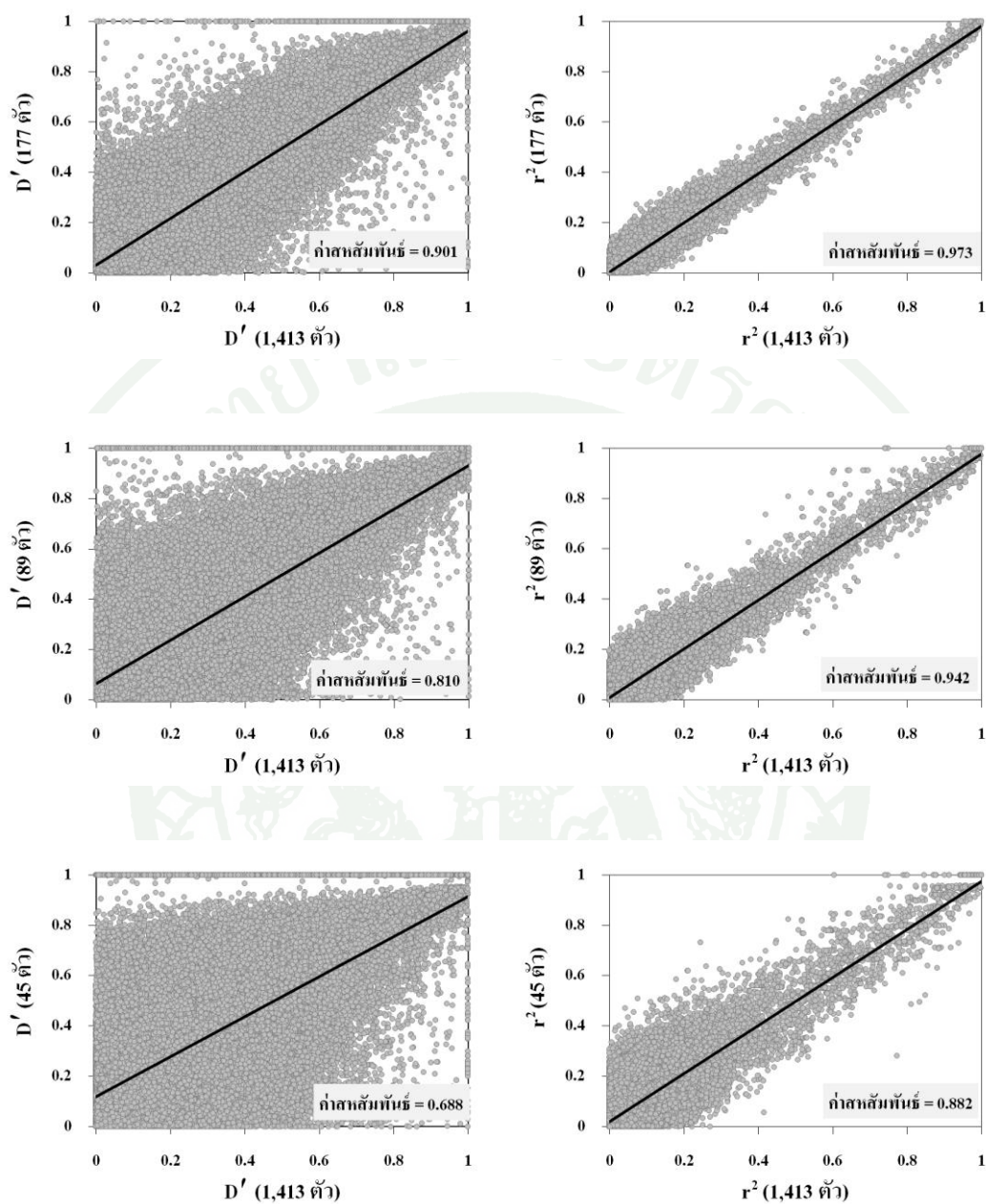
ภาพที่ 13 ความแตกต่างของระดับ r^2 ในกลุ่มประชากรที่มีขนาด แตกต่างกัน บนช่วงห่างระหว่างคู่สนิปส์ไม่เกิน 100 kb



ภาพที่ 14 ความแตกต่างของระดับ r^2 ในกลุ่มประชากรที่มีขนาด แตกต่างกัน บนช่วงห่างระหว่างคู่สนิปส์ไม่เกิน 1,000 kb และ 5 Mb



ภาพที่ 15 ความสัมพันธ์ระหว่างค่า LD ในกลุ่มประชากรโคนมทั้งหมด (1,413 ตัว) และประชากรกลุ่มทดสอบ (1,059, 707 และ 354 ตัว)



ภาพที่ 16 ความสัมพันธ์ระหว่างค่า LD ในกลุ่มประชากรโคนมทั้งหมด (1,413 ตัว) และประชากรกลุ่มทดสอบ (177, 89 และ 45 ตัว)

สรุปและข้อเสนอแนะ

สรุป

จากการศึกษาความผันแปรทางพันธุกรรมในระดับจีโนมของโคนม โดยใช้ข้อมูลสปีปส์จำนวน 8,573 ตำแหน่ง ที่กระจายตัวอยู่บนโครโมโซมร่างกายและโครโมโซมเพศของโคนมในประชากรโคนมหลากหลายพันธุ์ในประเทศไทยจำนวน 1,413 ตัว ซึ่งมีลักษณะการกระจายตัวของสปีปส์ที่ไม่สม่ำเสมอ และมีลักษณะอยู่รวมกันเป็นกลุ่มบนพื้นที่บางส่วนของโครโมโซมแต่ละคู่ โดยโครโมโซมคู่ที่ 1 มีตำแหน่งสปีปส์ที่กระจายบนพื้นที่โครโมโซม (158.162 Mb) และจำนวนสปีปส์มากที่สุด (532 ตำแหน่ง) และโครโมโซมคู่ที่ 20 มีความหนาแน่นสปีปส์มากที่สุด (3.969 ตำแหน่งต่อความยาวโครโมโซมที่เพิ่มขึ้น 1 Mb) ส่วนการกระจายตัวของค่า MAF บนโครโมโซมร่างกายและโครโมโซมเพศของโคนม มีสัดส่วนจำนวนสปีปส์เพิ่มขึ้นเมื่อค่า MAF เพิ่มขึ้น โดยสัดส่วนสปีปส์มากกว่าร้อยละ 75 มีค่า MAF มากกว่า 0.2 และสปีปส์ทั้งหมดบนจีโนมโคนมมีค่าเฉลี่ย MAF เท่ากับ 0.376

นอกจากนี้ สปีปส์จำนวน 8,207 ตำแหน่ง ที่กระจายตัวอยู่บนโครโมโซมร่างกายของโคนมแต่ละคู่ มีระดับ LD แตกต่างกัน โดยโครโมโซมคู่ที่ 20 มีค่าเฉลี่ย D' มากที่สุด (0.368) และโครโมโซมคู่ 16 มีค่าเฉลี่ย r^2 มากที่สุด (0.090) และระดับ LD บนโครโมโซมร่างกายทั้งหมดมีค่าเฉลี่ยเท่ากับ 0.263 (D') และ 0.049 (r^2) โดยระดับ LD มีค่าลดลงเมื่อระยะห่างระหว่างคู่สปีปส์เพิ่มขึ้น ($P < 0.05$) และการประมาณ LD จากค่า D' ได้รับอิทธิพลจากค่า MAF และขนาดประชากรที่ใช้ในการศึกษามากกว่า r^2 โดยขนาดประชากรโคนมที่เหมาะสมในประเทศไทยสำหรับการศึกษา LD ควรมากกว่า 177 ตัว ในการประมาณค่า D' และมากกว่า 89 ตัว ในการประมาณค่า r^2

ข้อเสนอแนะ

แนวทางการนำไปใช้ประโยชน์

ผลที่ได้จากการศึกษาวิจัยในครั้งนี้ สามารถช่วยให้ผู้ที่เกี่ยวข้องกับโคนมทราบถึงข้อมูล ความแตกต่างทางพันธุกรรมที่แท้จริงในระดับจีโนมของโคนมในประชากรโคนมหลากหลายพันธุ์ ที่ถูกเลี้ยงดูภายใต้สภาพแวดล้อมฟาร์มเกษตรกรในประเทศไทย ซึ่งข้อมูลเหล่านี้เป็นพื้นฐานสำคัญ ในการเรียนรู้และทำความเข้าใจ ในรูปแบบความผันแปรของสปีส์บนโครโมโซมโคนมแต่ละคู่ โดยสามารถนำไปใช้ประโยชน์ในการวางแผนการดำเนินงานและการบริหารจัดการ เพื่อใช้ ประโยชน์จากข้อมูลทางพันธุกรรม (สปีส์) ในการเพิ่มประสิทธิภาพทางพันธุกรรมสำหรับ ลักษณะการให้ผลผลิตของโคนมในประเทศไทย นอกจากนี้ยังแสดงให้เห็นถึงความสัมพันธ์ ระหว่างสปีส์บนโครโมโซมแต่ละคู่ ซึ่งมีส่วนช่วยในการพิจารณาการคัดเลือกสปีส์เพื่อใช้ ประโยชน์ในเพิ่มความสามารถในการตรวจสอบความสัมพันธ์ระหว่างสปีส์และลักษณะการ ให้ผลผลิต สำหรับการศึกษาความสัมพันธ์กับลักษณะการให้ผลผลิต และการประเมินความสามารถ ทางพันธุกรรมของโคนมในประเทศไทยได้อย่างเหมาะสม

แนวทางการศึกษาต่อ

ข้อมูลที่น่าสนใจในการศึกษาครั้งนี้ เป็นข้อมูลความแตกต่างทางพันธุกรรมสปีส์ ในระดับ จีโนมของโคนมในประชากรโคนมหลากหลายพันธุ์ในประเทศไทยจำนวน 8,573 ตำแหน่ง จาก โคนมทั้งหมด 1,413 ตัว ซึ่งเป็นจุดเริ่มต้นของการศึกษาด้านความแตกต่างทางพันธุกรรมสปีส์ใน ระดับจีโนมของโคนมในประเทศไทย ดังนั้น ข้อมูลความผันแปรของสปีส์บนโครโมโซมแต่ละคู่ และทั้งจีโนมของโคนม ควรได้รับการศึกษาต่อยอดในด้านความสัมพันธ์ระหว่างความผันแปรทาง พันธุกรรมกับลักษณะการให้ผลผลิตน้ำนม และการใช้ประโยชน์จากความผันแปรทางพันธุกรรม ในการประเมินค่าทางพันธุกรรมของโคนม เป็นต้น ซึ่งอาจช่วยให้สามารถเพิ่มความแม่นยำในการ พัฒนาระบบการคัดเลือกโคนมในประเทศไทย เนื่องจากเป็นข้อมูลความแตกต่างทางพันธุกรรมที่ แท้จริง

เอกสารและสิ่งอ้างอิง

กนกวรรณ จารุกัจฉา และ วรัญญา จตุพรประเสริฐ. 2550. สนิป: ความรู้พื้นฐานสู่การประยุกต์ใช้, **Thai Pharm Health Sci. J.** 2 (2): 166-174.

ทวี เหล่าดีม, ศกร คุณวุฒิฤทธิชน, ธนาทิพย์ สุวรรณโสภี และ Mauricio A. Elzo. 2557. รูปแบบความผันแปรของสลิปส์ในระดับจีโนมของโคนมไทยหลากหลายพันธุ์, น 17-24. ใน รายงานการประชุมทางวิชาการของมหาวิทยาลัยเกษตรศาสตร์ ครั้งที่ 52 (สาขาสัตว์). มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ.

พจน์ ฤทธิ์ไสว, ศกร คุณวุฒิฤทธิชน, ดนัย จัตุวา, ธนาทิพย์ สุวรรณโสภี และ Mauricio A. Elzo. 2557. สัดส่วนของพันธุ์โคและความมีอิทธิพลต่อการให้ผลผลิตน้ำนมของโคนมไทย, น 25-32. ใน รายงานการประชุมวิชาการของมหาวิทยาลัยเกษตรศาสตร์ ครั้งที่ 52 (สาขาสัตว์). มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ.

วรพรรณ สิทธิถาวร. 2549. เทคโนโลยีไมโครแอเรย์ของดีเอ็นเอ. **Thai Pharm Health Sci. J.** 1 (2): 139-145.

Ardlie, K.G., K.L. Lunetta and M. Seielstad. 2002. Patterns of linkage disequilibrium in the human genome. **Nat. Rev.** 3: 299-309.

Barrett, J.C., B. Fry, J. Maller and M.J. Daly. 2005. Haploview: analysis and visualization of LD and haplotype maps. **Bioinformatics** 21 (2): 263-265.

Bohmanova, J., M. Sargolzaei and F.S. Schenkel. 2010. Characteristics of linkage disequilibrium in North American Holsteins. **BMC Genomics** 11: 421.

Brooks, L.D. 2003. SNPs: Why Do We Care?, pp. 1-14. In P. Y. Kwok, **Single nucleotide polymorphisms methods and protocols.** USA.

Brown, T.A. 2010. **Gene Cloning & DNA Analysis**. 6th ed. Vivar Printing Sdn Bhd, Malaysia.

Bush, W.S. and J.H. Moore. 2012. Chapter 11: genome-wide association studies.

PLoS Comput. Biol. 8 (12): e1002822.

Chamberlain, A.J., H.C. McPartlan and M.E. Goddard. 2007. The number of loci that affect milk production traits in dairy cattle. **Genetics** 177: 1117-1123.

_____, B.J. Hayes, K. Savin, S. Bolormaa, H.C. McPartlan, P.J. Bowman, C. Van Der Jagt, S. MacEachern and M.E. Goddard. 2012. Validation of single nucleotide polymorphisms associated with milk production traits in dairy cattle. **J. Dairy Sci.** 95: 864–875.

Cole, J.B., G.R. Wiggans, L. Ma, T.S. Sonstegard, T.J. Lawlor Jr, B.A. Crooker, C.P. van Tassell, J. Yang, S. Wang, L.K. Matukumalli and Y. Da. 2011. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. **BMC Genomics** 12: 408.

Conrad, D.F., J.E.M. Keebler, M.A. Depristo, S.J. Lindsay, Y. Zhang, F. Cassals, Y. Idaghdour, C.L. Hartl, C. Torroja, K.V. Gerimella, M. Zilversmit, R. Cartwright, G. Rouleau, M. Daly, E.A. Stone, M.E. Hurles and P. Awadalla. 2012. Variation in genome-wide mutation rates within and between human families. **Nat. Genet.** 43 (7): 712–714.

de Roos, A.P.W., B.J. Hayes, R.J. Spelman and M.E. Goddard. 2008. Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. **Genetics** 179: 1503-1512.

- Dunning, A.M., F. Durocher, C.S. Healey, M.D. Teare, S.E. McBride, F. Carlomagno, C.F. Xu, E. Dawson, S. Rhodes, S. Ueda, E. Lai, R.N. Luben, E.J. Van Rensburg, A. Mannermaa, V. Kataja, G. Rennart, I. Dunham, I. Purvis, D. Easton and B.A.J. Ponder. 2000. The extent of linkage disequilibrium in four populations with distinct demographic histories. **Am. J. Hum. Genet.** 67: 1544-1554.
- Dziuda, D.M. 2010. **Data Mining Genomics and Proteomics: Analysis of Gene and Protein Expression Data.** Prentice Publishers, USA.
- Espigolan, A., F. Baldi, A.A Boligon, F.R.P. Souza, D.G.M. Gordo, R.L. Tonussi, D.F. Cardoso, H.N. Oliverira, H. Tonhati, M. Sargolzaei, F.S. Schenkel, R. Carneiro, J.A. Ferro and L.G. Albuquerque. 2013. Study of whole genome linkage disequilibrium in Nelore cattle. **BMC Genomics** 14: 305.
- Fadiel, A. and F. Naftolin. 2003. Microarray applications and challenges: a vast array of possibilities. **Int. Arch. Biosci.** 1111-1121.
- Fan, B., S.K. Onteru, Z.Q. Du, D.J. Garrick, K.J. Stalder and M.F. Rothschild. 2011. Composition and structure soundness traits in pig. **PLoS ONE** 6 (2): e14726.
- Farre, M., D. Michenletti and A. Ruiz-Herrera. 2013. Recombination rate and genomic shuffling in Human and Chimpanzee-a new twist in the chromosome speciation theory. **Mol. Biol. Evol.** 30 (4): 853-864.
- Frazer, K.A., S.S. Murray, N.J. Schork and E.J. Topol. 2009. Human genetic variation and its contribution to complex traits. **Nat. Rev. Genet.** 10: 241-251.
- Goddard, M.E. and B.J. Hayes. 2009. Mapping gene for complex traits in domestic animals and their use in breeding program. **Nat. Rev. Genet.** 10: 381-391.

Guo, J., H. Jorjani and O. Carlborg. 2012. A genome-wide association study using international breeding evaluation data identifies major loci affecting production traits and stature in the Brown Swiss cattle breed. **BMC Genet.** 13: 82.

Hayes, B. 2011. **Ben Hayes course notes.** Available Source:

http://snp.toulouse.inra.fr/~alegarra/ben_hayes_course/toulouse_course_notes.pdf. Oct 1, 2013.

Hayes, B.J., P.J. Bowman, A.J. Chamberlain and M.E. Goddard. 2009a. Invited review: genomic selection in dairy cattle: progress and challenges. **J. Dairy Sci.** 92: 433-443.

_____, _____, A.C. Chamberlain, K. Verbyla and M.E. Goddard. 2009b. Accuracy of genomic breeding values in multi-breed dairy cattle populations. **Genet. Sel. Evol.** 41: 51.

Hayes, D.W., J.I. Weller, M. Ron, M. Band, J.E. Beever, E. Feldmesser, Y. Da, G.R. Wiggins, P.M. VanRaden and H.A. Lewin. 1999. A genome scan for QTL influencing milk production and health traits in dairy cattle. **Physiol. Genomics** 1: 165-175.

Hill, W.G. and A. Robertson. 1968. Linkage disequilibrium in finite populations. **Theor. Appl. Genet.** 38 (6): 226-231.

Illumina Inc. 2013. **Revolutionizing genetics in the cattle industry.** Available Source:

<http://www.illumina.com/Documents/products/brochures/BovineProductBrochure.pdf>. May 1, 2013.

Khatkar, M.S., F.W. Nicholas, A.R. Collins, K.R. Zenger, J.A.L. Cavanagh, W. Barris, R.D.

Schnabel, J.F. Taylor and H.W. Raadsma. 2008. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel **BMC Genomics** 9: 187.

- Kolbehdari, D., Z. Wang, J.R. Grant, B. Murdoch, A. Prasad, Z. Xiu, E. Marques, P. Stothard and S.S. Moore. 2009. A whole genome scan to map QTL for milk production traits and somatic cell score in Canadian Holstein bulls. **J. Anim. Breed. Genet.** 126: 216–227.
- Koonawootrittriron, K., M.A. Elzo and T. Thongprapi. 2009. Genetic trends in a Holstein × other breeds multibreed dairy population in Central Thailand. **Livest. Sci.** 122: 186-192.
- Lewontin, R.C. 1964. The interaction of selection and linkage. I. general considerations; heterotic models. **Genetics** 49: 49-67.
- Lu, D., M. Sargolzaei, M. Kelly, C. Li, G.V. Voort, Z. Wang, G. Plastow, S. Moore and S.P. Miller. 2012. Linkage disequilibrium in Angus, Charolais and Crossbred beef cattle. **Front. Genet.** 3: 152.
- Meredith, B.K., F.J. Kearney, E.K. Finalay, D.G. Bradley, A.G. Fahey, D.P. Berry and D.J. Lynn. 2012. Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland. **BMC Genet.** 13: 21.
- Meuwissen, T.H.E., B.J. Hayes and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker mas. **Genetics** 157: 1819-1829.
- Microsoft Office. 2010. **Microsoft Excel 2010 Product Guide.** Microsoft Corp., USA.
- Miller, M.B. and Y.W. Tang. 2009. Basic concepts of microarrays and potential applications in clinical microbiology. **Clin. Microbiol. Rev.** 22 (4): 611-633.
- Mu, X.J., Z.J. Lu, Y. Kong, H.Y.K. Lam and M.B. Gerstein. 2011. Analysis of genomic variation in non-coding elements using population-scale sequencing data from the 1000 Genomes Project. **Nucleic Acids Res.** 39: 7058-7076.

- Ogawa, S., H. Matsuda, Y. Taniguchi, T. Watanabe, S. Nishmura, Y. Sugimoto and H. Iwaisaki. 2014. Effect of single nucleotide polymorphism marker density on degree of genetic variance explained and genomic evaluation for carcass traits in Japanese Black beef cattle. **BMC Genet.** 15: 15.
- Packer, M.J., M.P. Dauncey and C.A. Hunter. 2000. Sequence-dependent DNA structure: Tetranucleotide conformation maps. **J. Mol. Biol.** 295: 85-103.
- Pérez O'Brien., AM., G. Mészáros, Y.T. Utsunomiya, T.S. Sonstegard, J.F. Garcia, C.P. Van Tassell, R. Carvalheiro, M.V.B. da Silva and J. Solkner. 2014. Linkage disequilibrium levels in *Bos indicus* and *Bos taurus* cattle using medium and high density SNP chip data and different minor allele frequency distributions. **Livest. Sci.** 166: 121-132.
- Pritchard, J.K. and M. Przeworski. 2001. Review article Linkage disequilibrium in Human: models and data. **Am. J. Genet.** 69: 1-14.
- Pryce, J.E., S. Bolormaa, A.J. Chamberlain, P.J. Bowman, K. Savin, M.E. Goddard and B.J. Hayes. 2010. A validated genome-wide association study in 2 dairy cattle breeds for milk production and fertility traits using length haplotypes. **J. Dairy Sci.** 93: 3331-3345.
- Qanbari, S., E.C.G. Pimental, J. Tetens, G. Thaller, P. Lichtner, A.R. Sharifi and H. Simianer. 2010. The pattern of Linkage disequilibrium in German Holstein cattle. **Anim. Genet.** 41: 346-356.
- R Porto-Neto, L., J.W. Kias and A. Reverter. 2014. The extent of linkage disequilibrium in beff cattle breeds using high-density SNP genotypes. **Genet. Sel. Evol.** 46: 22.
- Sargolzaei, M., F.S. Schenkel, G.B. Jansen and L.R. Schaeffer. 2008. Extent of Linkage disequilibrium in Holstein cattle in North America. **J. Dairy Sci.** 91: 2106-2117.

SAS. 2003. **SAS OnlineDoc 9.1.3**. SAS Inc., Cary, NC, USA.

Schrooten, C. 2004. **Genomic Variation in Dairy Cattle-Identification and Use**. Ph.D. Thesis, Wageningen University, Netherlands.

Sinden, R.R. 1994. **DNA Structure and Function**. Academic Press, Inc, USA.

Tabangin, M.E., J.G. Woo and L.J. Martin. 2009. The effect of minor allele frequency on the likelihood of obtaining false positives. **BMC Genet.** 3: S41.

Teare, M.D., A.M. Dunning, F. Durocher, G. Rennart and D.F. Easton. 2002. Sampling distribution of summary linkage disequilibrium measures. **Ann. Hum. Genet.** 66: 223-233.

The Bovine HapMap Consortium. 2009. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. **Science** 324: 528.

Wall, J.D. and J.K. Pritchard. 2003. Haplotype block and linkage disequilibrium in the Human genome. **Net. Rev. Genet.** 4: 587-597.

Walters, R., C. Laurin and G.H. Lubke. 2012. An intergrated approach to reduce the impact of minor allele frequency and linkage disequilibrium on variable importance measures of genome-wide data. **Bioinformatics** 28 (20): 2615-2623.

Watson, J.D. and F.H.C. Crick. 1953. Molecular structure of deoxypentose nucleic acids. **Nature** 171: 738.

Weinstein, A. 1936. The theory of multiple strand crossing overs. **Genetics** 21: 155-199.

Wiener, P., D. Burton and J.R. Williams. 2004. Breed relationships and definition in British cattle: a genetic analysis. **Heredity** 93: 597-602.

Wiggans, G.R., T.A. Cooper, P.M. VanRaden, K.M. Olson and M.E. Tooker. 2012. Use of the Illumina Bovine3K Beadchip in dairy genomic evaluation. **J. Dairy Sci.** 95: 1552-1558.

Womack, J.E. 2005. Advances in livestock genomics: opening the barn door. **Genome Res.** 15: 1699-1705.

Yang, J., W. Zhu, J. Chen, Q. Zhang and S. Wu. 2014. Genome-wide two-marker linkage disequilibrium mapping of quantitative trait loci. **BMC Genet.** 15: 20.

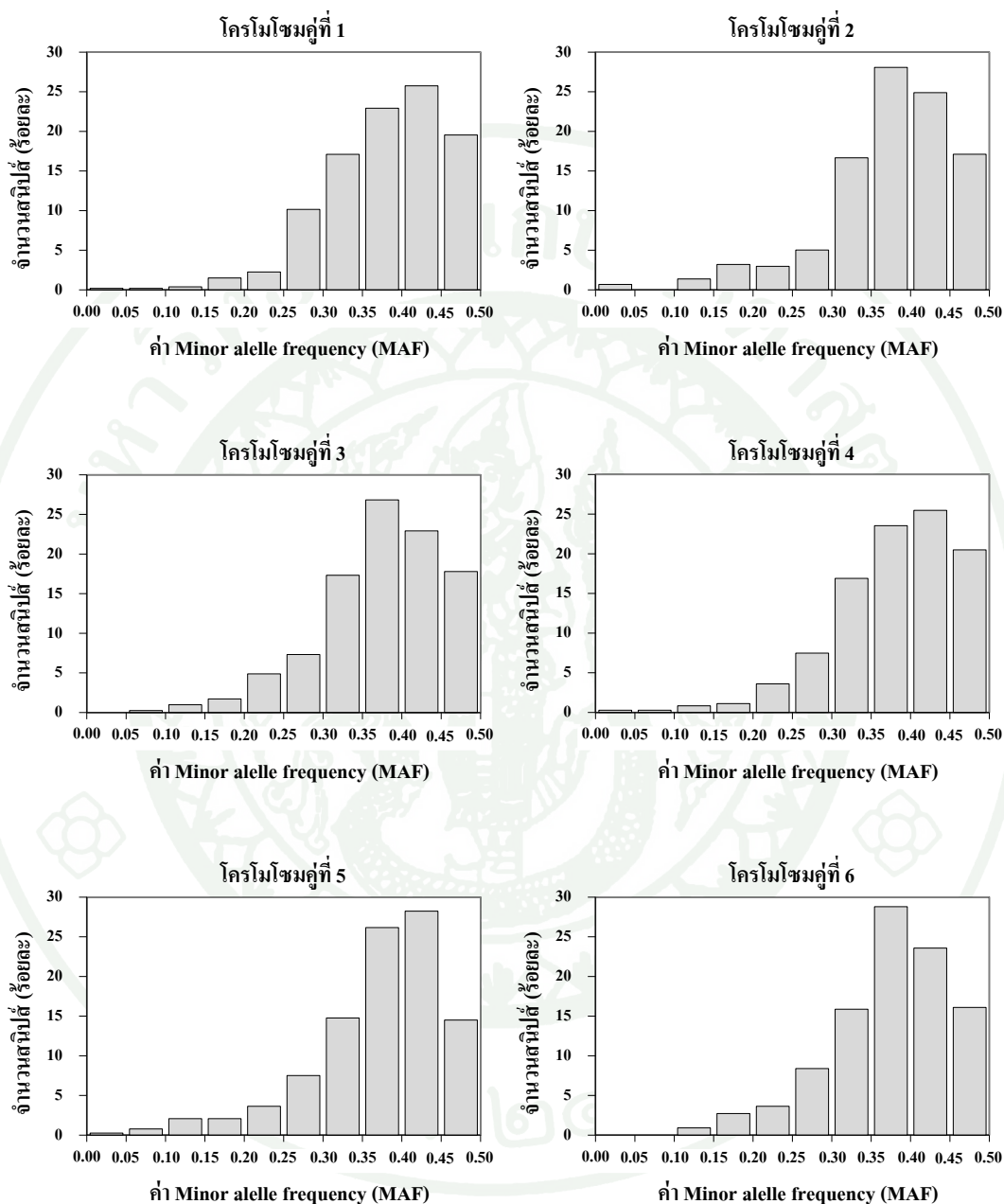
Zapata, C. 2013. Linkage disequilibrium measure for fine-scale mapping of disease loci are revisited. **Front. Genet.** 4: 228.

Zhou, L., X. Ding, Q. Zhang, Y. Wang, M.S. Lund and G. Su. 2013. Consistency of linkage disequilibrium between Chinese and Nordic Holsteins and genomic prediction for Chinese Holsteins using a joint reference population. **Genet. Sel. Evol.** 45: 7.

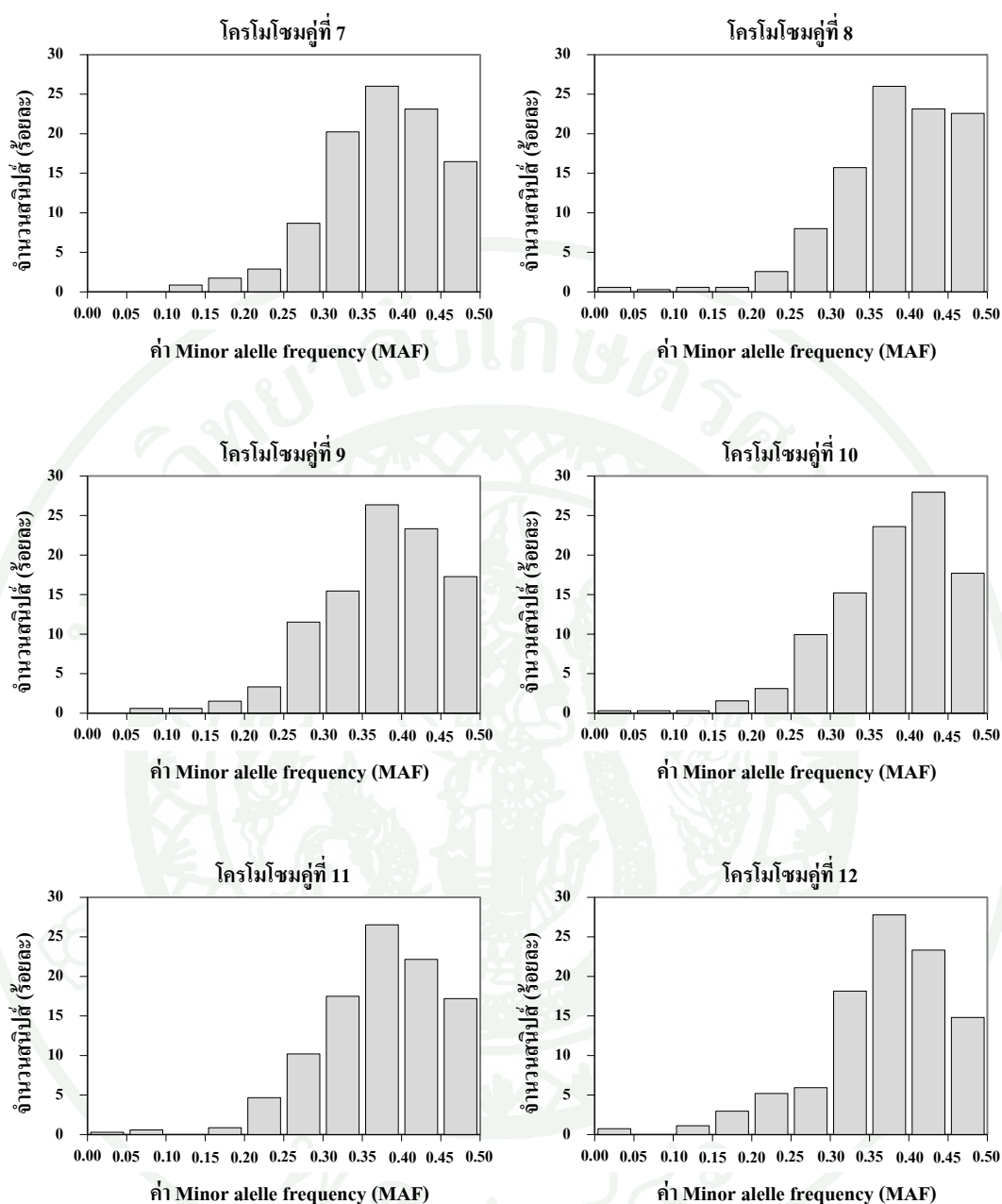


ภาคผนวก

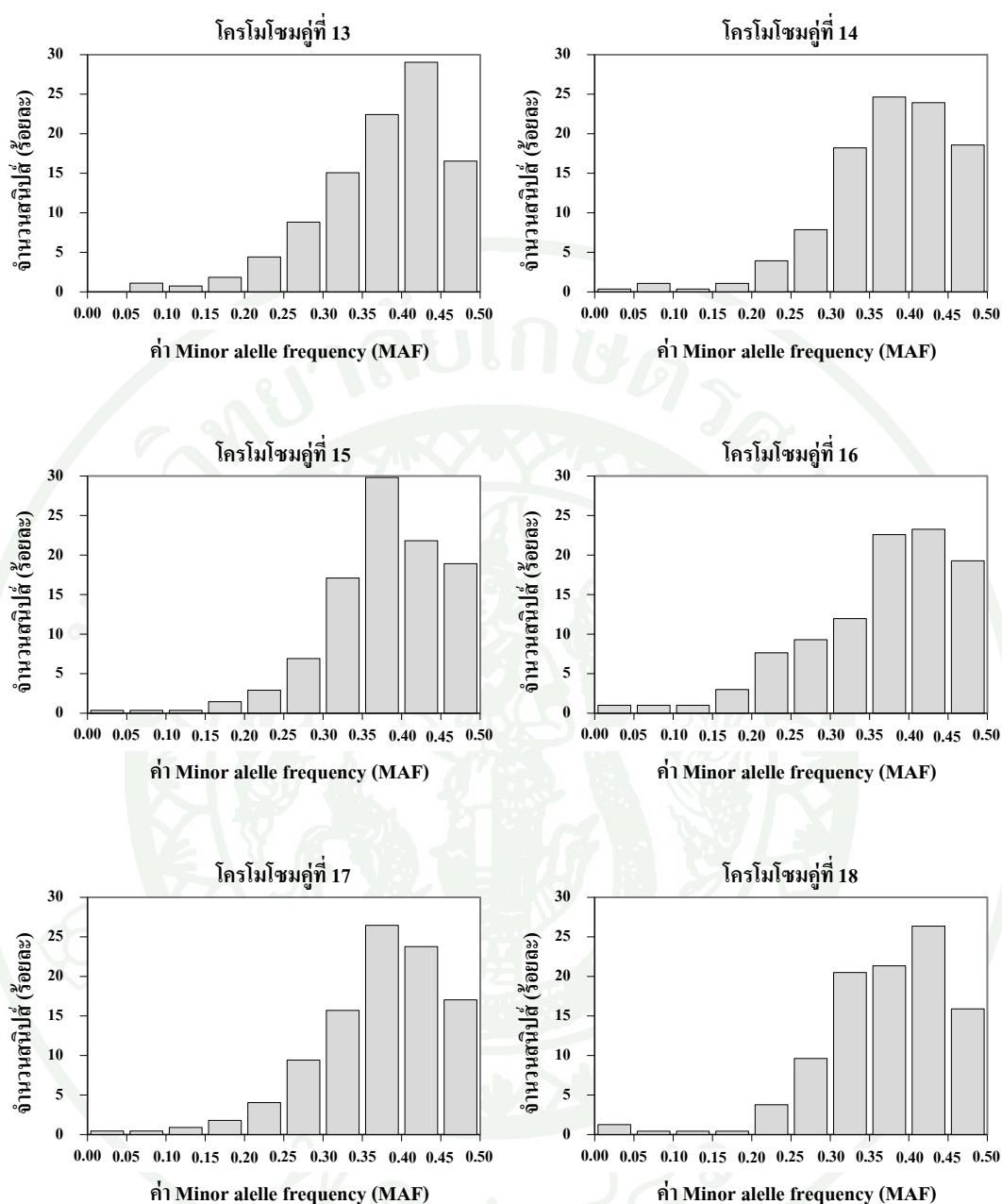
ภาพ ตารางผลการวิเคราะห์การกระจายตัวและความแปรปรวนของสปีส์



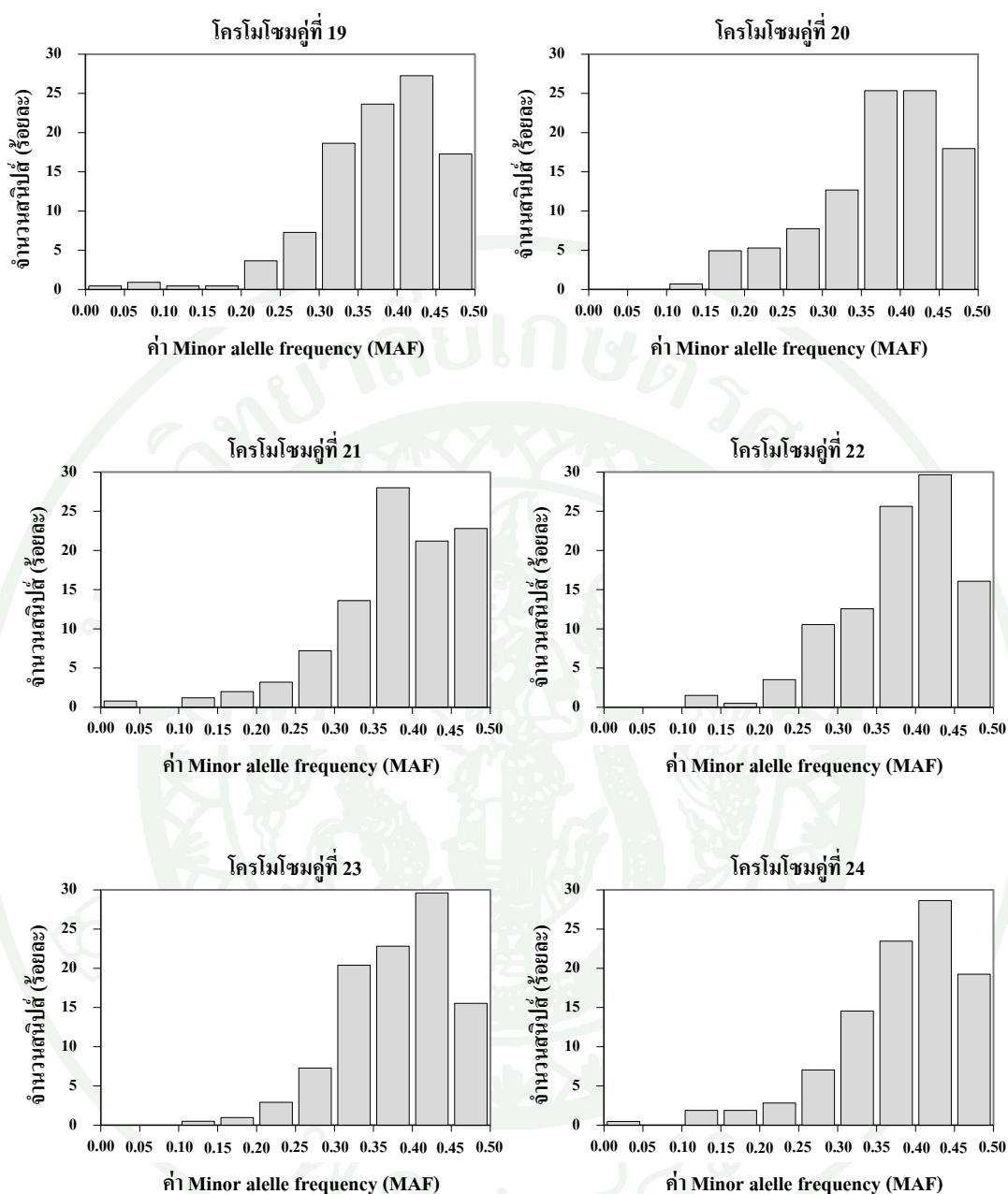
ภาพผนวกที่ 1 การกระจายตัวของสปีส์ MAF บนโครโมโซมคู่ที่ 1 ถึง 6 ในประชากรโคนม
หลากหลายพันธุ์ในประเทศไทย



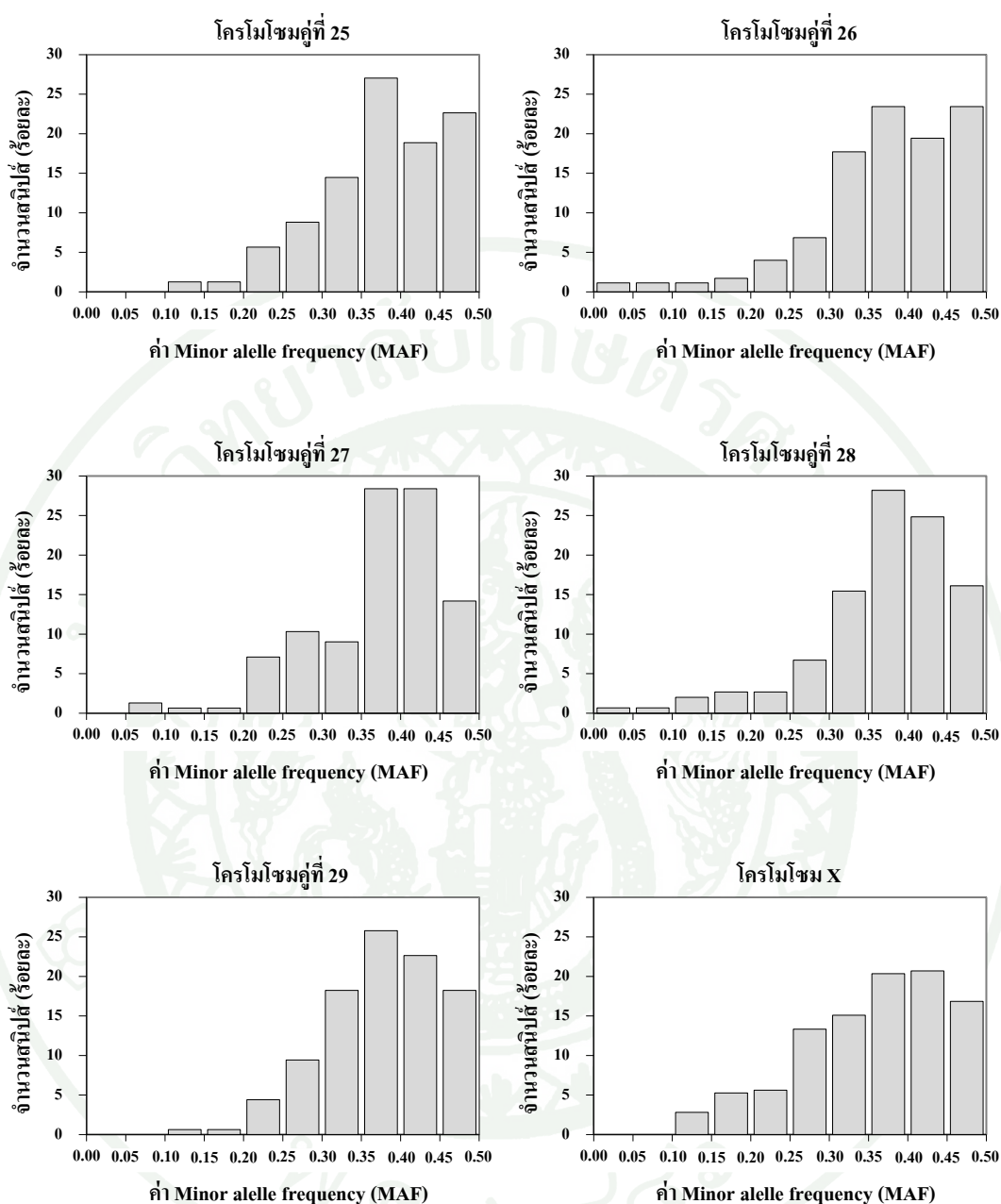
ภาพผนวกที่ 2 การกระจายตัวของสปีส์ MAF บนโครโมโซมคู่ที่ 7 ถึง 12 ในประชากรโคนม
หลากหลายพันธุ์ในประเทศไทย



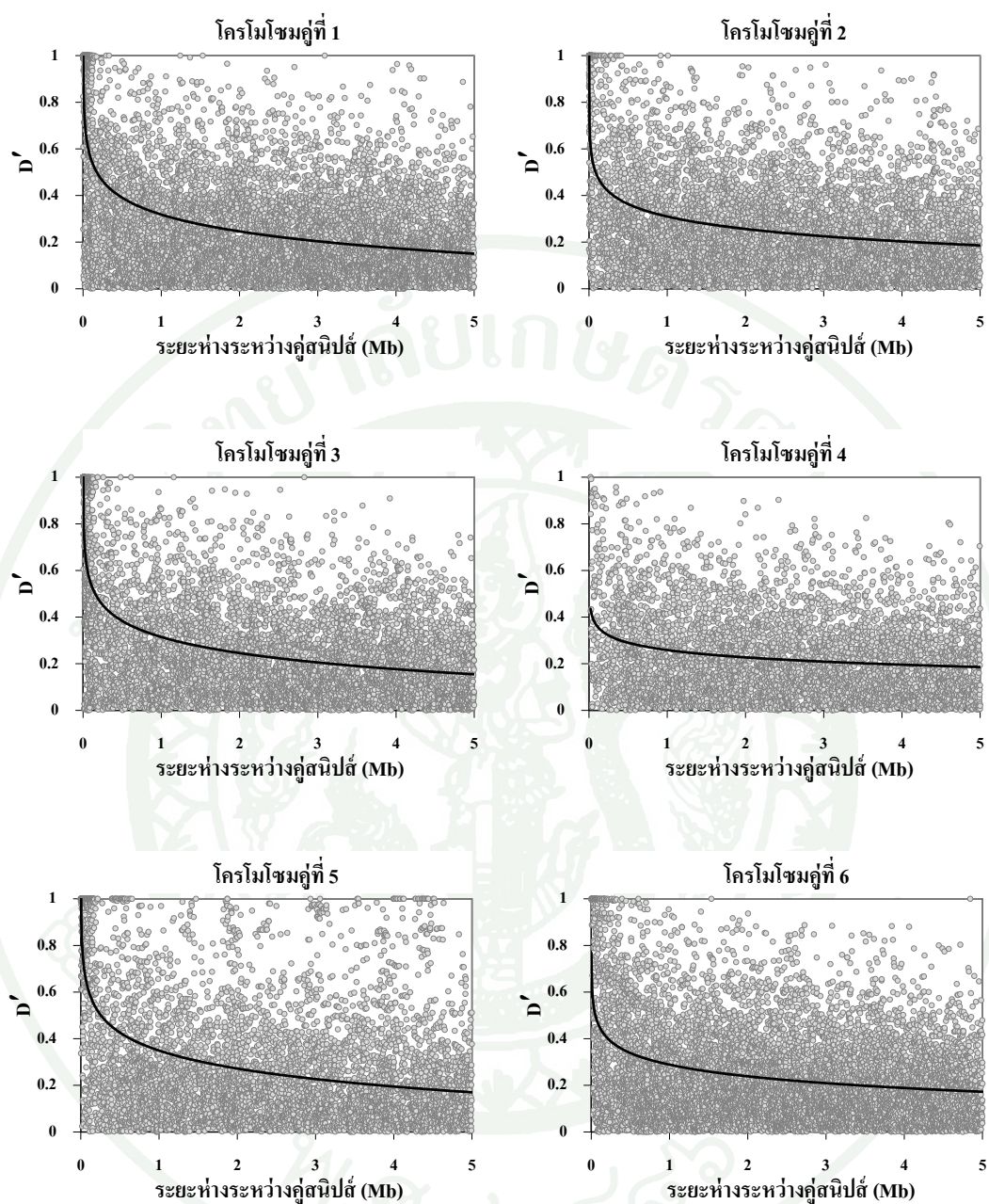
ภาพผนวกที่ 3 การกระจายตัวของสปีส์ MAF บนโครโมโซมคู่ที่ 13 ถึง 18 ในประชากรโคนม
หลากหลายพันธุ์ในประเทศไทย



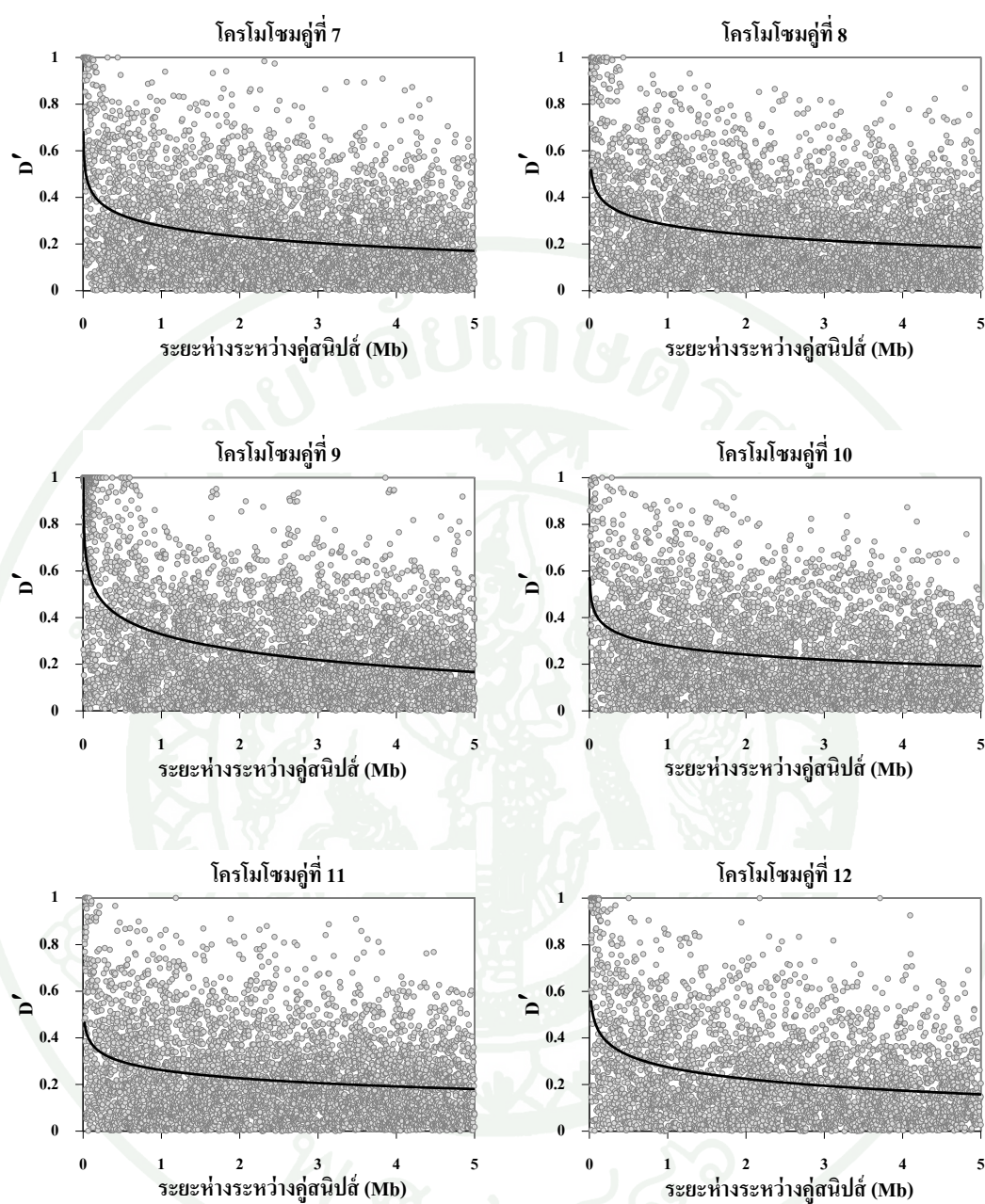
ภาพผนวกที่ 4 การกระจายตัวของสปีส์ MAF บนโครโมโซมคู่ที่ 19 ถึง 24 ในประชากรโคนม
หลากหลายพันธุ์ในประเทศไทย



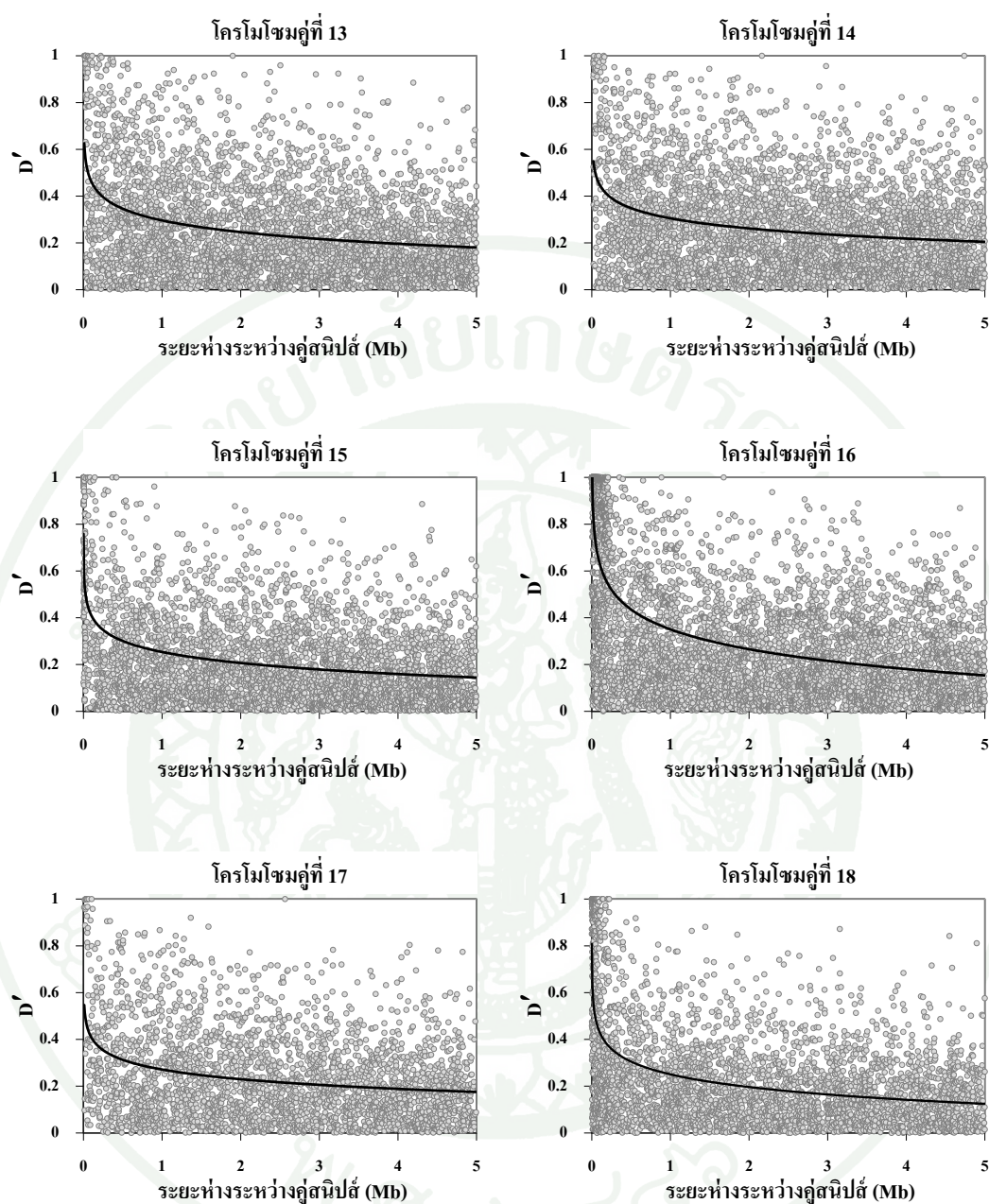
ภาพผนวกที่ 5 การกระจายตัวของสปีส์ MAF บนโครโมโซมคู่ที่ 25 ถึง X ในประชากรโคนม
หลากหลายพันธุ์ในประเทศไทย



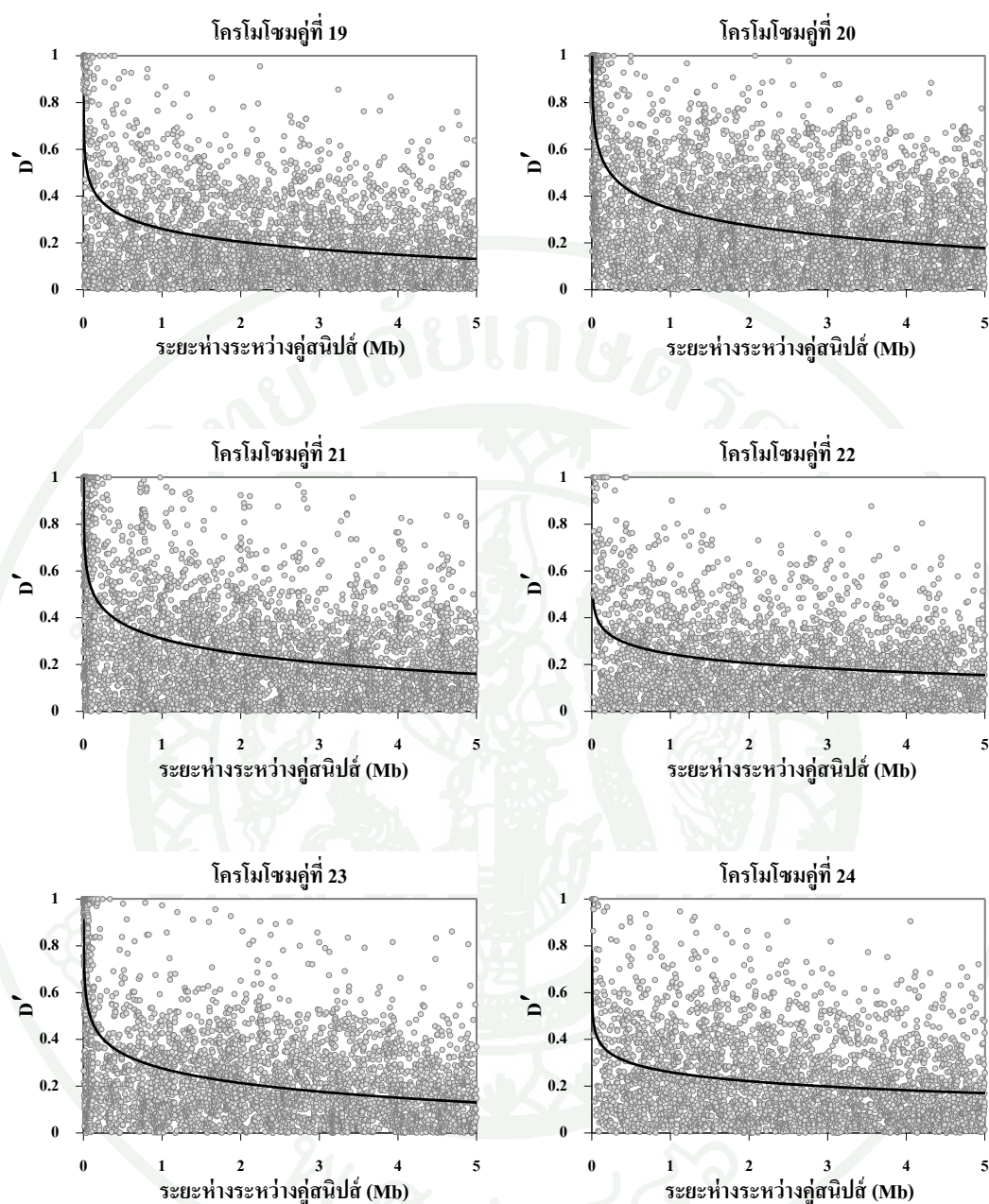
ภาพผนวกที่ 6 การกระจายตัวของ D' บนโครโมโซมร่างกายคู่ที่ 1 ถึง 6 ในประชากรโคนม
หลากหลายพันธุ์ในประเทศไทย



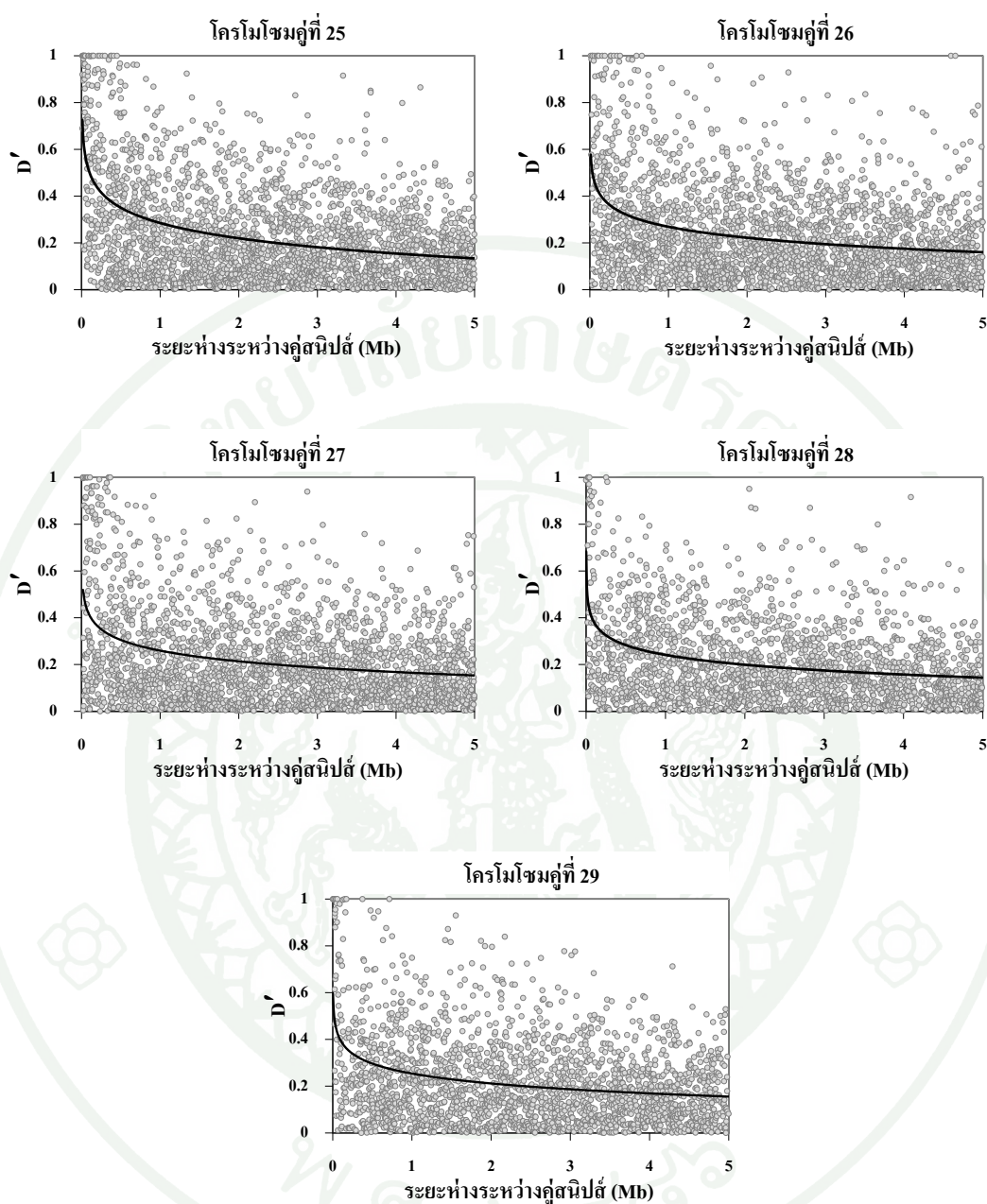
ภาพผนวกที่ 7 การกระจายตัวของ D' บนโครโมโซมร่างกายคู่ที่ 7 ถึง 12 ในประชากร โคนม
หลากหลายพันธุ์ในประเทศไทย



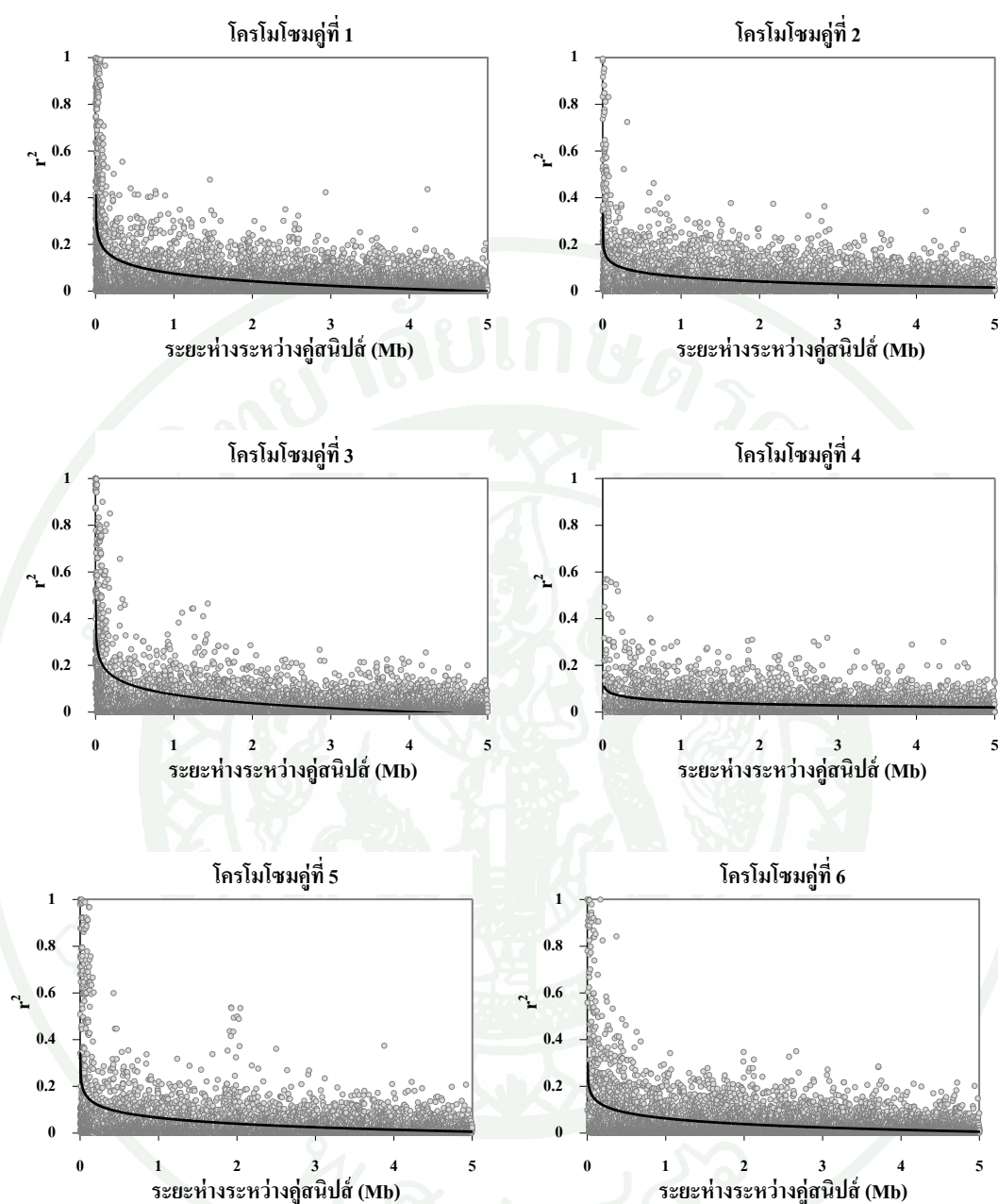
ภาพผนวกที่ 8 การกระจายตัวของ D' บนโครโมโซมร่างกายคู่ที่ 13 ถึง 18 ในประชากรโคนม
หลากหลายพันธุ์ในประเทศไทย



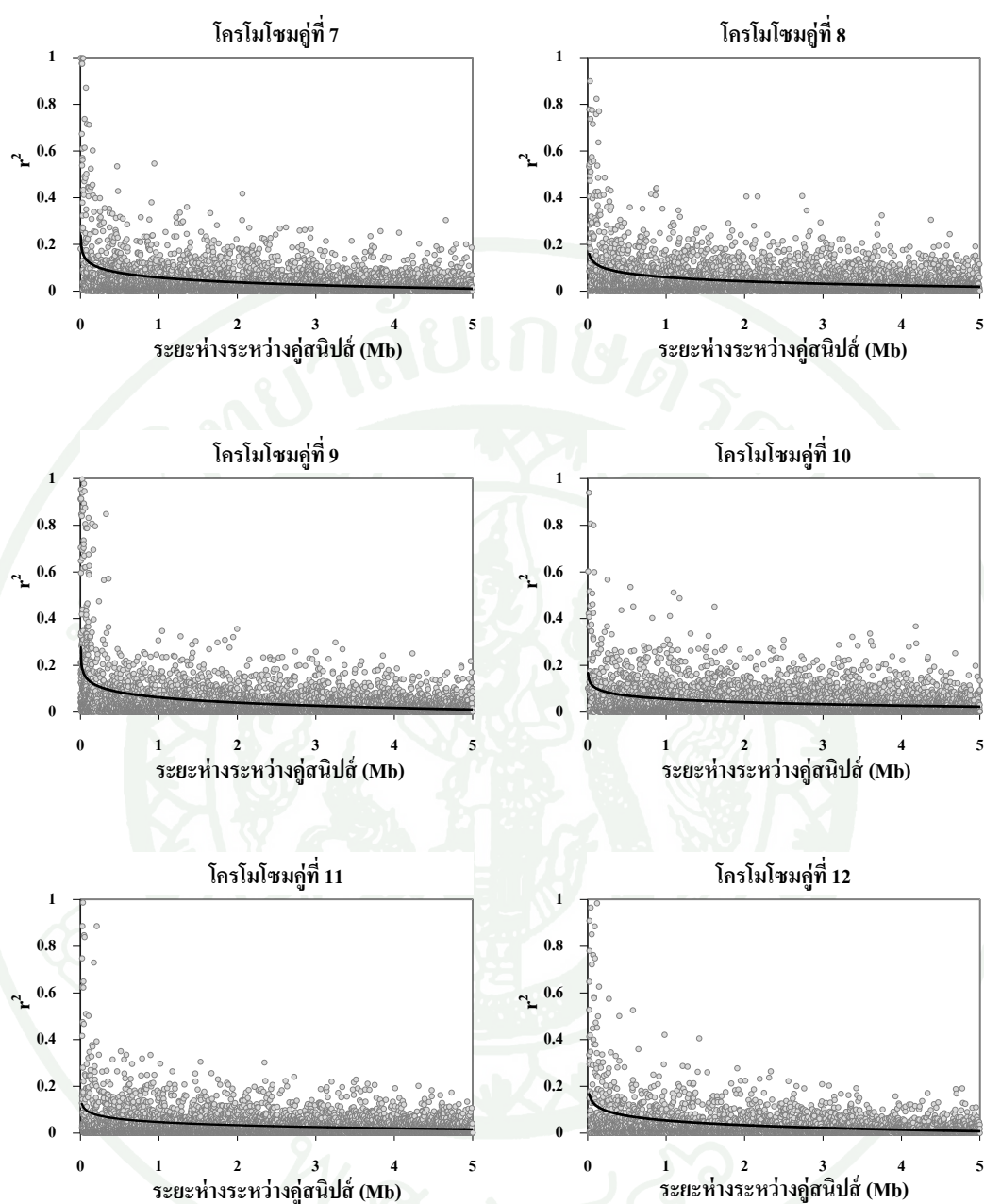
ภาพผนวกที่ 9 การกระจายตัวของ D' บนโครโมโซมร่างกายกลุ่มที่ 19 ถึง 24 ในประชากร โคนม
หลากหลายพันธุ์ในประเทศไทย



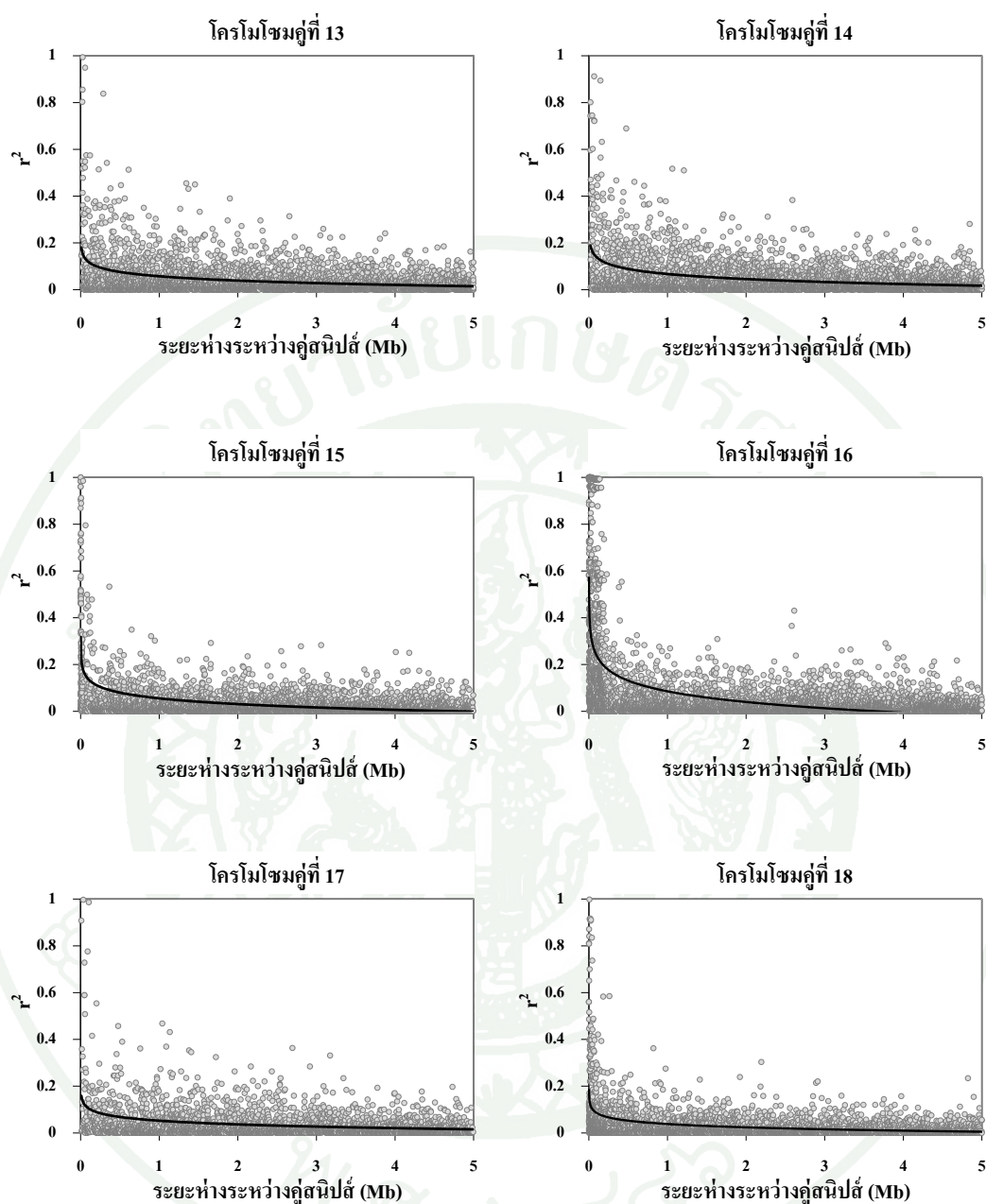
ภาพผนวกที่ 10 การกระจายตัวของ D' บนโครโมโซมร่างกายคู่ที่ 25 ถึง 29 ในประชากรโคนม
หลากหลายพันธุ์ในประเทศไทย



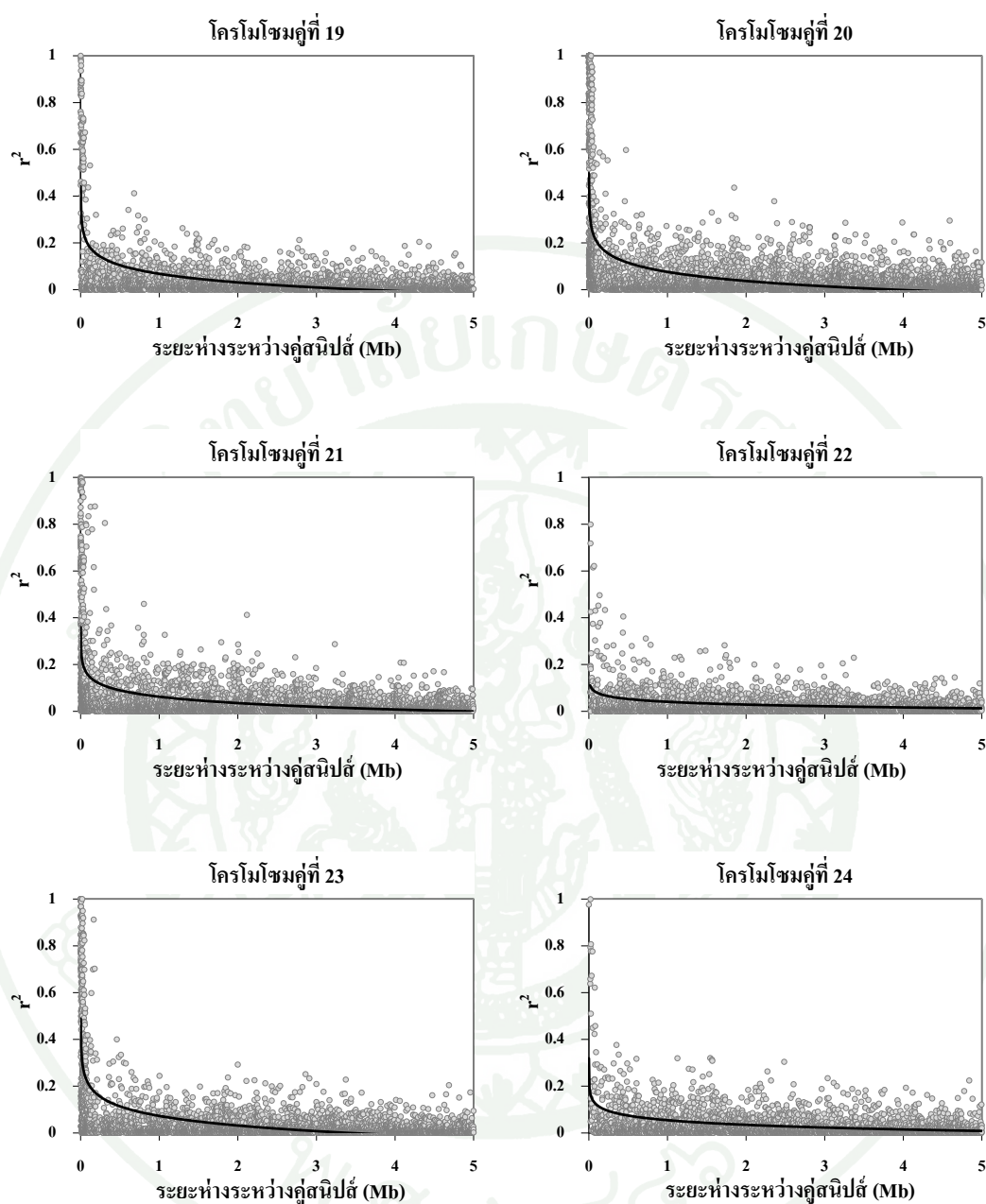
ภาพผนวกที่ 11 การกระจายตัวของ r^2 บนโครโมโซมร่างกายคู่ที่ 1 ถึง 6 ในประชากร ไก่เนื้อ
หลากหลายพันธุ์ในประเทศไทย



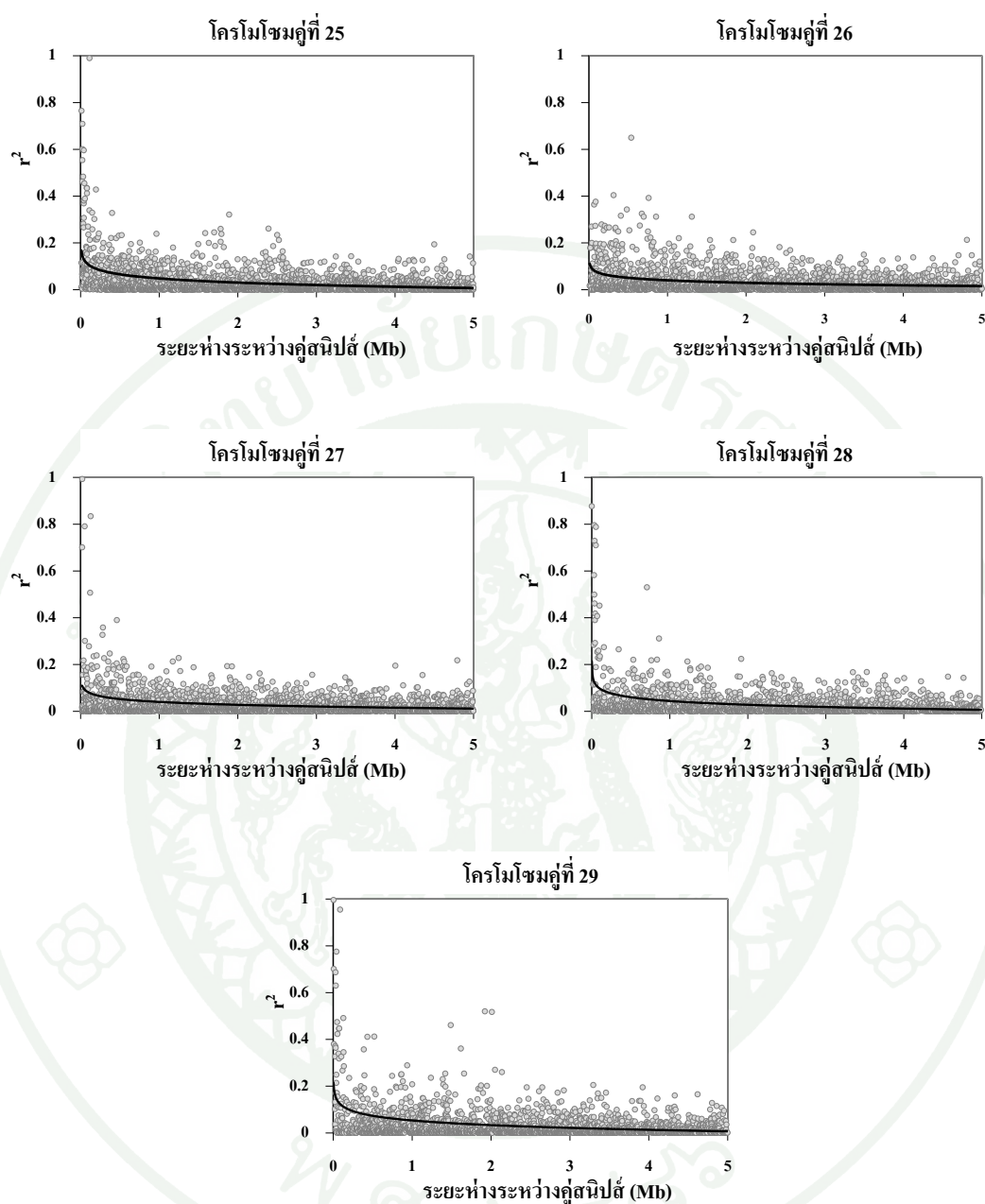
ภาพผนวกที่ 12 การกระจายตัวของ r^2 บนโครโมโซมร่างกายคู่ที่ 7 ถึง 12 ในประชากร โคนม
หลากหลายพันธุ์ในประเทศไทย



ภาพผนวกที่ 13 การกระจายตัวของ r^2 บนโครโมโซมร่างกายคู่ที่ 13 ถึง 18 ในประชากรโคนม
หลากหลายพันธุ์ในประเทศไทย



ภาพผนวกที่ 14 การกระจายตัวของ r^2 บนโครโมโซมร่างกายคู่ที่ 19 ถึง 24 ในประชากรโคนม
หลากหลายพันธุ์ในประเทศไทย



ภาพผนวกที่ 15 การกระจายตัวของ r^2 บนโครโมโซมร่างกายคู่ที่ 25 ถึง 29 ในประชากรโคนม
หลากหลายพันธุ์ในประเทศไทย

ตารางผนวกที่ 1 จำนวนคู่สนิปส์ที่มีระดับ LD สูงบนโครโมโซมร่างกายแต่ละคู่ของโคนม ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย

โครโมโซม คู่ที่	จำนวนคู่สนิปส์ ทั้งหมด	จำนวนคู่สนิปส์ (ร้อยละ)			
		$D' = 1$	$D' > 0.8$	$r^2 = 1$	$r^2 > 0.2$
1	8,750	238 (2.72)	549 (6.27)	0 (0.00)	507 (5.79)
2	6,477	70 (1.08)	244 (3.77)	0 (0.00)	231 (3.57)
3	7,104	151 (2.13)	525 (7.39)	27 (0.38)	456 (6.42)
4	4,992	2 (0.04)	41 (0.82)	0 (0.00)	74 (1.48)
5	5,911	269 (4.55)	598 (10.12)	4 (0.07)	233 (3.94)
6	8,699	51 (0.59)	231 (2.66)	2 (0.02)	295 (3.39)
7	4,865	11 (0.23)	82 (1.69)	0 (0.00)	140 (2.88)
8	4,870	11 (0.23)	92 (1.89)	0 (0.00)	147 (3.02)
9	4,826	76 (1.57)	274 (5.68)	0 (0.00)	179 (3.71)
10	4,605	4 (0.09)	63 (1.37)	0 (0.00)	149 (3.24)
11	4,937	7 (0.14)	61 (1.24)	0 (0.00)	85 (1.72)
12	3,576	20 (0.14)	79 (2.21)	0 (0.00)	86 (2.40)
13	3,930	10 (0.25)	102 (2.60)	0 (0.00)	125 (3.18)
14	4,090	5 (0.12)	89 (2.18)	0 (0.00)	153 (3.74)
15	4,079	22 (0.54)	97 (2.38)	2 (0.05)	120 (2.94)
16	5,969	314 (5.26)	868 (14.54)	32 (0.54)	685 (11.48)
17	2,903	6 (5.26)	38 (1.31)	0 (0.00)	65 (2.24)
18	3,666	48 (1.31)	183 (4.99)	0 (0.00)	97 (2.65)
19	3,437	31 (0.90)	112 (3.26)	1 (0.03)	123 (3.58)
20	5,875	281 (4.78)	875 (14.89)	27 (0.46)	523 (8.90)
21	4,353	149 (3.42)	394 (9.05)	1 (0.02)	254 (5.84)
22	2,933	11 (0.38)	27 (0.92)	0 (0.00)	44 (1.50)
23	4,025	83 (2.06)	304 (7.55)	6 (0.15)	308 (7.65)
24	3,214	5 (0.16)	37 (1.15)	0 (0.00)	62 (1.93)
25	2,472	21 (0.85)	73 (2.95)	0 (0.00)	46 (1.86)
26	2,489	18 (0.72)	57 (2.29)	0 (0.00)	33 (1.33)
27	2,301	10 (0.43)	54 (2.35)	0 (0.00)	22 (0.96)
28	2,018	6 (0.30)	26 (1.29)	0 (0.00)	28 (1.39)
29	2,192	10 (0.46)	33 (1.51)	0 (0.00)	48 (2.19)
ทั้งหมด	127,618	1,940 (1.50)	6,208 (4.79)	102 (0.08)	5,318 (4.10)

**การเขียนชุดคำสั่งในโปรแกรม SAS สำหรับใช้ในการวิเคราะห์ความผันแปรทาง
พันธุกรรมสปีส์ในระดับจีโนมของโคนม ในประชากรโคนมหลากหลายพันธุ์ในประเทศไทย**

DATA SNPs;

INFILE 'D:\TWL-MS-Program\TWL-THESIS\DATA_Thesis\Variation

SNPs\SNP_SUMARY_20140714_SASinput.txt' DLM='09'x FIRSTOBS=2 LRECL=200000;

INPUT SNPID\$ Chr\$ Posit NoCalls CallFreq AA AB BB A B MAF;

/*

SNPID = ID of SNPs **AA** = Frequency of SNPs genotype AA

Chr = Chromosome (Autosomes 1 – 29; Sex Chromosome) **AB** = Frequency of SNPs genotype AB

Posit = Physical position (base pair) of SNPs **BB** = Frequency of SNPs genotype BB

at each chromosome **A** = Frequency of SNPs allele A

NoCalls = No call rate of the SNPs **B** = Frequency of SNPs allele B

CallFreq = Fraction of the number SNPs that have call rate **MAF** = Minor allele frequency

*/

/*

*/

PROC SORT; BY Chr;

IF MAF <= 0.05 THEN X = 1;

PROC UNIVARIATE; BY Chr;

IF MAF > 0.05 THEN X = 2;

VAR Posit NoCalls Calls AA AB BB A B

IF MAF > 0.10 THEN X = 3;

MAF;

IF MAF > 0.15 THEN X = 4;

RUN;

ประวัติการศึกษาและการทำงาน

ชื่อ นายทวี เหล่าคิ้ม
เกิดวันที่ 11 ธันวาคม 2532
สถานที่เกิด อำเภอแก่งคร้อ จังหวัดชัยภูมิ
ประวัติการศึกษา วท.บ. (สัตวศาสตร์) มหาวิทยาลัยมหาสารคาม
ตำแหน่งปัจจุบัน -
สถานที่ทำงานปัจจุบัน -
ผลงานดีเด่นและ/หรือรางวัลทางวิชาการ -
ทุนการศึกษาที่ได้รับ ทุนสนับสนุนการศึกษาจากสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยี
แห่งชาติ (สวทช.) พ.ศ. 2555 - 2557

ผลงานวิจัยในวิทยานิพนธ์ที่ได้รับการตีพิมพ์เผยแพร่

1) วารสารวิชาการ

Laodim, T., S. Koonawootrittriron, M.A. Elzo and T. Suwanasopee. 2014. Genome-Wide Linkage Disequilibrium in a Thai Multibreed Dairy Cattle Population. **J. Anim Breed. Genet.** (Submitted)

ทวี เหล่าดีม, ศกร คุณวุฒิตถิธรณ, ธนาทิพย์ สุวรรณโสภี และ Mauricio A. Elzo. 2557. รูปแบบความผันแปรของสปีส์ในระดับจีโนมของโคนมไทยหลากหลายพันธุ์, น 17-24. ใน การประชุมทางวิชาการของมหาวิทยาลัยเกษตรศาสตร์ ครั้งที่ 52 (สาขาสัตว). มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ.

2) วารสารหรือสื่ออื่นๆ

ทวี เหล่าดีม, ศกร คุณวุฒิตถิธรณ, ธนาทิพย์ สุวรรณโสภี และ Mauricio A. Elzo. 2557. รูปแบบความผันแปรจีโนมของสปีส์ในประชากรโคนมไทยหลากหลายพันธุ์. หน้า 66-67. ใน ค่าการผสมพันธุ์โคนม 2557. องค์การส่งเสริมกิจการโคนมแห่งประเทศไทย, กระทรวงเกษตรและสหกรณ์. [ISSN 1905-7504]