วารสารแก่นเกษตร
THAIJO

Content List Available at ThaiJo

Khon Kaen Agriculture Journal

Journal Home Page : https://li01.tci-thaijo.org/index.php/agkasetkaj

การทำนายค่าพันธุกรรมจีโนมของลักษณะการให้ผลผลิตน้ำนมสำหรับโคนมไทยโดยใช้วิธีการแบบขั้นตอนเดียวด้วยโมเดลวันทดสอบการถดถอยแบบสุ่ม

Genomic prediction of milk production traits for Thai dairy cattle using single-step approach with random regression test-day model

สายัณห์ บัวบาน[1]*, สมศักดิ์ เปรมปรีดิ์[1] และ มนต์ชัย ดวงจินดา[2]

Sayan Buaban[1]*, Somsak Prempree[1] and Monchai Duangjinda[2]

[1] สำนักเทคโนโลยีชีวภาพการผลิตปศุสัตว์ กรมปศุสัตว์ ต. บางกะดี อ.เมือง จ.ปทุมธานี 12000
[1] Bureau of Biotechnology in Livestock Production, Department of Livestock Development, Bangkadee, Muang, Pathumthani, 12000
[2] สาขาวิชาสัตวศาสตร์ คณะเกษตรศาสตร์ มหาวิทยาลัยขอนแก่น อ.เมือง จ.ขอนแก่น 40002
[2] Department of Animal Science, Faculty of Agriculture, Khon Kaen University, Muang, Khon Kaen 40002

บทคัดย่อ: การทำนายค่าพันธุกรรมจีโนมโดยใช้วิธีการแบบขั้นตอนเดียว (single-step genomic BLUP, ssGBLUP) ได้แสดงให้เห็นว่ามีประสิทธิภาพสูงกว่าวิธีการแบบหลายขั้นตอน และการวิเคราะห์ทางสถิติของลักษณะการให้ผลผลิตน้ำนมด้วยโมเดลวันทดสอบการถดถอยแบบสุ่ม (random regression test-day model, RR-TDM) มีข้อดีที่ชัดเจนกว่าโมเดลอื่น ๆ การวิจัยครั้งนี้จึงมีวัตถุประสงค์เพื่อประเมินความสามารถในการทำนายค่าพันธุกรรมของลักษณะการให้ผลผลิตน้ำนมโดยใช้วิธีการแบบขั้นตอนเดียวด้วยโมเดลวันทดสอบการถดถอยแบบสุ่ม (single-step random regression test-day model, SS-RR-TDM) เปรียบเทียบกับวิธีการแบบดั้งเดิมโดย RR-TDM ในประชากรโคนมไทย และศึกษาผลกระทบจากการใช้ข้อมูลจีโนไทป์ของแม่โคต่อการทำนายค่าพันธุกรรมจีโนมสำหรับพ่อโคที่อายุยังน้อย ชุดข้อมูลของลักษณะปริมาณน้ำนม (n = 170,666) และลักษณะองค์ประกอบน้ำนม (n = 160,526) ได้มาจากแม่โคนมในรอบการให้นมครั้งแรกที่คลอดลูกระหว่างเดือนพฤศจิกายน 2536 ถึงเดือนมีนาคม 2560 จำนวน 24,858 และ 23,201 ตัว ตามลำดับ ข้อมูลจีโนไทป์ที่ได้จาก Illumina Bovine SNP50 BeadChip สำหรับลักษณะปริมาณน้ำนมมีจำนวน 876 ตัว และลักษณะองค์ประกอบน้ำนมมีจำนวน 868 ตัว เราตัดข้อมูลในช่วงหกปีสุดท้าย และสัตว์ที่ใช้ตรวจสอบความแม่นยำกำหนดให้เป็นพ่อโคที่ไม่มีลูกสาวในชุดข้อมูลที่ถูกตัดทอน ค่า EBVs (estimated breeding values) ที่ใช้พันธุ์ประวัติแบบดั้งเดิมทำนายด้วย RR-TDM ส่วนค่าพันธุกรรมจีโนม (genomic estimated breeding values, GEBVs) ทำนายด้วย SS-RR-TDM เปรียบเทียบวิธีการทำนายโดยใช้ความสามารถในการทำนายค่าพันธุกรรมสำหรับพ่อโคที่มีอายุยังน้อย ได้แก่ ความแม่นยำเชิงทฤษฎี ความแม่นยำเชิงการตรวจสอบ และความไม่อคติ ความแม่นยำเชิงทฤษฎีได้จากการแปลงเมตริกซ์สัมประสิทธิ์ของสมการโมเดลแบบผสม (mixed model equations, MME) ในขณะที่ความแม่นยำเชิงการตรวจสอบได้จากสัมประสิทธิ์สหสัมพันธ์เพียร์สันระหว่าง dEBV (deregressed proofs) จากชุดข้อมูลทั้งหมด และ G(EBV) จากชุดข้อมูลที่ลดลง ความไม่อคติจะพิจารณาจากค่าสัมประสิทธิ์การถดถอยที่คำนวณตามโมเดลการถดถอยเชิงเส้น (มีค่าเข้าใกล้ 1) ผลการศึกษาแสดงให้เห็นว่าการทำนายค่าทางพันธุกรรมจีโนมโดยใช้จีโนไทป์พ่อโคเพียงอย่างเดียวด้วย SS-RR-TDM ให้ค่าความแม่นยำเชิงทฤษฎี และความแม่นยำเชิงการตรวจสอบเพิ่มขึ้นโดยเฉลี่ย 0.22 และ 0.06 จุด ตามลำดับเมื่อเปรียบเทียบกับ RR-TDM สำหรับทุกลักษณะการให้ผลผลิตน้ำนม ขณะที่เมื่อนำข้อมูลจีโนไทป์ของแม่โคมาร่วมใช้ในการทำนายให้ความแม่นยำเชิงทฤษฎี และความแม่นยำเชิงการตรวจสอบเพิ่มขึ้นอีกโดยเฉลี่ย 0.02 และ 0.07 จุดตามลำดับ ค่าสัมประสิทธิ์การถดถอยที่ได้จาก SS-RR-TDM มีค่าเข้าใกล้ 1 มากกว่า RR-TDM และการทำนาย GEBV ที่เกินจริงมีแนวโน้มลดลงเมื่อมี

* Corresponding author: buaban_ai@hotmail.com

Received: date; December 29, 2020 Accepted: June 9, 2021 date; Published: date August , 2021

การใช้ข้อมูลจีโนไทป์ของแม่โคร่วมในการทำนาย การวิจัยครั้งนี้ช่วยยืนยันความเป็นไปได้ในการใช้ SS-RRTDM ในการทำนายพันธุกรรมจีโนมสำหรับลักษณะการให้ผลผลิตน้ำนมของโคนมไทยในอนาคต และควรใช้ข้อมูลจีโนไทป์ของทั้งพ่อและแม่โคในการทำนายค่าพันธุกรรมจีโนมในประชากรโคนมไทย
คำสำคัญ: โคนมไทย; โมเดลในวันทดสอบการถดถอยแบบสุ่ม; ลักษณะการให้ผลผลิตน้ำนม; วิธีการแบบขั้นตอนเดียว

ABSTRACT: In genomic prediction, the single-step genomic BLUP (ssGBLUP) has been demonstrated to outperform multi-step methods. In statistical analyses of milk production traits, the random regression test-day model (RR-TDM) has clear advantages over other models. This study aimed to evaluate the feasibility of using the single-step random regression test-day model (SS-RR-TDM) in genomic prediction of milk production traits, in comparison with the pedigree-based RR-TDM, and to investigate an effect of genotyped cows on the accuracy of genomic prediction for young bulls. Data of milk yield (n=170,666) and milk components (n= 160,526) were from 24,858 and 23,201 cows in first lactation, calving between November 1993 and March 2018, respectively. Additionally, 876 and 868 bulls and cows of each data set were genotyped using Illumina Bovine SNP50 BeadChip. We cut off the data in the last six years, and the validation animals were defined as genotyped bulls with no daughters in the truncated set. Estimated breeding values (EBVs) were obtained with the traditional pedigree-based RR-TDM, and genomic estimated breeding values (GEBVs) were estimated with SS-RR-TDM. The prediction methods were compared with the genetic predictive ability for young bulls, namely theoretical accuracy, validation accuracy, and unbiasedness. Theory accuracies were obtained by inverting the coefficient matrix of the mixed model equations (MME) whereas validation accuracies were measured by the Pearson correlation between de-regressed EBV from the full dataset and (G)EBV predicted with the reduced dataset. The unbiasedness is determined by the regression coefficient calculated according to the linear regression model (closed to 1). For prediction of all milk production traits using only bull genotypes, on average, SS-RR-TDM increased theoretical accuracies by 0.22 and validation accuracies by 0.07, compared with RR-TDM. With cow genotypes, the extra increase was 0.02 and 0.07 for theoretical accuracies and validation accuracies, respectively. The regression coefficient obtained from SS-RR-TDM was closer to 1 more than RR-TDM. The inflation of GEBV tended to be reduced using cow genotypes in prediction. This study confirmed the feasibility of SS-RR-TDM, thus enhancing the strategies for genomic prediction of milk production traits in the future for Thai dairy population. We recommend that the evaluation should include the genotype data of both the sires and cows.
Keywords: Thai dairy cattle; random regression test-day model; milk production traits; single-step genomic BLUP

บทนำ

การทำนายค่าพันธุกรรมเป็นเครื่องมือที่สำคัญสำหรับใช้ในการจัดลำดับและคัดเลือกสัตว์ตามพันธุกรรมในแผนการปรับปรุงพันธุ์ อันจะทำให้การผลิตสัตว์ในรุ่นถัดไปเป็นตามวัตถุประสงค์ของการปรับปรุงพันธุ์ โดยเฉพาะลักษณะการให้ผลผลิตน้ำนมซึ่งเป็นลักษณะสำคัญทางเศรษฐกิจที่มีผลต่อการทำกำไรของการเลี้ยงโคนม โคนมที่มีการให้ผลผลิตน้ำนมเพิ่มขึ้นทั้งปริมาณ และคุณภาพมีความสัมพันธ์โดยตรงกับรายได้ที่มากขึ้น (Seegers et al., 2003) ลักษณะนี้เป็นลักษณะที่ควบคุมด้วยยีนหลายคู่ (polygenic traits) ที่ได้รับผลกระทบจากปัจจัยหลายประการ และหลายยีน ซึ่งแต่ละยีนมีผลกระทบเล็กน้อยต่อการแสดงออกของลักษณะ (Snelling et al., 2013) การปรับปรุงด้านการจัดการและโภชนาการ พร้อมกับการคัดเลือกพันธุกรรมอย่างเข้มข้นสามารถเพิ่มการผลิตน้ำนมได้ ในครั้งแรกลักษณะการให้ผลิตน้ำนมมีการทำนายค่าพันธุกรรมบนพื้นฐานของลักษณะการแสดงออกแบบสะสมโดยใช้โมเดลรอบการให้นม (lactation model, LAM) จนถึงปัจจุบันการทำนายค่าพันธุกรรมของลักษณะดังกล่าวได้มุ่งเน้นไปที่การใช้โมเดลวันทดสอบการถดถอยแบบสุ่ม (random regression test-day model, RR-TDM) ซึ่งเป็นโมเดลที่สามารถอธิบายปัจจัยทางพันธุกรรมและทางสภาพแวดล้อมทั้งหมดที่มีผลกระทบต่อการเปลี่ยนแปลงข้อมูลฟีโนไทป์ของแม่โคได้โดยตรงในแต่ละวันทดสอบตลอดระยะการให้นม (Oliveira et al., 2019a) ด้วยวิธีการทำนายค่าพันธุกรรมแบบไม่เอนเอียงเชิงเส้นตรงที่ดีที่สุด (Best Linear Unbiased Prediction, BLUP) (Henderson, 1984) โดยใช้ข้อมูลฟีโนไทป์ร่วมกับข้อมูลพันธุ์ประวัติ สำหรับการทำนายค่าการผสมพันธุ์ (estimated breeding value, EBV) เพื่อใช้ในการคัดเลือกแบบดั้งเดิม (Traditional pedigree-

based selection) ส่งผลให้การทำนายค่า EBV ด้วย RR-TDM มีความแม่นยำสูงกว่าเมื่อเปรียบเทียบกับ LAM หรือวิธีการทางสถิติอื่นๆ (Kistemaker, 1997; Buaban and Sanpote, 2010) อย่างไรก็ตามวิธีการนี้ยังคงมีข้อจำกัดในการใช้งานอยู่บ้าง เช่น ในการคัดเลือกสัตว์ให้เป็นไปตามวัตถุประสงค์มักใช้ระยะเวลานาน และได้ผลช้า ในทศวรรษที่ผ่านมาได้มีความก้าวหน้าของเทคโนโลยีจีโนม ส่งผลให้การจำแนกข้อมูลของความผันแปรทางพันธุกรรมในระดับนิวคลีโอไทด์ หรือ single nucleotide polymorphisms (SNPs) ของสัตว์รายตัวสามารถทำได้ครั้งละจำนวนมาก ครอบคลุมทั่วทั้งจีโนม และมีราคาถูกลงอย่างต่อเนื่อง (VanRaden, 2020) จึงได้นำการคัดเลือกด้วยจีโนม (genomic selection, GS) ซึ่งหมายถึงการคัดเลือกจากการทำนายค่าพันธุกรรมจีโนม (genomic estimated breeding value, GEBV) โดยใช้ข้อมูล SNPs ร่วมกับข้อมูลฟีโนไทป์และข้อมูลพันธุ์ประวัติ (Legarra et al., 2009; Misztal et al., 2009) มาแทนการคัดเลือกแบบดั้งเดิม โดยเฉพาะในโคนมที่เกิดความสำเร็จในการพัฒนาในหลายประเทศ เช่น ประเทศสหรัฐอเมริกา แคนาดา สหราชอาณาจักร นิวซีแลนด์ ออสเตรเลีย ฝรั่งเศส เนเธอร์แลนด์ เยอรมนี และประเทศสแกนดิเนเวีย (Silva et al., 2014) การคัดเลือกด้วยจีโนมมีข้อดีหลายอย่างเมื่อเปรียบเทียบกับการคัดเลือกแบบดั้งเดิม เช่น มีศักยภาพในการเพิ่มความก้าวหน้าทางพันธุกรรมของลักษณะที่สำคัญทางเศรษฐกิจสำหรับโคนมในแต่ละปีเพิ่มขึ้น (Doublet et al., 2019) ผ่านความแม่นยำในการทำนายค่าพันธุกรรม และลดช่วงอายุจากการที่สามารถคัดเลือกพ่อแม่พันธุ์โคนมได้ในขณะที่ยังไม่มีข้อมูลของลูกสาว (Hayes et al., 2009) ที่ทำได้ในช่วงต้นของชีวิตสัตว์ หรือบางครั้งก่อนที่สัตว์จะเกิด นอกจากนี้การคัดเลือกด้วยจีโนมสามารถลดต้นทุนในการสร้างพ่อพันธุ์ได้เป็นอย่างมากเมื่อเทียบกับโปรแกรมการทดสอบลูกสาวแบบดั้งเดิม โดยทั่วไปวิธีการทำนายค่าพันธุกรรมจีโนมสามารถแบ่งได้เป็นสองประเภทหลัก: วิธีการแบบหลายขั้นตอน (multi-step approach) และแบบขั้นตอนเดียว (single-step approach) โดยวิธีแบบขั้นตอนเดียว (single-step genomic BLUP, ssGBLUP) เป็นวิธีการที่ได้รับความนิยมในปัจจุบัน และมีข้อดีกว่าวิธีการทำนายแบบหลายขั้นตอน (Misztal et al., 2020) เนื่องจากสามารถทำนายสัตว์ที่มีจีโนไทป์ และไม่มีจีโนไทป์ทั้งหมดได้ในเวลาเดียวกันได้ทั้งหมด ลดความซับซ้อนในทางปฏิบัติ และมีอคติในการทำนายน้อยลง นอกจากนี้วิธีการแบบขั้นตอนเดียวยังให้การทำนายที่มีความแม่นยำมากขึ้น และมีคุณสมบัติที่สามารถใช้ร่วมกับ RR-TDM สำหรับลักษณะการให้ผลผลิตน้ำนมที่มีข้อมูลต่อเนื่องได้ (Kang et al., (2017) จากการศึกษาก่อนหน้านี้แสดงให้เห็นว่าการทำนายค่าพันธุกรรมโดยวิธีการแบบขั้นตอนเดียวด้วยโมเดลวันทดสอบการถดถอยแบบสุ่ม (single-step random regression test-day model, SS-RR-TDM) มีความเป็นไปได้ที่จะให้ความแม่นยำมากขึ้น และมีอคติของการทำนายน้อยลง (Oliveira et al., 2019b) การทำนายค่าพันธุกรรมสำหรับลักษณะการให้ผลผลิตน้ำนมในประเทศไทย สำนักเทคโนโลยีชีวภาพการผลิตปศุสัตว์ กรมปศุสัตว์ ได้ดำเนินการด้วย RR-TDM โดยใช้ข้อมูลฟีโนไทป์ร่วมกับข้อมูลพันธุ์ประวัติมาตั้งแต่ปี 2556 (กรมปศุสัตว์, 2556) และในปัจจุบันได้มีการรวบรวมข้อมูลจีโนไทป์จากทั้งพ่อพันธุ์ที่ผ่านการพิสูจน์ด้วยการทดสอบลูกสาว และจากแม่โค ซึ่งหากมีการผสมผสานระหว่าง RR-TDM เข้ากับ ssGBLUP (SS-RR-TDM) ในการทำนายค่าพันธุกรรม น่าจะเพิ่มความแม่นยำ และมีอคติของการทำนายน้อยลง และส่งผลทำให้ความสำเร็จในการปรับปรุงพันธุ์โคนมของประเทศเป็นไปอย่างรวดเร็ว อย่างไรก็ตามในประเทศไทยมีพ่อพันธุ์ที่ผ่านการทดสอบด้วยข้อมูลของลูกสาวไม่มากนัก การเพิ่มข้อมูลจีโนไทป์ของแม่พันธุ์ที่มีข้อมูลฟีโนไทป์ร่วมในการทำนายค่าพันธุกรรมแทนที่จะใช้จีโนไทป์ของพ่อพันธุ์เพียงอย่างเดียวน่าจะทำให้การทำนายค่าพันธุกรรมจีโนมมีความน่าเชื่อถือมากขึ้น ดังนั้นการวิจัยครั้งนี้จึงมีวัตถุประสงค์เพื่อประเมินความสามารถในการทำนายค่าพันธุกรรมของลักษณะการให้ผลผลิตน้ำนมโดยใช้วิธีการแบบขั้นตอนเดียวด้วยโมเดลวันทดสอบการถดถอยแบบสุ่ม (SS-RR-TDM) เปรียบเทียบกับวิธีการแบบดั้งเดิมโดย RR-TDM ในประชากรโคนมไทย และศึกษาผลกระทบจากการใช้ข้อมูลจีโนไทป์ของแม่โคต่อการทำนายค่าพันธุกรรมจีโนมสำหรับพ่อโคที่อายุยังน้อย

วิธีการศึกษา

ข้อมูลการให้อาหาร

โคนมในประเทศไทยมักเลี้ยงในสภาพแวดล้อมที่เป็นฟาร์มขนาดเล็กถึงขนาดกลางเป็นส่วนใหญ่ อาหารและวิธีการให้อาหารมีความแตกต่างกันไปในแต่ละภูมิภาค เมื่อพิจารณาจากคุณภาพอาหารข้นที่ใช้ เกษตรกรจำนวนมากจะให้อาหารข้น 1 กก. การให้ปริมาณผลิตน้ำนม

2 กก. ต่อวันเป็นบรรทัดฐานในการปฏิบัติ ส่วนอาหารหยาบที่ให้ประกอบด้วย หญ้าในเขตร้อน ฟางข้าว และของเหลือจากการเกษตรบางส่วน องค์ประกอบของส่วนผสมของอาหารข้นขึ้นอยู่กับความพร้อมในท้องถิ่นและราคา อย่างไรก็ตามการขาดแคลนอาหารหยาบเป็นเรื่องวิกฤติในฤดูหนาวและฤดูร้อน และเกษตรกรต้องซื้อหญ้าแห้งหรือฟางข้าว หรือเพิ่มการใช้อาหาร TMR (Total Mixed Ration) หรือ Complete Ration (CR) ซึ่งเป็นอาหารผสมสำเร็จรูปที่เกิดจากการนำอาหารหยาบ และอาหารข้นมาผสมกันในอัตราส่วนที่เหมาะสม

ข้อมูลฟีโนไทป์ และพันธุ์ประวัติ

ข้อมูลในวันทดสอบรายเดือน (monthly test-day records) ของแม่โคนมในรอบการให้นมครั้งแรกที่คลอดลูกระหว่างเดือนพฤศจิกายน 2536 ถึงเดือนมีนาคม 2560 จากฐานข้อมูลโคนมของสำนักเทคโนโลยีชีวภาพการผลิตปศุสัตว์ กรมปศุสัตว์ และเพื่อให้ได้ชุดข้อมูลที่มีความเหมาะสมในการวิเคราะห์จึงได้ตรวจสอบข้อมูลตามเงื่อนไขของ Buaban et al. (2020) ข้อมูลที่ผ่านการตรวจสอบแล้วจะเรียกว่าเป็นชุดข้อมูลทั้งหมด (full data set) ซึ่งมีข้อมูลในวันทดสอบถึงปี 2560 ข้อมูลในวันทดสอบประกอบด้วย 7 ลักษณะ: ปริมาณน้ำนม (milk yield, MY, กก.) ปริมาณไขมัน (fat yield, FY, กก.), ปริมาณโปรตีน (protein yield, PY, กก.) ปริมาณของแข็งทั้งหมด (total solid yield, TY, กก.) เปอร์เซ็นต์ไขมัน (fat percentage, FP, %) เปอร์เซ็นต์โปรตีน (protein percentage, PP, %) และเปอร์เซ็นต์ของแข็งทั้งหมด (total solid percentage, TP, %) เพื่อให้ข้อมูลมีความเหมาะสมในการวิเคราะห์ และเนื่องจากในวันทดสอบเดียวกันบางลักษณะไม่มีการบันทึก ผู้วิจัยจึงแบ่งลักษณะการให้ผลผลิตน้ำนมที่ศึกษาออกเป็น 2 กลุ่ม: ลักษณะปริมาณน้ำนม (MY) และลักษณะองค์ประกอบน้ำนม (FY, PY, TY, FP, PP และ TP) โดยกลุ่มลักษณะปริมาณน้ำนมมีจำนวน 170,666 ข้อมูล (24,858 ตัว) และกลุ่มลักษณะองค์ประกอบน้ำนมมีข้อมูล 160,526 ข้อมูล (23,201 ตัว) โดยชุดข้อมูลที่ได้ในแต่ละกลุ่มจะสืบย้อนกลับพันธุ์ประวัติไปถึงสามชั่วอายุจากแม่โคที่มีข้อมูล สำหรับการตรวจสอบพ่อโคที่อายุยังน้อย ผู้วิจัยจะต้องเตรียมชุดข้อมูลเป็น 2 ชุด คือ ชุดข้อมูลทั้งหมด และชุดข้อมูลที่ลดลง (reduced data set) โดยชุดข้อมูลทั้งสองจะต้องอยู่ในฐานพันธุกรรมเดียวกัน ชุดข้อมูลที่ลดลงถูกดึงมาจากชุดข้อมูลทั้งหมดโดยตัดข้อมูลในวันทดสอบในช่วงหกปีสุดท้ายออกเพื่อต้องการให้พ่อโคที่อายุยังน้อยที่ใช้ตรวจ (validation of young bulls) ไม่มีข้อมูลฟีโนไทป์ของลูกสาว ดังนั้นชุดข้อมูลนี้จึงมีข้อมูลในวันทดสอบถึงปี 2554 นอกจากนี้พ่อโคที่อายุยังน้อยที่ใช้ตรวจสอบยังต้องมีความแม่นยำจากการทำนายค่า EBV สำหรับแต่ละลักษณะจากชุดข้อมูลทั้งหมดมากกว่า 75%

ข้อมูลจีโนไทป์

ข้อมูลจีโนไทป์ที่ได้จากการจำแนกข้อมูลของความผันแปรทางพันธุกรรมในระดับนิวคลีโอไทด์ (SNPs) ด้วย Illumina BovineSNP50 BeadChip (Illumina Inc., San Diego, CA, USA) ที่เป็นเวอร์ชัน 2 (มีจำนวน 54,609 SNPs) หรือ เวอร์ชัน 3 (มีจำนวน 53,218 SNPs) โดยเลือกใช้ SNPs ที่เหมือนกันทั้งสองเวอร์ชัน จำนวน 50,908 SNPs ไปผ่านการควบคุมคุณภาพ ซึ่งพิจารณาด้วย call rate > 0.9 สำหรับทั้ง SNPs และจีโนมสัตว์ MAF (minor allele frequency) > 0.05 การออกจากสมดุล Hardy-Weinberg (ความแตกต่างระหว่างความถี่ที่คาดหวัง และความถี่ที่สังเกต) < 0.15 การทดสอบความขัดแย้งระหว่างคู่พ่อแม่ลูกหลาน (parent-progeny conflict) โดยโปรแกรม PREGSF90 (Misztal, 2018) ได้นำมาใช้ในการควบคุมคุณภาพ SNPs และตัวอย่าง ภายหลังจากการควบคุมคุณภาพ ยังคงเหลือข้อมูล SNPs สำหรับใช้ในการวิเคราะห์จำนวน 41,977 และ 41,975 SNPs จากสัตว์ที่มีจีโนไทป์จำนวน 876 และ 868 ตัว สำหรับลักษณะปริมาณน้ำนม และลักษณะองค์ประกอบน้ำนม ตามลำดับ รายละเอียดของชุดข้อมูลทั้งหมดที่ใช้ในการศึกษานี้แสดงไว้ใน **Table 1**

Table 1 Number of farms, animals and records used in this study

Items	Number of animals and records	
	Milk yield traits	Milk component traits
Number of farms		
Full data set	1,905	1,764
Reduced data set	1,523	1,379
Animals in pedigree	51,128	48,834
Animals with records		
Full data set	24,858	23,201
Reduced data set	15,404	13,768
Test-day records		
Full data set	170,666	160,526
Reduced data set	97,970	88,000
Genotyped animals	876	868
Bulls	152	152
Cows	724	716
SNP information	41,977	41,975

โมเดลทางสถิติ

ทุกลักษณะที่ศึกษาครั้งนี้จะทำนายค่าพันธุกรรมโดยใช้วิธีการแบบดั้งเดิมด้วย RR-TDM และวิธีการแบบขั้นตอนเดียวด้วย SS-RR-TDM โมเดลที่ใช้มีความคล้ายคลึงกับที่ใช้ในการทำนายค่าพันธุกรรมของกรมปศุสัตว์ (กรมปศุสัตว์, 2562; Buaban et al., 2020)

RR-TDM เป็นวิธีการแบบดั้งเดิมที่ใช้ในการทำนายค่าพันธุกรรมโดยใช้ข้อมูลฟีโนไทป์ ร่วมกับข้อมูลพันธุ์ประวัติของสัตว์แต่ละตัวด้วยวิธี Best Linear Unbiased Prediction (BLUP) โดยได้ EBV เป็นผลลัพท์โดยตรง และใช้เป็นโมเดลอ้างอิงในการศึกษาครั้งนี้ ซึ่งมีโมเดลทางสถิติดังนี้:

$$\mathbf{y} = \mathbf{X_1 b_1} + \mathbf{X_2 b_2} + \mathbf{Vh} + \mathbf{Za} + \mathbf{Wp} + \mathbf{e}, \quad (1)$$

โดยที่ $\mathbf{y}$ เป็นเวกเตอร์ของบันทึกข้อมูลในวันทดสอบของลักษณะที่ทำการศึกษา (MY, FY, PY, TY, FP, PP และ TP) $\mathbf{b_1}$ เป็นเวกเตอร์ของอิทธิพลคงที่ซึ่งประกอบด้วย หน่วยผสมเทียม-ปี-ฤดูกาล กลุ่มพันธุ์-กลุ่มอายุที่คลอด $\mathbf{b_2}$ เป็นเวกเตอร์ของค่าสัมประสิทธิ์การถดถอยคงที่ของ Legendre polynomials (**LPs**) ที่ซ้อนอยู่ในกลุ่มพันธุ์-กลุ่มอายุที่คลอด $\mathbf{h}$ เป็นเวกเตอร์ของอิทธิพลสุ่มเนื่องจาก ฝูง-ปี-เดือนทดสอบ $\mathbf{a}$ และ $\mathbf{p}$ เป็นเวกเตอร์ของสัมประสิทธิ์การถดถอยแบบสุ่มของ LPs ที่ซ้อนกันภายในอิทธิพลเนื่องจากพันธุกรรม และสิ่งแวดล้อมถาวร ตามลำดับ และ $\mathbf{e}$ เป็นเวกเตอร์ของอิทธิพลของความคลาดเคลื่อน เมตริกซ์ $\mathbf{X_1}$, $\mathbf{X_2}$, $\mathbf{V}$, $\mathbf{Z}$ และ $\mathbf{W}$ เป็น incidence matrices ที่สอดคล้องกันกับอิทธิพลที่กล่าวมาข้างต้น ลำดับของ LPs ซึ่งเป็นค่าแปรผันร่วมของปัจจัยคงที่ และปัจจัยสุ่ม ที่คำนวณจากวันให้นม (DIM) ตามที่กำหนดโดย Gengler et al. (1999) การศึกษาในครั้งนี้ใช้ LPs ลำดับที่ 3 (constant, linear, quadratic และ cubic) สำหรับลักษณะปริมาณน้ำนม และองค์ประกอบน้ำนม โดยจะใช้ LPs ลำดับเดียวกันสำหรับอิทธิพลการถดถอยแบบคงที่และแบบสุ่ม กลุ่มพันธุ์จำแนกเป็น 3 กลุ่ม ตามระดับสายเลือดโฮลสไตน์ ฟรีเซี่ยน (HF) ประกอบด้วย HF < 87.50 %, 87.50 ≤ HF < 93.75 และ HF ≥ 93.75% กลุ่มอายุที่คลอดลูกจำแนกเป็น 7 กลุ่ม โดยแต่ละกลุ่มห่างกันทุกๆ 3 เดือน กลุ่มแรกมีอายุน้อยกว่า 25 เดือน และกลุ่มสุดท้ายมีอายุ

มากกว่า 39 เดือน ฤดูกาลจำแนกเป็น 3 กลุ่ม คือ ฤดูหนาว (พฤศจิกายน-กุมภาพันธ์) ฤดูร้อน (มีนาคม-มิถุนายน) และฤดูฝน (กรกฎาคม-ตุลาคม)

เพื่อลดความซับซ้อนของโมเดลจึงสมมุติให้ตัวแปรของอิทธิพลของความคลาดเคลื่อนมีค่าเท่ากันตลอดการให้นม และกำหนดให้อิทธิพลแบบสุ่มมีการกระจายแบบปกติด้วยค่าเฉลี่ยเท่ากับ 0 และโครงสร้างความแปรปรวนร่วมของโมเดลกำหนดให้เป็นดังนี้

$$Var\begin{bmatrix}\mathbf{h}\\ \mathbf{a}\\ \mathbf{p}\\ \mathbf{e}\end{bmatrix}=\begin{bmatrix}\mathbf{I}\sigma^2_{htm} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{A}\otimes\mathbf{G}_0 & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{I}\otimes\mathbf{P}_0 & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma^2_e\end{bmatrix} \quad (2)$$

โดยที่ $\mathbf{G}_0$ และ $\mathbf{P}_0$ เป็นเมตริกซ์ความแปรปรวนร่วม 4 × 4 ของสัมประสิทธิ์การถดถอยแบบสุ่มสำหรับอิทธิพลเนื่องจากพันธุกรรมและสภาพแวดล้อมถาวร ตามลำดับ $\mathbf{A}$ เป็นเมตริกซ์ความสัมพันธ์ระหว่างสัตว์ในพันธุ์ประวัติ (numerator relationship matrix) $\mathbf{I}$ เป็นเมตริกซ์เอกลักษณ์ (Identity matrix) $\otimes$ เป็น Kronecker product σ^2_{htm} เป็นความแปรปรวนร่วมของฝูง-ปี-เดือนทดสอบ และ σ^2_e เป็นความแปรปรวนร่วมของความคลาดเคลื่อน

SS-RR-TDM เป็นวิธีการที่ใช้ในการทำนายค่าพันธุกรรมโดยใช้ข้อมูลฟีโนไทป์ ข้อมูลพันธุ์ประวัติของสัตว์แต่ละตัว และข้อมูลจีโนไทป์พร้อมกันในการวิเคราะห์โดยได้ GEBV เป็นผลลัพท์โดยตรง วิธีการนี้มีโมเดลเดียวกันกับ RR-TDM ยกเว้นการแทนที่ $\mathbf{A}$ ด้วย $\mathbf{H}$ ซึ่งเป็นการรวมเมตริกซ์ความสัมพันธ์เครือญาติทางพันธุ์ประวัติเข้ากับเมตริกซ์ความสัมพันธ์ทางจีโนม (genomic-based relationship matrices**,** $\mathbf{G}$) (Legarra et al., 2009) โดยสามารถเขียนในรูปผกผันดังนี้:

$$\mathbf{H}^{-1}=\mathbf{A}^{-1}+\begin{bmatrix}0 & 0\\ 0 & \tau\mathbf{G}^{-1}-\omega\mathbf{A}^{-1}_{22}\end{bmatrix} \quad (3)$$

เมื่อ $\mathbf{H}^{-1}$ คือเมตริกซ์ผกผันของความสัมพันธ์ทางพันธุกรรมที่เอาเมตริกซ์ผกผันของความสัมพันธ์ทางจีโนม ($\mathbf{G}^{-1}$) รวมเข้ากับเมตริกซ์ผกผันของความสัมพันธ์เครือญาติทางพันธุ์ประวัติเฉพาะตัวสัตว์ที่มีข้อมูลจีโนไทป์ ($\mathbf{A}^{-1}_{22}$), $\mathbf{A}^{-1}$ คือเมตริกซ์ผกผันของความสัมพันธ์เครือญาติทางพันธุ์ประวัติ τ และ ω คือ ปัจจัยปรับ (weighting factors), $\mathbf{G}$ คือเมตริกซ์ของความสัมพันธ์ทางทางจีโนม (genomic relationship matrix) ที่สร้างตามวิธีการแรกของ VanRaden (2008) โดย $\mathbf{G}=0.95\frac{\mathbf{zz'}}{2\sum_{i=1}^{m}p_i(1-p_i)}+0.05\mathbf{A}_{22}$ (เพื่อที่จะหลีกเลี่ยงปัญหา singularity) เมื่อ $\mathbf{Z}$ คือเมตริกซ์ของจีโนไทป์ที่ปรับตามความถี่ของอัลลีลแล้ว โดย $\mathbf{Z}=\mathbf{M}-\mathbf{P}$, $\mathbf{M}$ คือ centered genotypes (เช่น −1, 0, และ 1 ที่เป็นตัวแทนของ AA, Aa, และ aa, ตามลำดับ), $\mathbf{P}$ คือความถี่ของอัลลีลสำหรับ SNPs $\mathbf{i}$ (p_i) ในคอลัมน์ที่ i ที่แสดงตาม $2(p_i - 0.5)$, เมื่อ p_i คือความถี่ของอัลลีลที่โลกัส i และ m คือ จำนวนของ SNPs ทั้งหมด

ความแตกต่างระหว่าง $\mathbf{A}_{22}$ และ $\mathbf{G}$ อาจเกิดขึ้นหลายเหตุผลที่ควรพิจารณา เช่น ความไม่สมบูรณ์ของพันธุ์ประวัติ ความผิดพลาดของพันธุ์ประวัติ การกำหนดจีโนไทป์ที่ไม่ถูกต้อง และคุณภาพของจีโนไทป์ที่ไม่ดี อาจทำให้เกิดความแตกต่างดังกล่าว ส่งผลให้เกิดการมีอคติของค่าพันธุกรรมจีโนมสำหรับสัตว์ที่มีอายุยังน้อย เพื่อหลีกเลี่ยงปัญหาดังกล่าว และควบคุมอคติของการ convergence จึงมีการกำหนดน้ำหนัก (ω) ให้แก่ $\mathbf{A}^{-1}_{22}$ และในการวิเคราะห์เบื้องต้นน้ำหนักของ ω = 0.50 ให้ค่าอคติน้อยเนื่องจากค่าสัมประสิทธิ์การถดถอย (regression coefficients, b) ของ de-regressed EBV เข้าใกล้ 1 ซึ่งมีความเหมาะสมกับประชากรที่ศึกษา

การวิเคราะห์ข้อมูล

1. การประมาณองค์ประกอบความแปรปรวน และการทำนายค่าพันธุกรรม

องค์ประกอบความแปรปรวนของลักษณะที่ศึกษาถูกวิเคราะห์ด้วย RR-TDM แบบดั้งเดิมโดยใช้ชุดข้อมูลทั้งหมด และค่าองค์ประกอบความแปรปรวนที่ประมาณได้ในประชากรที่ศึกษานี้จะถูกใช้สำหรับทำนายค่าพันธุกรรมทั้งใน RR-TDM และ SS-RR-TDM โดยใช้โปรแกรม AIREMLF90 เพื่อประมาณองค์ประกอบความแปรปรวน และใช้โปรแกรม BLUPF90 เพื่อแก้สมการโมเดลแบบผสม (mixed model equation, MME) (Misztal et al., 2018) ความแปรปรวนของความคลาดเคลื่อนจากการทำนาย (prediction error variance,

PEV) ของสัมประสิทธิ์การถดถอยแบบสุ่มจะคำนวณได้ด้วยค่าผกผันของเมตริกซ์ด้านซ้ายของ MME จากโปรแกรม BLUPF90 ค่าพันธุกรรมของสัตว์แต่ละตัวจะถูกนำเสนอจากผลผลิตที่ 305 วัน โดยค่าพันธุกรรมที่ 305 วัน ได้กำหนดจากผลรวมของค่าพันธุกรรมระหว่างวันให้นมที่ 5 ถึง 305 สำหรับลักษณะปริมาณการให้ผลผลิตน้ำนมตามวิธีการของ Jamrozik et al. (1997) สำหรับลักษณะเปอร์เซ็นต์องค์ประกอบน้ำนม (FP, PP และ PT) ผลรวมของค่าพันธุกรรมในรอบการให้นมจะถูกแทนที่ด้วยค่าเฉลี่ย

ค่าอัตราพันธุกรรมต่อวัน (h_d^2) สำหรับลักษณะปริมาณการให้ผลผลิตน้ำนม และลักษณะเปอร์เซ็นต์องค์ประกอบน้ำนม คำนวณได้ดังนี้:

$$h_d^2 = \frac{\sigma_{a(d)}^2}{\sigma_{a(d)}^2+\sigma_{htm}^2+\sigma_{p(d)}^2+\sigma_e^2} \quad \textbf{(4)}$$

เมื่อ $\sigma_{a(d)}^2$ และ $\sigma_{p(d)}^2$ เป็นความแปรปรวนและความแปรปรวนร่วม ((co)variances) ทางพันธุกรรม และสิ่งแวดล้อมแบบถาวรสำหรับแต่ละวัน ตามลำดับ

ค่าอัตราพันธุกรรมสำหรับลักษณะปริมาณการให้ผลผลิตน้ำนม และลักษณะเปอร์เซ็นต์องค์ประกอบน้ำนม ที่ 305 วัน (h_{305d}^2) คำนวณได้ดังนี้:

$$h_{305d}^2 = \frac{\hat{\sigma}_{a(305)}^2}{301\hat{\sigma}_{htm}^2+\hat{\sigma}_{a(305)}^2+\hat{\sigma}_{p(305)}^2+301\hat{\sigma}_e^2}, \quad \text{สำหรับลักษณะปริมาณการให้ผลผลิตน้ำนม } \textbf{(5)}$$

$$h_{305d}^2 = \left(\frac{\hat{\sigma}_{a(305)}^2}{\hat{\sigma}_{htm}^2+\hat{\sigma}_{a(305)}^2+\hat{\sigma}_{p(305)}^2+\hat{\sigma}_e^2}\right) \quad \text{สำหรับลักษณะเปอร์เซ็นต์องค์ประกอบน้ำนม } \textbf{(6)}$$

เมื่อ $\hat{\sigma}_{a(305)}^2$และ $\hat{\sigma}_{p(305)}^2$ สำหรับลักษณะปริมาณการให้ผลผลิตน้ำนม เป็นความแปรปรวนและความแปรปรวนร่วมทางพันธุกรรมและสิ่งแวดล้อมแบบถาวรที่รวมกันตั้งแต่วันให้นมที่ 5 ถึง 305 ตามลำดับ (Jamrozik and Schaeffer, 1997; Jakobsen et al., 2002) ส่วน $\hat{\sigma}_{a(305)}^2$และ $\hat{\sigma}_{p(305)}^2$ สำหรับเปอร์เซ็นต์องค์ประกอบน้ำนม เป็นค่าความแปรปรวนและความแปรปรวนร่วมทางพันธุกรรม และสิ่งแวดล้อมแบบถาวรเฉลี่ยตั้งแต่วันให้นมที่ 5 ถึง 305 ตามลำดับ

PEV ของค่าพันธุกรรมที่ 305 วันจะคำนวณตามฟังก์ชันของ PEV ของสัมประสิทธิ์การถดถอยแบบสุ่ม และตัวแปรร่วมของ LPs (Mrode, 2014)

2. การตรวจสอบค่าพันธุกรรมจีโนมสำหรับพ่อโคที่อายุยังน้อย (การเปรียบเทียบวิธีการทำนายจากโมเดล RR-TDM กับ SS-RR-TDM)

การศึกษาครั้งนี้ได้ประเมินความสามารถในการทำนายค่าพันธุกรรมสำหรับพ่อโคที่มีอายุยังน้อย โดยใช้การตรวจสอบความแม่นยำกับข้อมูลที่ลดลง (cross-validation) ค่าการทำนายทางพันธุกรรมคือค่าเฉลี่ยของพ่อแม่ (PA) ใน RR-TDM และค่า GEBV ใน SS-RR-TDM พ่อโคที่ทำการตรวจสอบ (validation bulls) จะพิจารณาจากพ่อโคหนุ่มที่อายุยังน้อยในชุดข้อมูลที่ลดลงที่มีเฉพาะข้อมูลจีโนไทป์ และมีค่าความแม่นยำจากการทำนายค่า EBV สำหรับแต่ละลักษณะจากชุดข้อมูลทั้งหมดมากกว่า 75% พ่อโคที่ทำการตรวจสอบความแม่นยำมีจำนวนถึง 28 ตัว ดังแสดงใน Table 3 ในลำดับแรกนั้นจะคำนวณค่า EBV ที่ 305 วัน สำหรับพ่อโคที่ทำการตรวจสอบตาม RR-TDM โดยใช้ข้อมูลทั้งหมด จากนั้นจะสามารถคำนวณค่า de-regressed EBV (dEBV) สำหรับพ่อโคแต่ละตัวโดยใช้สูตรของ (VanRaden et al., 2009) ดังต่อไปนี้

$$\mathrm{dEBV} = \mathrm{PA} + \frac{\mathrm{EBV}-\mathrm{PA}}{\mathrm{R}^2_{\mathrm{dEBV}}} \quad \textbf{(7)}$$

โดยที่ $\mathrm{R}^2_{\mathrm{dEBV}}$ คือความเชื่อมั่นของค่า EBV ที่ไม่รวมการสนับสนุน (contribution) จากค่า EBV เฉลี่ยของพ่อแม่ (parent average, PA) โดยค่า dEBV จะถูกนำมาใช้เป็นเกณฑ์มาตรฐานในการตรวจสอบ จากนั้นเราคำนวณค่า GEBV (และ PA) ที่ 305 วัน สำหรับพ่อโคที่ทำการตรวจสอบแต่ละตัวจาก SS-RR-TDM (และ RR-TDM) โดยใช้ข้อมูลที่ลดลง ในการทดสอบความสามารถในการทำนายจะใช้เกณฑ์ 3 ข้อ

ดังต่อไปนี้ คือ ความแม่นยำเชิงทฤษฎี (theoretical accuracy, $\mathbf{r_t}$) ความแม่นยำเชิงการตรวจสอบ (validation accuracy, $\mathbf{r_v}$) และความไม่อคติของการทำนาย (unbiasedness of prediction, $\mathbf{b}$)

ความแม่นยำเชิงทฤษฎี ($\mathbf{r_t}$) เป็นค่าความแม่นยำของสัตว์แต่ละตัวของค่า GEBV (และ PA) ที่ 305 วัน ที่คำนวณได้จาก PEV (Misztal and Wiggans, 1988):

$$\mathbf{r_t} = \sqrt{1 - \frac{\mathbf{PEV}}{\hat{\sigma}^2_{a(305)}}} \quad \textbf{(8)}$$

เมื่อ $\sigma^2_{a(305)}$ เป็นค่าความแปรปรวนทางพันธุกรรมที่ประมาณได้จาก RR-TDM ตั้งแต่วันให้นมที่ 5 ถึง 305 $\mathbf{PEV}$ เป็นความแปรปรวนของความคลาดเคลื่อนจากการทำนายสำหรับสัตว์แต่ละตัว และแต่ละลักษณะซึ่งคำนวณได้จากค่ายกกำลังที่สองของค่าความเคลื่อนมาตรฐาน (SEP[2]) จากโปรแกรม AIREMLF90 (Misztal et al., 2018) โดยเฉพาะอย่างยิ่งสำหรับการวิเคราะห์ด้วยโมเดลการถดถอยแบบสุ่ม ความแปรปรวนของความคลาดเคลื่อนจากการทำนายสำหรับสัตว์แต่ละตัวที่ i และแต่ละลักษณะตามวิธีการด้านล่าง:

$$\mathbf{PEV_i} = \mathbf{z_t'C^{ii}}\,, \quad \textbf{(9)}$$

เมื่อ $\mathbf{z_t}$ เป็นเวคเตอร์ของ orthogonal covariates ที่เกี่ยวข้องกับฟังก์ชัน LPs และ $\mathbf{C^{ii}}$ เป็นเซตย่อยของการผกผันของ MME ที่สอดคล้องกับอิทธิพลของจีโนไทป์สำหรับสัตว์ i

ความแม่นยำเชิงการตรวจสอบ ($\mathbf{r_v}$) เป็นค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson correlation coefficient) ระหว่าง dEBV และ GEBV (หรือ PA) ที่ปรับด้วยรากที่สองของความเชื่อมั่นของ dEBV ($\sqrt{\mathbf{R^2_{dEBV}}}$) ดังแสดงตามสมการด้านล่าง

$$\mathbf{r_v} = \frac{\mathbf{corr(dEBV, GEBV\ or\ PA)}}{\sqrt{\mathbf{R^2_{dEBV}}}} \quad \textbf{(10)}$$

ความไม่อคติของการทำนาย ($\mathbf{b}$) เป็นค่าสัมประสิทธิ์การถดถอยที่คำนวณตามโมเดลการถดถอยเชิงเส้นดังนี้:

$$\mathbf{dEBV} = \mu + \mathbf{b(GEBV\ or\ PA)} + \mathbf{e} \quad \textbf{(11)}$$

โดยที่ μ คือจุดตัด (intercept) และ $\mathbf{e}$ คือความคลาดเคลื่อน ค่าสัมประสิทธิ์การถดถอยที่มีค่าต่ำกว่า และสูงกว่า 1 บ่งชี้ว่า GEBV (หรือ PA) มีค่าสูงเกินจริง (inflation) และต่ำเกินจริง (deflation) ตามลำดับ ทั้งนี้การวิเคราะห์ด้วยวิธีการแบบขั้นตอนเดียว (SS-RR-TDM) นั้น จะมีการดำเนินการจำนวนสองครั้งสำหรับแต่ละชุดข้อมูลที่มีและไม่มีจีโนไทป์ของแม่โค เพื่อเปรียบเทียบผลกระทบจากการมีส่วนร่วมของแม่โคต่อการทำนายค่าพันธุกรรมจีโนม

ผลการศึกษาและวิจารณ์

ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และค่าอัตราพันธุกรรม

ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และค่าอัตราพันธุกรรม (ค่าเฉลี่ยรายวัน และในรอบการให้นมที่ 305 วันจากการวิเคราะห์ทีละลักษณะด้วย RR-TDM) สำหรับลักษณะการให้ผลผลิตน้ำนมในรอบการให้นมแรกของประชากรโคนมไทยแสดงใน **Table 2** ค่าเฉลี่ยของ TD-MY, TD-FY, TD-PY, TD-TY, TD-FP, TD-PP และ TD-TP มีค่าเท่ากับ 13.89 ± 4.53 กก. 0.49 ± 0.20 กก. 0.44 ± 0.14 กก. 1.69 ± 0.55 กก. 3.56 ± 0.96 %, 3.13 ± 0.39 % และ 12.12 ± 1.10 % ตามลำดับ โดยทั่วไปชุดข้อมูลที่ลดลงจะมีค่าเฉลี่ยของแต่ละลักษณะใกล้เคียงกับชุดข้อมูลทั้งหมด ส่วนเบี่ยงเบนมาตรฐานของแต่ละลักษณะมีค่าความผันแปรเล็กน้อย ค่าอัตราพันธุกรรมโดยเฉลี่ยรายวันของทุกลักษณะอยู่ในระดับต่ำถึงปานกลาง โดยลักษณะปริมาณการให้ผลผลิตน้ำนม (0.21 - 0.36) ให้ค่าสูงกว่าองค์ประกอบน้ำนม (0.10 - 0.16) นอกจากนี้ค่าอัตราพันธุกรรมของลักษณะปริมาณการให้ผลผลิตน้ำนมในรอบการให้นมที่ 305 วันนั้นสูงกว่าค่าอัตราพันธุกรรมโดยเฉลี่ยต่อวันของลักษณะที่สอดคล้องกันโดยมีค่าอยู่ในช่วง 0.43 - 0.54

ค่าอัตราพันธุกรรมโดยเฉลี่ยต่อวันที่ประเมินได้ของปริมาณน้ำนม และโปรตีนสูงกว่าไขมันซึ่งสอดคล้องกับผลของ Ben Zaabza et al. (2018) โดยเฉพาะอย่างยิ่งค่าอัตราพันธุกรรมของปริมาณน้ำนมในการศึกษานี้ (0.35) ซึ่งมีค่าอยู่ในช่วง 0.17 ถึง 0.41 (ไม่แสดง) ใกล้เคียง

กับช่วงที่ได้รับจากโคนมเอธิโอเปีย (0.32: 0.17 - 0.42; Gebreyohannes et al., 2016) และโคนมไทย (0.36: 0.29 - 0.44; Buaban et al., 2020) ค่าอัตราพันธุกรรมโดยเฉลี่ยต่อวันนั้นมีค่าต่ำกว่าค่าอัตราพันธุกรรมที่ 305 วัน ซึ่งสาเหตุหนึ่งอาจเป็นผลมาจากการกำจัดความแปรปรวนของลักษณะที่แสดงออกภายในสัตว์โดยการรวบรวมข้อมูลในวันทดสอบหลายบันทึกไปเป็นข้อมูลที่ 305 วันบันทึกเดียว (Abdullahpour et al., 2013) องค์ประกอบของความผันแปรนี้สามารถเพิ่มความแปรปรวนของความคลาดเคลื่อนของโมเดลวันทดสอบได้อย่างมากหากอิทธิพลในโมเดลการวิเคราะห์ไม่ได้ถูกพิจารณา ด้วยเหตุผลนี้ในสถานการณ์ของภูมิอากาศที่หลากหลาย การเปลี่ยนแปลงด้านสิ่งแวดล้อม การจัดการและระบบการให้อาหาร เช่น ในประเทศไทย ลักษณะเกี่ยวกับไขมันในนมซึ่งสัตว์มีความไวสูงต่อปัจจัยเหล่านี้ โมเดลวันทดสอบจะให้ความแปรปรวนของความคลาดเคลื่อนมากกว่า และอัตราพันธุกรรมที่ต่ำกว่า ปัญหานี้อาจลดลงได้โดยการให้แหล่งข้อมูลใหม่ในโมเดลของการวิเคราะห์ เช่น คุณภาพอาหาร ซึ่งอาจอธิบายการเปลี่ยนแปลงภายในตัวสัตว์ หรือเกี่ยวข้องกับการจัดการกับฝูงสัตว์ที่คล้ายกันมากขึ้น (Abdullahpour et al., 2013) ความแตกต่างของค่าอัตราพันธุกรรมสำหรับลักษณะปริมาณการให้ผลผลิตทั้งหมดในการศึกษาต่างๆ อาจเป็นเพราะความแตกต่างของประชากร วิธีการวิเคราะห์ โมเดลที่แตกต่างกัน และอิทธิพลที่รวมอยู่ในโมเดล (Machado et al., 1999) นอกจากนี้ความแตกต่างอาจเกิดจากระบบการให้อาหาร และความเครียดจากความร้อนในประเทศไทยยังสามารถอธิบายค่าอัตราพันธุกรรมที่ตรงกันข้ามกันสำหรับปริมาณการให้ผลผลิต ฟาร์มโคนมของประเทศไทยมีการกระจายอยู่ทั่วทุกภูมิภาค แหล่งอาหารหยาบสำหรับโคนมประกอบด้วยผลพลอยได้ทางการเกษตร ข้าวโพด หญ้า และฟางข้าว ในภาคเหนือ ภาคกลาง ตะวันตก และตะวันออกตามลำดับ ส่วนใหญ่แล้วในประเทศไทยมีการให้อาหารแบบแยกส่วนมากที่สุด ภายใต้ระบบการให้อาหารนี้ เกษตรกรใช้อาหารข้นสูงแตกต่างกันตามปริมาณและคุณภาพของอาหารหยาบที่หาได้ง่าย นอกจากนี้ โคนมยังเลี้ยงในโรงเรือนเปิดซึ่งมีสภาพอากาศร้อนชื้นตลอดทั้งปี ดังนั้นโคนมจึงพยายามปรับตัวเพื่อการให้ผลผลิตภายใต้ความหลากหลายของธรรมชาติของประเทศไทย เราอาจคาดการณ์ได้ว่าโคนมในเขตร้อนแสดงศักยภาพทางพันธุกรรมที่แตกต่างกัน (Buaban et al., 2020)

Table 2 Mean, standard deviation (SD), average daily heritability estimates (h_d^2), and 305-d heritability estimates (h_{305d}^2) of milk production traits in the data set used for genetic evaluation

		Full data set				Reduced data set	
Traits[1]	Unit	Mean	SD	h_d^2	h_{305d}^2	Mean	SD
TD-MY	kg	13.89	4.53	0.35	0.49	13.35	4.44
TD-FY	kg	0.49	0.20	0.21	0.43	0.48	0.19
TD-PY	kg	0.44	0.14	0.36	0.54	0.42	0.15
TD-TY	kg	1.69	0.55	0.36	0.52	1.64	0.55
TD-FP	%	3.56	0.96	0.10	0.10	3.58	0.91
TD-PP	%	3.13	0.39	0.16	0.16	3.16	0.41
TD-TP	%	12.12	1.10	0.15	0.15	12.16	1.08

[1]TD-MY =Test day-milk yield; TD-FY=Test day-fat yield; TD-PY=Test day-protein yield; TD-TY=Test day-total solid yield; TD-FP=Test day-fat percentage; TD-PP=Test day-protein percentage; TD-TP=Test day-total solid percentage

การตรวจสอบค่าพันธุกรรมจีโนมสำหรับพ่อโคที่อายุยังน้อย

ความแม่นยำเชิงทฤษฎี

Table 3 แสดงค่าเฉลี่ยของความแม่นยำเชิงทฤษฎี (r_t) ของ PA จาก RR-TDM และ ค่าพันธุกรรมโดยการทำนายค่าพันธุกรรมจีโนม (GEBV) จาก SS-RR-TDM สำหรับประชากรที่ตรวจสอบความแม่นยำโดยใช้ลักษณะปริมาณการให้ผลผลิตน้ำนม และองค์ประกอบน้ำนมที่ 305 วัน ซึ่งคำนวณจากการทำนายค่าพันธุกรรมจีโนมด้วยวิธีการถดถอยแบบสุ่มที่พิจารณาข้อมูลจีโนไทป์จากพ่อโคอย่างเดียว และจากทั้งพ่อโคและแม่โค

Table 3 Averages (SD) of theoretical accuracies of parent average (PA) from the pedigree-based RR-TDM and genomic estimated breeding value (GEBV) from SS-RR-TDM for the validation bulls using 305-d milk production and milk components, which were calculated from random regression methods considering genotype of only bulls and bulls and cows in genomic evaluation

Traits[1/]	No. of validation bulls	PA	GEBV			
			Only Bulls		Bulls and Cows	
		r_t[2/]	r_t	Gain[3/]	r_t	Gain
305-d MY	28	0.51 (0.11)	0.71 (0.04)	0.20	0.73 (0.04)	0.22
305-d FY	23	0.48 (0.12)	0.70 (0.04)	0.22	0.72 (0.04)	0.24
305-d PY	28	0.51 (0.11)	0.71 (0.04)	0.20	0.73 (0.04)	0.22
305-d TY	28	0.50 (0.11)	0.71 (0.04)	0.20	0.73 (0.04)	0.23
average-FP	9	0.43 (0.15)	0.68 (0.05)	0.25	0.69 (0.05)	0.26
average-PP	25	0.48 (0.11)	0.70 (0.04)	0.22	0.72 (0.04)	0.24
average-TP	27	0.48 (0.11)	0.70 (0.04)	0.22	0.72 (0.04)	0.24
	Mean	0.48	0.70	0.22	0.72	0.24

[1/]305-d MY=305-d milk yield; 305-d FY=305-d fat yield; 305-d PY=305-d protein yield; 305-d TY=305-d total solid yield; average-FP=average fat percentage; average-PP=average protein percentage; average-TP=average total solid percentage

[2/]r_t= theoretical accuracy

[3/]Gain= r_t of GEBV - r_t of PA

ความแม่นยำเชิงทฤษฎีในการทำนายค่าพันธุกรรมจีโนมทุกลักษณะที่พิจารณาจีโนไทป์ทั้งจากพ่อโคอย่างเดียว และจากทั้งพ่อโคและแม่โคเปรียบเทียบกับ PA มีค่าเพิ่มขึ้น (gain) โดยเฉลี่ยเท่ากับ 0.22 และ 0.24 หรือประมาณร้อยละ 46 และ 50 ตามลำดับ โดยความแม่นยำเชิงทฤษฎีเพิ่มขึ้นโดยเฉลี่ย 0.02 จุด หรือประมาณร้อยละ 9.10 เมื่อทำนายค่าพันธุกรรมจีโนม (GEBV) โดยใช้ข้อมูลจีโนไทป์จากทั้งพ่อโคและแม่โคเปรียบเทียบกับสถานการณ์ที่มีพ่อโคเพียงอย่างเดียว

ความแม่นยำเชิงทฤษฎีแตกต่างกันไปตามค่าอัตราพันธุกรรมของแต่ละลักษณะ ลักษณะที่มีอัตราพันธุกรรมสูง (MY, PY และ TY) จะมีความแม่นยำเชิงทฤษฎีสูงขึ้นมากกว่าลักษณะที่มีค่าพันธุกรรมต่ำ (FP, PP และ TP) อย่างไรก็ตามลักษณะที่มีค่าอัตราพันธุกรรมต่ำมีแนวโน้มที่จะได้รับความแม่นยำเชิงทฤษฎีสำหรับการทำนายค่าพันธุกรรมจีโนมเพิ่มมากขึ้นเมื่อเปรียบเทียบกับ PA ผลของการเพิ่มความแม่นยำเชิงทฤษฎีสำหรับทุกลักษณะสอดคล้องกับการศึกษาอื่น ๆ (VanRaden et al., 2009; Ding et al., 2013; Jattawa et al., 2016;

Oliveira, 2019b) ซึ่งมีเหตุผลเกี่ยวข้องกับปริมาณข้อมูลที่มีอยู่ใน MME ดังนั้นการรวมเอาข้อมูลจีโนมซึ่งเป็นอีกแหล่งข้อมูลหนึ่งในการทำนายค่าพันธุกรรม GEBV ด้วย SS-RR-TDM ทำให้มีความแม่นยำเชิงทฤษฎีเพิ่มขึ้นเมื่อเปรียบเทียบกับ PA ที่ได้จาก RR-TDM โดยทั่วไป ความแม่นยำเชิงทฤษฎีของ GEBV ที่ประเมินได้ในการศึกษานี้อยู่ในระดับปานกลางถึงสูง (0.68 - 0.73) ซึ่งเกี่ยวข้องกับการลดลงของค่าอคติของการทำนาย (PEV) เนื่องจากการใช้วิธีการแบบขั้นตอนเดียว และเกณฑ์ที่ใช้ในการเลือกสัตว์รวมเข้าไปอยู่ในประชากรอ้างอิง และประชากรที่ตรวจสอบ ที่มีความแม่นยำสูง ผลของการศึกษานี้ยังแสดงให้เห็นถึงความสัมพันธ์ที่ดีระหว่างความแม่นยำเชิงทฤษฎี และค่าอัตราพันธุกรรม ลักษณะที่มีค่าพันธุกรรมต่ำจะมีค่า EBV หรือ GEBV ที่มีความความแม่นยำต่ำสอดคล้องกับ Luan at al. (2009) ในโคนมสีแดงในกลุ่มประเทศนอร์ดิก สำหรับ FP ที่มีค่าความแม่นยำเชิงทฤษฎีที่ต่ำกว่านั้นเนื่องจากค่าอัตราพันธุกรรมที่ต่ำกว่า และมีสัตว์จำนวนน้อยกว่าในเมตริกซ์ $\mathbf{G}$ เมื่อเปรียบเทียบกับลักษณะอื่น ๆ ซึ่งสะท้อนถึงข้อมูลในการทำนายมีจำนวนน้อย และเป็นที่ทราบกันดีว่าหนึ่งในปัจจัยที่มีผลต่อความแม่นยำของ EBV หรือ GEBV คือจำนวนข้อมูลฟีโนไทป์ของลักษณะ (Hayes et al., 2009) ในการศึกษาครั้งนี้จำนวนข้อมูลฟีโนไทป์ที่ใช้ในการทำนายค่าพันธุกรรมสำหรับ FP นั้นน้อยกว่าลักษณะอื่น ๆ จึงทำให้ความแม่นยำของ EBV หรือ GEBV สำหรับ FP ที่ได้ต่ำกว่าลักษณะอื่น ๆ

ความแม่นยำเชิงการตรวจสอบ

Table 4 แสดงค่าเฉลี่ยของความแม่นยำเชิงการตรวจสอบของค่าเฉลี่ยพ่อแม่ (PA) จาก RR-TDM และค่าพันธุกรรมโดยการทำนายค่าพันธุกรรมจีโนม (GEBV) จาก SS-RR-TDM สำหรับประชากรที่ตรวจสอบความแม่นยำโดยใช้ลักษณะปริมาณการให้ผลผลิตน้ำนมและองค์ประกอบน้ำนมที่ 305 วัน ซึ่งคำนวณจากการทำนายค่าพันธุกรรมจีโนมด้วยวิธีการถดถอยแบบสุ่มที่พิจารณาข้อมูลจีโนไทป์จากพ่อโคอย่างเดียว และจากทั้งพ่อโคและแม่โค โดยทั่วไปรากที่สองของความเชื่อมั่นเฉลี่ยของ dEBV ($\mathbf{r}_{\mathbf{dEBV}_{\mathbf{mean}}}$) ของพ่อโคที่ตรวจสอบสำหรับแต่ละลักษณะอยู่ในช่วง 0.84 ถึง 0.88 สำหรับลักษณะการให้ผลผลิตน้ำนม ซึ่งมีแนวโน้มที่สอดคล้องตามค่าอัตราพันธุกรรมสำหรับแต่ละลักษณะ (ลักษณะที่มีค่าอัตราพันธุกรรมที่สูงกว่านั้นจะมีค่าความแม่นยำสูงกว่า) ความแม่นยำเชิงการตรวจสอบที่ประเมินได้สำหรับ GEBV โดยใช้ SS-RR-TDM มีค่าสูงกว่าที่ประเมินโดย PA โดยใช้ RR-TDM ในทุกลักษณะที่ศึกษา ความแม่นยำเชิงการตรวจสอบของ PA มีค่าอยู่ในช่วง 0.06-0.34 และ GEBV มีค่าอยู่ในช่วง 0.12-0.39 และ 0.22-0.40 สำหรับการใช้ข้อมูลจีโนไทป์จากพ่อโคอย่างเดียว และจากทั้งพ่อโคและแม่โค ตามลำดับ ความแม่นยำเชิงการตรวจสอบที่เพิ่มสูงขึ้นเมื่อวิเคราะห์ด้วย SS-RR-TDM จากการใช้ข้อมูลจีโนไทป์จากพ่อโคอย่างเดียว และจากทั้งพ่อโคและแม่โค เมื่อเปรียบเทียบกับ RR-TDM นั้นมีค่าโดยเฉลี่ยเพิ่มขึ้น 0.06 และ 0.13 จุด หรือประมาณร้อยละ 33.30 และ 72.2 ตามลำดับ (**Table 4**) โดยความแม่นยำเชิงการตรวจสอบเพิ่มขึ้นโดยเฉลี่ย 0.07 จุด เมื่อทำนายค่า GEBV โดยใช้ข้อมูลจีโนไทป์จากทั้งพ่อโคและแม่โคเปรียบเทียบกับสถานการณ์ที่มีพ่อโคเพียงอย่างเดียว

Table 4 Validation realized accuracies of parent average (PA) from the pedigree-based RR-TDM and genomic estimated breeding value (GEBV) from SS-RR-TDM for the validation bulls using 305-d milk production and milk components, which were calculated from random regression methods considering genotype of only bulls and bulls and cows in genomic evaluation

Traits[1/]	No. of validation bulls	$r_{dEBV_{mean}}$[2/]	PA	GEBV			
				Only Bulls		Bulls and Cows	
			r_v[3/]	r_v	Gain[4/]	r_v	Gain
305-d MY	28	0.87	0.24	0.31	0.07	0.37	0.13
305-d FY	23	0.87	0.15	0.18	0.03	0.33	0.18
305-d PY	28	0.87	0.22	0.28	0.06	0.31	0.09
305-d TY	28	0.87	0.11	0.18	0.07	0.27	0.16
average-FP	9	0.88	0.06	0.12	0.06	0.22	0.16
average-PP	25	0.84	0.34	0.39	0.05	0.40	0.06
average-TP	27	0.85	0.15	0.25	0.10	0.25	0.10
	Mean	0.86	0.18	0.24	0.06	0.31	0.13

[1/]305-d MY=305-d milk yield; 305-d FY=305-d fat yield; 305-d PY=305-d protein yield; 305-d TY=305-d total solid yield; average-FP=average fat percentage; average-PP=average protein percentage; average-TP=average total solid percentage

[2/]$r_{dEBV_{mean}}$ = the average of the square root of the reliability of de-regressed proof for validation bulls (dEBV)

[3/]r_v = validation accuracy

[4/]Gain=r_v of GEBV-r_v of PA

การรวมเอาข้อมูลจีโนมในการทำนายค่าพันธุกรรมจีโนมด้วย SS-RR-TDM เพิ่มความแม่นยำเชิงการตรวจสอบของ GEBV สำหรับเกือบทุกลักษณะอาจเป็นเพราะ GEBV ยังคำนึงถึงเงื่อนไขการสุ่มตัวอย่าง Mendelian และเมตริกซ์ **H** เชื่อมโยงสัตว์ที่มีจีโนไทป์แต่ไม่มีฟีโนไทป์กับสัตว์ที่ไม่มีจีโนไทป์แต่มีฟีโนไทป์ ถ้ามีการเชื่อมโยงผ่านทางพันธุ์ประวัติ (Garcia et al, 2018) ดังนั้นจึงเป็นการประมาณค่าความดีเด่นทางพันธุกรรมของสัตว์ที่อายุยังน้อยซึ่งเป็นประโยชน์มากกว่าเมื่อเทียบกับ PA การค้นพบนี้สอดคล้องกับที่รายงานในการศึกษาอื่นๆ โดยใช้วิธีการแบบขั้นตอนเดียวกันเหมือนกันด้วยโมเดลในวันทดสอบการถดถอยแบบสุ่มหลายรอบการให้นมในการทำนายค่าพันธุกรรมจีโนมของโคนมสีแดงในกลุ่มประเทศนอร์ดิก (Koivula et al., 2015) และการทำนายค่าพันธุกรรมจีโนมแบบสองขั้นตอนในโคโฮลสไตน์ของแคนาดา (Oliveira et al., 2019b) แต่มีค่าความแม่นยำต่ำกว่า ปัจจัยสำคัญที่ทำให้ค่าความแม่นยำเชิงการตรวจสอบของการศึกษาครั้งนี้แตกต่างจากการศึกษาอื่นนั้น ได้แก่ การมีประชากรอ้างอิงขนาดเล็ก และองค์ประกอบในแง่ของแม่โคส่วนใหญ่ที่มีความแม่นยำของข้อมูลฟีโนไทป์ต่ำกว่าพ่อโคทดสอบในขบวนการทดสอบลูกหลานในประเทศที่พัฒนาแล้ว

ความไม่อคติของการทำนาย

Table 5 แสดงค่าสัมประสิทธิ์การถดถอยซึ่งเป็นตัวบ่งชี้วัดความไม่มีอคติของการทำนายค่าเฉลี่ยของพ่อแม่ (PA) โดยใช้ RR-TDM และค่าพันธุกรรมจีโนม (GEBV) โดยใช้ SS-RR-TDM สำหรับประชากรที่ตรวจสอบโดยใช้ลักษณะปริมาณการให้ผลผลิตน้ำนม และ

องค์ประกอบน้ำนมที่ 305 วัน ซึ่งคำนวณจากวิธีการถดถอยแบบสุ่มที่พิจารณาข้อมูลจีโนไทป์จากพ่อโคอย่างเดียว และจากทั้งพ่อโคและแม่โคในการทำนายค่าพันธุกรรมจีโนม

Table 5 Regression coefficients (b_1) and standard error (SE) of parent average (PA) from the pedigree-based RR-TDM and genomic estimated breeding value (GEBV) from SS-RR-TDM for the validation population using 305-d milk production and milk components, which were calculated from random regression methods considering genotype of only bulls and bulls and cows in genomic evaluation

Traits[1/]	No. of validation bulls	PA		GEBV			
				Only Bulls		Bulls and Cows	
		b_1	SE	b_1	SE	b_1	SE
305-d MY	28	0.23	0.21	0.45	0.31	0.47	0.27
305-d FY	23	0.19	0.31	0.33	0.45	0.58	0.42
305-d PY	28	0.18	0.18	0.32	0.26	0.35	0.24
305-d TY	28	0.21	0.21	0.43	0.31	0.45	0.27
average-FP	9	0.19	1.36	0.48	1.36	0.48	1.66
average-PP	25	0.54	0.38	0.67	0.39	0.75	0.43
average-TP	25	0.30	0.38	0.44	0.42	0.39	0.46
Mean		0.26		0.45		0.50	

[1/]305-d MY=305-d milk yield; 305-d FY=305-d fat yield; 305-d PY=305-d protein yield; 305-d TY=305-d total solid yield; average-FP=average fat percentage; average-PP=average protein percentage; average-TP=average total solid percentage

ตามหลักการแล้วค่าสัมประสิทธิ์การถดถอยใกล้เคียง 1 จะบ่งชี้ว่าในการทำนายความแตกต่างทางพันธุกรรมระหว่างสัตว์มีความไม่อคติมาก ค่าสัมประสิทธิ์การถดถอยของ PA สำหรับทุกลักษณะการให้ผลผลิตน้ำนม มีค่าอยู่ในช่วง 0.19 - 0.54 ในขณะที่ค่าสัมประสิทธิ์การถดถอยของ GEBV มีค่าอยู่ในช่วง 0.32 – 0.67 และ 0.35 – 0.75 สำหรับการใช้ข้อมูลจีโนไทป์จากพ่อโคอย่างเดียว และจากทั้งพ่อโคและแม่โคในการทำนายค่าพันธุกรรมจีโนม ตามลำดับ ในการศึกษาครั้งนี้สำหรับลักษณะการให้ผลผลิตน้ำนมมีค่าสัมประสิทธิ์การถดถอยของทั้ง GEBV และ PA ต่ำกว่า 1 ซึ่งบ่งชี้ว่าค่าทำนายที่ได้นั้นสูงเกินจริง ส่งผลให้การทำนายค่าพันธุกรรมของลูกโคนั้นได้เกินจริงเมื่อเทียบกับการแสดงออกของลูกโคเหล่านั้นในอนาคต แต่อย่างไรก็ตามค่าสัมประสิทธิ์การถดถอยที่ได้จาก SS-RR-TDM ของลักษณะการให้ผลผลิตน้ำนมโดยเฉลี่ยมีค่าเข้าใกล้ 1 มากกว่า RR-TDM จากผลการศึกษาบ่งชี้ว่าการเปลี่ยนแปลงการทำนายแบบดั้งเดิมโดยใช้พันธุ์ประวัติ (PA) มาเป็นการทำนายค่าพันธุกรรมจีโนม (GEBV) จะช่วยเพิ่มความแม่นยำในการประมาณค่าพันธุกรรมให้ดีขึ้น เมื่อเปรียบเทียบค่าสัมประสิทธิ์การถดถอยของ GEBV ที่ประมาณได้ในการศึกษานี้กับการศึกษาอื่น ๆ ที่ใช้วิธีเดียวกัน (Koivula et al., 2015; Baba et al., 2017) และที่มีการปรับขนาดเพื่อรวมเมตริกซ์ความสัมพันธ์จีโนม ($\mathbf{G}$) และเมตริกซ์ความสัมพันธ์ทางพันธุ์ประวัติ ($\mathbf{A}$) ซึ่งทำให้ความไม่อคติสูงขึ้นเช่นเดียวกัน (Christensen et al., 2012; Koivula et al., 2015) แต่การศึกษานี้มีจำนวนข้อมูลเนื่องจากพ่อโคที่ใช้ในการตรวจสอบความแม่นยำน้อย ค่าสัมประสิทธิ์การถดถอยของ GEBV ที่ประมาณได้จึงไม่ใกล้เคียงกับ 1 หากประชากรอ้างอิงที่มีขนาดใหญ่ขึ้นน่าจะส่งผลให้ความแม่นยำในการทำนายเพิ่มขึ้น และสามารถเพิ่มความไม่อคติได้ โดยรวมแล้ว SS-RR-TDM มีประสิทธิภาพเหนือกว่า RR-TDM ในแง่ความแม่นยำเชิงทฤษฎี ความแม่นยำเชิงการตรวจสอบ และความไม่อคติ (สัมประสิทธิ์การถดถอย, $\mathbf{b_1}$) โมเดล RR-TDM ใช้ข้อมูลพันธุ์ประวัติ ดังนั้นความแม่นยำที่ได้นั้นมาจากเพียงความสัมพันธ์ทางพันธุ์ประวัติ ($\mathbf{A}$) ระหว่างสัตว์แต่ละตัว โมเดล SS-RR-TDM ใช้ข้อมูลจีโนมร่วมด้วย ดังนั้นความแม่นยำที่ได้

เพิ่มขึ้นมาจากทั้งความสัมพันธ์ทางพันธุกรรม (**G**) ที่เพิ่มขึ้นระหว่างสัตว์แต่ละตัวและจากความเชื่อมโยงแบบไม่สมดุลของเครื่องหมายทางพันธุกรรม (linkage disequilibrium, LD) นอกจากนี้การใช้เมตริกซ์ **H** ใน SS-RR-TDM เทียบเท่ากับการกำหนดจีโนไทป์ที่หายไปของสัตว์แต่ละตัวที่มีฟีโนไทป์แต่ไม่มีข้อมูลจีโนไทป์ และทำให้ชุดข้อมูลอ้างอิงมีขนาดใหญ่ขึ้น ดังนั้น SS-RR-TDM จึงให้ความแม่นยำที่สูงขึ้น แสดงให้เห็นว่าการใช้ GEBV ที่ได้จากวิธีการแบบขั้นตอนเดียวด้วยโมเดลการถดถอยแบบสุ่มจึงเป็นค่าการทำนายที่เข้าใกล้ความเป็นจริงมากกว่าและมีประสิทธิภาพในการทำนายค่าพันธุกรรมจีโนมสำหรับทุกลักษณะในโคนมไทย

ผลกระทบจากการมีส่วนร่วมของแม่โคกับการทำนายค่าพันธุกรรมจีโนม

โดยรวมแล้วค่าสัมประสิทธิ์การถดถอยของลักษณะการให้ผลผลิตน้ำนมสำหรับการทำนายค่า GEBV ที่ใช้เฉพาะจีโนไทป์พ่อโคอย่างเดียว และจีโนไทป์ของทั้งพ่อโคและแม่โคมีค่าเฉลี่ยท่ากับ 0.45 และ 0.50 ตามลำดับ (Tables 5) และเมื่อพิจารณาค่าการทำนายที่ได้รับจาก SS-RR-TDM โดยการเพิ่มข้อมูลจีโนไทป์ของแม่โคแสดงให้เห็นว่ามีความแม่นยำเพิ่มขึ้นโดยเฉลี่ย 0.02 และ 0.07 จุด สำหรับความแม่นยำเชิงทฤษฎี และความแม่นยำเชิงการตรวจสอบ ตามลำดับ เมื่อเทียบกับการใช้ข้อมูลจีโนไทป์ของพ่อโคอย่างเดียว (Tables 3 และ 4) ชี้ให้เห็นว่าการใช้จีโนไทป์ของทั้งพ่อโคและแม่โคในการในการทำนายมีความไม่อคติมากกว่าการใช้จีโนไทป์พ่อโคอย่างเดียว การทำนาย GEBV ที่เกินจริงมีแนวโน้มลดลงเมื่อมีการใช้ข้อมูลจีโนไทป์ของแม่โคร่วมในการทำนาย ซึ่งอาจเกี่ยวข้องกับการเพิ่มจำนวนของข้อมูลที่มีอยู่สำหรับการวิเคราะห์ หรือเกี่ยวข้องกับค่า EBV ที่ได้รับจากแม่โคมีความน่าเชื่อถือ (Uemoto et al., 2017) ในประชากรอ้างอิงจำเป็นต้องมีค่า EBV ที่น่าเชื่อถือสำหรับใช้ในประมาณอิทธิพลของเครื่องหมายทางพันธุกรรม หรือการทำนายค่า GEBV ได้อย่างถูกต้อง ขณะที่ในประชากรที่ตรวจสอบจำเป็นต้องมีรายละเอียดของข้อมูลที่น่าเชื่อถือสำหรับการตรวจสอบ GEBV ได้อย่างถูกต้อง โดยทั่วไปแม่โค 3-10 ตัว ที่มีจีโนไทป์จะให้รายละเอียดของข้อมูลเท่ากันกับพ่อโคที่ผ่านการทดสอบแล้วเพียง 1 ตัว ขึ้นอยู่กับค่าอัตราพันธุกรรมของแต่ละลักษณะ (Boichard et al., 2015; Weller, 2016) การค้นพบที่ได้ในครั้งนี้สอดคล้องกับการศึกษาก่อนหน้านี้ที่แสดงให้เห็นว่าการเพิ่มแม่โคในประชากรอ้างอิงช่วยเพิ่มความแม่นยำเชิงทฤษฎี (Gao et al., 2015; Su et al., 2016; Oliveira et al., 2019b) และความความถูกเชิงตรวจสอบ (Tsuruta et al., 2013; Koivula et al., 2014; Cooper et al., 2015; Uemoto et al., 2017) อย่างมีนัยสำคัญ แต่มีงานวิจัยบางฉบับที่ให้ผลตรงกันข้าม (Baba et al., 2017; Olivera et al., 2019b) อย่างไรก็ตามค่าสัมประสิทธิ์การถดถอยที่ได้ในการศึกษานี้มีค่าน้อยกว่า 1 มาก ซึ่งแตกต่างจากการศึกษาอื่น ๆ ที่ใช้ SS-RR-TDM (Koivula et al., 2015; Baba et al., 2017) ทั้งนี้อาจเนื่องมาจากการมีพ่อพันธุ์อยู่อย่างจำกัดจึงไม่สามารถเก็บตัวอย่างของพ่อพันธุ์โดยสุ่มเพื่อตรวจจีโนไทป์(Mäntysaari et al., 2010) มีข้อมูลสัตว์ที่จีโนไทป์น้อย (Legarra and Reverter, 2018) และความเข้ากันไม่ได้ (incompatibility) ระหว่าง **G** และ $\mathbf{A}_{22}$ (ความแตกต่างโดยเฉลี่ยขององค์ประกอบระหว่างเมทริกซ์ทั้งสอง) (Misztal et al., 2017) ความเข้ากันไม่ได้อาจเกิดจาก $\mathbf{A}_{22}$ไม่สมบูรณ์เนื่องจากพันธุ์ประวัติที่ขาดหายไป (Misztal et al., 2017) แม้ว่าค่า GEBV ที่ทำนายได้จะมีค่าสูงเกินจริง แต่เรายังคงสามารถจัดอันดับสัตว์เพื่อคัดเลือกสัตว์ที่มีจีโนไทป์และอายุยังน้อยได้โดยพิจารณาจากค่า GEBV แม้ว่าเราจะไม่สามารถทำนายแนวโน้มทางพันธุกรรมในอนาคตของพ่อพันธุ์ได้อย่างแม่นยำ หรือเปรียบเทียบพ่อโคหนุ่มที่จะคัดเลือก (young candidate bulls) กับพ่อพันธุ์ที่ผ่านการพิสูจน์แล้วได้เนื่องจาก GEBV ที่มีค่าสูงเกินจริง แต่หากเพิ่มจำนวนสัตว์ที่ใช้ในการทดสอบความถูกต้องมากขึ้นจะทำให้การทำนาย GEBV ที่ได้มีค่าใกล้เคียงกับความเป็นจริงในอนาคต ปัจจุบันนี้การทำนายค่าพันธุกรรมจีโนมโดยทั่วไปสำหรับโคนมสายพันธุ์โฮลสไตน์ได้รวมเอาเฉพาะพ่อโคเพียงอย่างเดียวทั้งในประชากรอ้างอิงและประชากรที่ตรวจสอบ เนื่องจากมีความซับซ้อนและกังวลเกี่ยวกับอคติในการทำนายแม่โค การเพิ่มความแม่นยำในการรวมจีโนไทป์ของแม่โคในประชากรอ้างอิงอาจขึ้นอยู่กับวิธีการวิเคราะห์ การใช้โมเดลแบบหลายลักษณะพร้อมกันช่วยเพิ่มความแม่นยำของลักษณะที่มีจำนวนข้อมูลที่จำกัด (Guo et al. 2014; Bauer et al. 2015) นอกจากนี้การวิเคราะห์โดยใช้วิธีการแบบขั้นตอนเดียวจะเพิ่มความแม่นยำของค่า GEBV และเพิ่มความไม่อคติเมื่อเปรียบเทียบกับวิธีการทำนายอื่น ๆ (Uemoto et al., 2017)

การศึกษาครั้งนี้แสดงให้เห็นว่าโมเดล SS-RR-TDM มีความเป็นไปได้ที่จะนำไปใช้ในการทำนายค่าพันธุกรรมจีโนมสำหรับลักษณะการให้ผลผลิตน้ำนมในประเทศไทย ซึ่งเป็นวิธีที่รวมข้อมูลฟีโนไทป์เข้ากับข้อมูลจีโนไทป์โดยตรง ทำให้ได้ค่า GEBV ที่มีความแม่นยำสูงขึ้น เนื่องจากปัจจุบันจำนวนของพ่อโคที่มีจีโนไทป์มีค่อนข้างน้อยทำให้การสร้างเมตริกซ์ **G** ผกผันทำได้ง่าย แต่ในอนาคตเมื่อมีจำนวนของพ่อโคที่มีจีโนไทป์มากขึ้น การสร้างเมตริกซ์ **G** ผกผันอาจทำได้ยากขึ้น จึงมีการเสนออัลกอริทึมเพื่อแก้ปัญหา โดย Legarra and Ducrocq (2012) และ Liu et al. (2013) ซึ่งทำให้ โมเดล SS-RR-TDM ถูกนำไปใช้อย่างกว้างขวาง

สรุป

การทำนายค่าพันธุกรรมจีโนมด้วยการใช้ข้อมูลจีโนมร่วมกับข้อมูลฟีโนไทป์ และข้อมูลพันธุ์ประวัติโดยใช้วิธีการแบบขั้นตอนเดียวด้วยโมเดลวันทดสอบการถดถอยแบบสุ่ม (SS-RR-TDM) ให้ทั้งความแม่นยำเชิงทฤษฎีและความแม่นยำเชิงการตรวจสอบที่สูงกว่า และมีความไม่อคติที่สูงกว่าการทำนายค่าพันธุกรรมด้วยการใช้ข้อมูลฟีโนไทป์ และข้อมูลพันธุ์ประวัติโดยใช้โมเดลวันทดสอบการถดถอยแบบสุ่มแบบดั้งเดิม (RR-TDM) สำหรับการทำนายพ่อโคที่อายุยังน้อยทุกลักษณะการให้ผลผลิตน้ำนม การใช้จีโนไทป์พ่อโคเพียงอย่างเดียวในการทำนายค่าพันธุกรรม SS-RR-TDM ให้ค่าความแม่นยำเชิงทฤษฎี และความแม่นยำเชิงการตรวจสอบเพิ่มขึ้นโดยเฉลี่ย 0.22 และ 0.06 จุดตามลำดับ เมื่อเปรียบเทียบกับ RR-TDM ขณะที่การนำเอาข้อมูลจีโนไทป์ของแม่โคมาร่วมใช้ในการทำนายค่าพันธุกรรมจีโนมมีผลทำให้ความแม่นยำเชิงทฤษฎี และความแม่นยำเชิงการตรวจสอบเพิ่มขึ้นอีกโดยเฉลี่ย 0.02 และ 0.07 จุดตามลำดับ ค่าสัมประสิทธิ์การถดถอยที่ได้จาก SS-RR-TDM มีค่าเข้าใกล้ 1 มากกว่า RR-TDM และการทำนาย GEBV ที่เกินจริงมีแนวโน้มลดลงเมื่อมีการใช้ข้อมูลจีโนไทป์ของแม่โคร่วมในการทำนาย ดังนั้นการวิจัยครั้งนี้ช่วยยืนยันความเป็นไปได้ในการใช้ SS-RRTDM ในการทำนายค่าพันธุกรรมจีโนมสำหรับลักษณะการให้ผลผลิตน้ำนมของโคนมไทยในอนาคต และควรใช้ข้อมูลจีโนไทป์ของทั้งพ่อโคและแม่โคในการทำนายพันธุกรรมจีโนมในประชากรโคนมไทยซึ่งมีพ่อโคจำนวนน้อยที่ผ่านการทดสอบด้วยข้อมูลของลูกสาวแล้ว รวมทั้งเพิ่มจำนวนแม่โคที่มีจีโนไทป์เพื่อเพิ่มความแม่นยำของค่า GEBV สำหรับใช้ในการคัดเลือก

คำขอบคุณ

การศึกษาครั้งนี้เป็นส่วนหนึ่งของโครงการ “การประเมินพันธุกรรมจีโนมเพื่อการปรับปรุงพันธุกรรมของโคนมพันธุ์โฮลสไตน์ในประเทศไทย” ซึ่งเป็นความร่วมมือระหว่างสำนักเทคโนโลยีชีวภาพการผลิตปศุสัตว์ กรมปศุสัตว์ และมหาวิทยาลัยขอนแก่น โดยได้รับงบประมาณสนับสนุนส่วนหนึ่งจากสำนักงานพัฒนาการวิจัยการเกษตร ผู้วิจัยขอขอบคุณสำนักเทคโนโลยีชีวภาพการผลิตปศุสัตว์สำหรับการจัดหาข้อมูลจีโนไทป์และข้อมูลผลผลิตในวันทดสอบ

เอกสารอ้างอิง

กรมปศุสัตว์. 2556. สมุดพ่อพันธุ์กรมปศุสัตว์ 2556. สำนักเทคโนโลยีชีวภาพการผลิตปศุสัตว์ กรมปศุสัตว์ พิมพ์ที่ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย 69 หน้า. แหล่งข้อมูล: http://biotech.dld.go.th/webnew/Data/Sire_Summary/2556/Sire_Summary2556.pdf. ค้นเมื่อ 15 ธันวาคม 2020.

กรมปศุสัตว์. 2562. สมุดพ่อพันธุ์กรมปศุสัตว์ 2562. สำนักเทคโนโลยีชีวภาพการผลิตปศุสัตว์ กรมปศุสัตว์ พิมพ์ที่ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย 69 หน้า. แหล่งข้อมูล: http://biotech.dld.go.th/webnew/Data/Sire_Summary/2562/Sire_Summary2562.pdf. ค้นเมื่อ 15 ธันวาคม 2020.

Abdullahpour, R., M. M. Shahrbabak, A. Nejati-Javaremi, R. V. Torshizi, and R. Mrode. 2013. Genetic analysis of milk yield, fat and protein content in Holstein dairy cows in Iran: Legendre polynomials random regression model applied. Archiv fur Tierzucht. 56: 497–508.

Baba, T., Y. Gotoh, S. Yamaguchi, S. Nakagawa, H. Abe, Y. Masuda, and T. Kawahara. 2017. Application of single-step genomic best linear unbiased prediction with a multiple-lactation random regression test-day model for Japanese Holsteins. Animal Science Journal. 88: 1226–1231.

Bauer, J., J. Pribyl, and L. Vostrý. 2015. Short communication: Reliability of single-step genomic BLUP breeding values by multi-trait test-day model analysis. Journal of Dairy Science. 98: 4999–5003.

Ben Zaabza, H., A. Ben Gara, and B. Rekik. 2018. Genetic analysis of milk production traits of Tunisian Holsteins using random regression test-day model with Legendre polynomials. Asian-Australasian Journal of Animal Sciences. 31: 636–642.

Boichard, D., V. Ducrocq, and S. Fritz. 2015. Sustainable dairy cattle selection in the genomic era. Invited review. Journal of Animal Breeding and Genetics. 132: 135–143.

Buaban, S. and J. Sanpote. 2010. Comparison of breeding values predicted with lactation model and test-day model in crossbred dairy cattle. Journal of Biotecnology Livestock Production. 5: 33–59.

Buaban, S., S. Puangdee, M. Duangjinda, and W. Boonkum. 2020. Estimation of genetic parameters and trends for production traits of dairy cattle in Thailand using a multiple-trait multiple-lactation test day model. Asian-Australas Journal of Animal Science. 33: 1387–1399.

Christensen, O. F., P. Madsen, B. Nielsen, T. Ostersen, and G. Su. 2012. Single-step methods for genomic evaluation in pigs. Animal. 6: 1565–1571.

Cooper, T. A., G. R. Wiggans, and P. M. VanRaden. 2015. Short communication: Analysis of genomic predictor population for Holstein dairy cattle in the United States-Effects of sex and age. Journal of Dairy Science. 98: 2785–2788.

Ding, X., Z. Zhang, X. Li, S. Wang, X. Wu, and D. Sun. 2013. Accuracy of genomic prediction for milk production traits in the Chinese Holstein Population using reference population consisting of cows. Journal of Dairy Science. 96: 5315–5323.

Doublet, A.-C., P. Croiseau, S. Fritz, A. Michenet, C. Hozé, C. Danchin-Burge, D. Laloë, and G. Restoux. 2019. The impact of genomic selection on genetic diversity and genetic gain in three french dairy cattle breeds. Genetics Selection Evolution. 51: 52.

Gao, H., P. Madsen, U. Nielsen, G. Aamand, G. Su, K. Byskov, and J. Jensen. 2015. Including different groups of genotyped females for genomic prediction in a Nordic Jersey population. Journal of Dairy Science. 98: 9051–9059.

Garcia, A. L. S., B. Bosworth, G. Waldbieser, I. Misztal, S. Tsuruta, and D. A. L. Lourenco. 2018. Development of genomic predictions for harvest and carcass weight in channel catfish. Genetics Selection Evolution. 50: 66.

Gebreyohannes, G., S. Koonawootrittriron, M. A. Elzo, T. Suwanasopee. 2016. Estimation of genetic parameters using a random regression monthly test-day model in an Ethiopian dairy cattle population. Agriculture and Natural Resources. 50: 64–70.

Gengler, N., A. Tijani, G. R. Wiggans, and I. Misztal. 1999. Estimation of (co)variances function coefficient for test-day yield with expectation-maximization restricted maximum likelihood algorithm. Journal of Dairy Science. 82: 1849.e1–1849.e23.

Hayes, B. J., P. J. Bowman, A. Chamberlain, and M. Goddard. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. Journal of Dairy Science. 92: 433–443.

Henderson, C.R. 1984. Applications of linear models in animal breeding, University of Guelph.

Jakobsen, J. H., P. Madsen, J. Jensen, J. Pedersen, L. Christensen, and D. Sorensen. 2002. Genetic parameters for milk production and persistency for danish Holsteins estimated in random regression models using REML. Journal of Dairy Science. 85: 1607–1616.

Jamrozik, J., L. Schaeffer, and J. Dekkers. 1997. Genetic evaluation of dairy cattle using test day yields and random regression model. Journal of Dairy Science. 80: 1217–1226.

Jattawa, D., M. A. Elzo, S. Koonawootrittriron, and T. Suwanasopee. 2016. Genomic-polygenic and polygenic evaluations for milk yield and fat percentage using random regression models with Legendre polynomials in a Thai multibreed dairy population. Livestock Science. 188: 133–141.

Kang, H., L. Zhou, R. Mrode, Q. Zhang, and J. F. Liu. 2017. Incorporating the single-step strategy into a random regression model to enhance genomic prediction of longitudinal traits. Heredity. 119: 459–467.

Kistemaker, G. J. 1997. The comparison of random regression test day models and a 305-day for evaluation of milk yield in dairy cattle. PhD Thesis. University of Guelph, Guelph, ON, Canada.

Koivula, M., I. Strandén, G. P. Aamand, and E. A. Mäntysaari. 2014. Effect of cow reference group on validation accuracy of genomic evaluation. p.1-3. In: Proceedings of 10th world congress of genetics applied to livestock production. August 17-22, 2014, Vancouver, Canada.

Koivula, M., I. Strandén, J. Pösö, G. Aamand, and E. Mäntysaari. 2015. Single-step genomic evaluation using multitrait random regression model and test-day data. Journal of Dairy Science. 98: 2775–2784.

Legarra, A., and A. Reverter. 2018. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. Genetics Selection Evolution. 50: 53.

Legarra, A., and V. Ducrocq. 2012. Computational strategies for national integration of phenotypic, genomic, and pedigree data in a single-step best linear unbiased prediction. Journal of Animal Science. 95: 4629–4645.

Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. Journal of Dairy Science. 92: 4656–4663.

Liu, R., Y. Sun, G. Zhao, F. Wang, D. Wu, M. Zheng, J. Chen, L. Zhang, Y. Hu, and J. Wen. 2013. Genome-wide association study identifies loci and candidate genes for body composition and meat quality traits in Beijing-You chickens. PLoS ONE. 8: e61172.

Luan, T., J. A. Woolliams, S., M. Kent, M. Svendsen, and T. H. Meuwissen. 2009. The accuracy of Genomic Selection in Norwegian red cattle assessed by cross-validation. Genetics. 183: 1119–1126.

Machado, S. G., M. A. R. Freitas, and C. H. Gadini. 1999. Genetic parameters of test day milk yields of Holstein cows. Genetics and Molecular Biology. 22: 383–386.

Mäntysaari, E., Z. Liu, and P. VanRaden. 2010. Interbull validation test for genomic evaluations. p.17-22. In: Proceedings of Interbull bulletin. March 4-5, 2010, Paris, France.

Misztal, I. and G. R. Wiggans. 1988. Approximation of Prediction Error Variance in Large-Scale Animal Models. Journal of Dairy Science. 71: 27–32.

Misztal, I., A. Legarra, and I. Aguilar. 2009. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. Journal of Dairy Science. 92: 4648–4655.

Misztal, I., D. Lourenco, and A. Legarra. 2020. Current status of genomic evaluation. Journal of Animal Science. 98: skaa101.

Misztal, I., H. L. Bradford, D.A.L. Lourenco, S. Tsuruta, Y. Masuda, A. Legarra, and T.J. Lawlor. 2017. Studies on inflation of GEBV in single-step GBLUP for type. Interbull bulletin. 51: 38–42.

Misztal, I., S. Tsuruta, D. Lourenco, Y. Masuda, I. Aguilar, A. Legarra, and Z. Vitezica. 2018. Manual for blupf90 family of programs. University of Georgia. Available: http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all2.pdf. Accessed February 1, 2021.

Mrode, R. A. 2014. Linear models for the prediction of animal breeding values, 3rd edition. Wallingford, United Kingdom. CAB International.

Oliveira, H. R., L. F. Brito, D. A. L. Lourenco, F. F. Silva, J. Jamrozik, L. R. Schaeffer, and F. S. Schenkel. 2019a. Invited review: Advances and applications of random regression models: From quantitative genetics to genomics. Journal of Dairy Science. 102: 7664–7683.

Oliveira, H. R., L. F. Brito, F. F. Silva, D. A. L. Lourenco, J. Jamrozik, and F. S. Schenkel. 2019b. Genomic prediction of lactation curves for milk, fat, protein, and somatic cell score in Holstein cattle. Journal of Dairy Science. 102: 452–463.

Seegers, H., C. Fourichon, and F. Beaudeau. 2003. Production effects related to mastitis and mastitis economics in dairy cattle herds. Veterinary Research. 34: 475–491.

Silva, M. V. B., D. J. A. dos Santos, S. A. Boison, A. T. H. Utsunomiya, A. S. Carmo, T. S. Sonstegard, J. B. Cole, and C. P. Van Tassell. 2014. The development of genomics applied to dairy breeding. Livestock Science. 166: 66–75.

Snelling, W. M., R. A. Cushman, J. W. Keele, C. Maltecca, M. G. Thomas, and M. R. S Fortes, A. Reverter. 2013. Breeding and genetics symposium: networks and pathways to guide genomic selection. Journal of Animal Science. 91: 537–552.

Su, G. M. P., U. S. Nielsen, G. P. Aamand, G. Wiggans, and B. Guldbrandtsen. 2016. Sharing referenced data and including cows in the reference population improve genomic predictions in Danish Jersey. Animal. 10: 1067–1075.

Tsuruta, S., I. Misztal, and T. J. Lawlor. 2013. Short communication: Genomic evaluations of final score for US Holsteins benefit from the inclusion of genotypes on cows. Journal of Dairy Science. 96: 3332–3335.

Uemoto, Y., T. Osawa, and J. Saburi. 2017. Effect of genotyped cows in the reference population on the genomic evaluation of Holstein cattle. Animal. 11(3): 382–393.

VanRaden, P. 2020. Symposium review: How to implement genomic selection. Journal of Dairy Science. 103: 5291–5301.

VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. Journal of Dairy Science. 91: 4414–4423.

VanRaden, P. M., C. P. Van Tassel, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F.S. Schenkel. 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. Journal of Dairy Science. 92: 16–24.

Weller, J. I. 2016. Genomic Selection in Animals. John Wiley & Sons Inc., Hoboken, NJ.