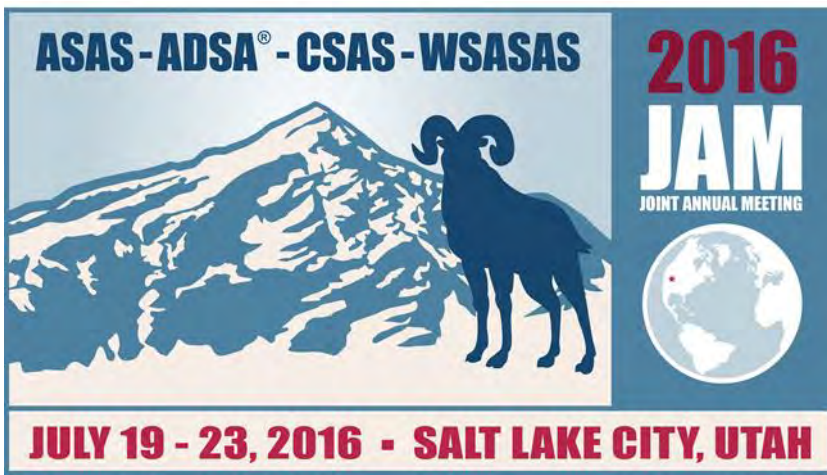




# 13707: Genomic and polygenic evaluations for milk and fat yields in Holstein upgraded Thai dairy cattle

Danai Jattawa<sup>1</sup>, Mauricio A. Elzo<sup>2</sup>, Skorn Koonawootrittriron<sup>1</sup>, and Thanathip Suwanasopee<sup>1</sup>

Kasetsart University, Chatuchak, Bangkok 10900, Thailand<sup>1</sup>

University of Florida, Gainesville, FL 32611-0910, USA<sup>2</sup>



## ABSTRACT

The objectives of this study were to compare variance components, genetic parameters, prediction accuracies, and ranking of animals for 305-day milk yield (MY) and 305-day fat yield (FY) in a Holstein upgraded Thai dairy population using two genomic models and a polygenic model (**PM**). One genomic model utilized 7,656 SNP (**GM7K**) and the other one used 74,144 actual and imputed 80K SNP (**GM80K**). Phenotypic and pedigree data were from 8,361 first-lactation cows located in 810 farms that had their first calving between 1989 and 2014. Variance components and genetic parameters were estimated using REML procedures. Fixed effects were contemporary group (herd-year-season), calving age and heterosis. Random effects were animal additive genetic and residual. Estimates of variance components and heritabilities for MY and FY from **GM80K** were the highest, followed by those from **GM7K**, and **PM** had the lowest values. Correlations estimates between MY and FY were similar across models. The **GM80K** yielded higher prediction accuracies (38.8% for MY and 31.9% for FY) than **GM7K** (36.7% for MY, and 31.4% for FY) and **PM** (31.5% for MY, and 24.4% for FY). Different MY and FY EBV rankings existed across models. The highest rank correlations were between **GM80K** and **GM7K** (0.90 for MY, and 0.91 for FY), followed by those between **GM80K** and **PM** (0.84 for MY, and 0.83 for FY), and the lowest ones were between GM7K and PM (0.80 for MY, and 0.80 for FY). *Animal rankings from GM80K should be preferred because its EBV had higher accuracy than EBV from GM7K and PM. Faster selection responses for MY and FY would be expected from GM80K than from GM7K and PM in this Holstein upgraded population.*

## INTRODUCTION

The availability of thousands of genotypes across the genome has provided valuable information for the characterization and evaluation of livestock animals. Genomic evaluations that utilize pedigree, phenotypes, and genotypes have increased accuracies of prediction and rates of genetic progress in animal breeding programs. Currently, genomic evaluation is widely utilized in the livestock industry, especially in dairy cattle. Conversely, dairy cattle in Thailand are genetically evaluated using only pedigree and phenotypic information. To improve the accuracy of genetic evaluation and speed up selection response for milk yield, fat yield, and other dairy traits, a national project for “the development of a national genomic evaluation system in Thailand” was started in 2012. To develop a national genomic evaluation system, variance components, genetic parameters, prediction accuracies, and ranking of animals need to be compared among genomic and conventional polygenic models using Thai dairy information. *Thus, the objectives of this study were to compare variance components, genetic parameters, prediction accuracies, and ranking of animals for 305-day milk yield and 305-day fat yield in the Holstein upgraded Thai dairy population using two genomic models with different sets of SNP genotypes and a polygenic model.*

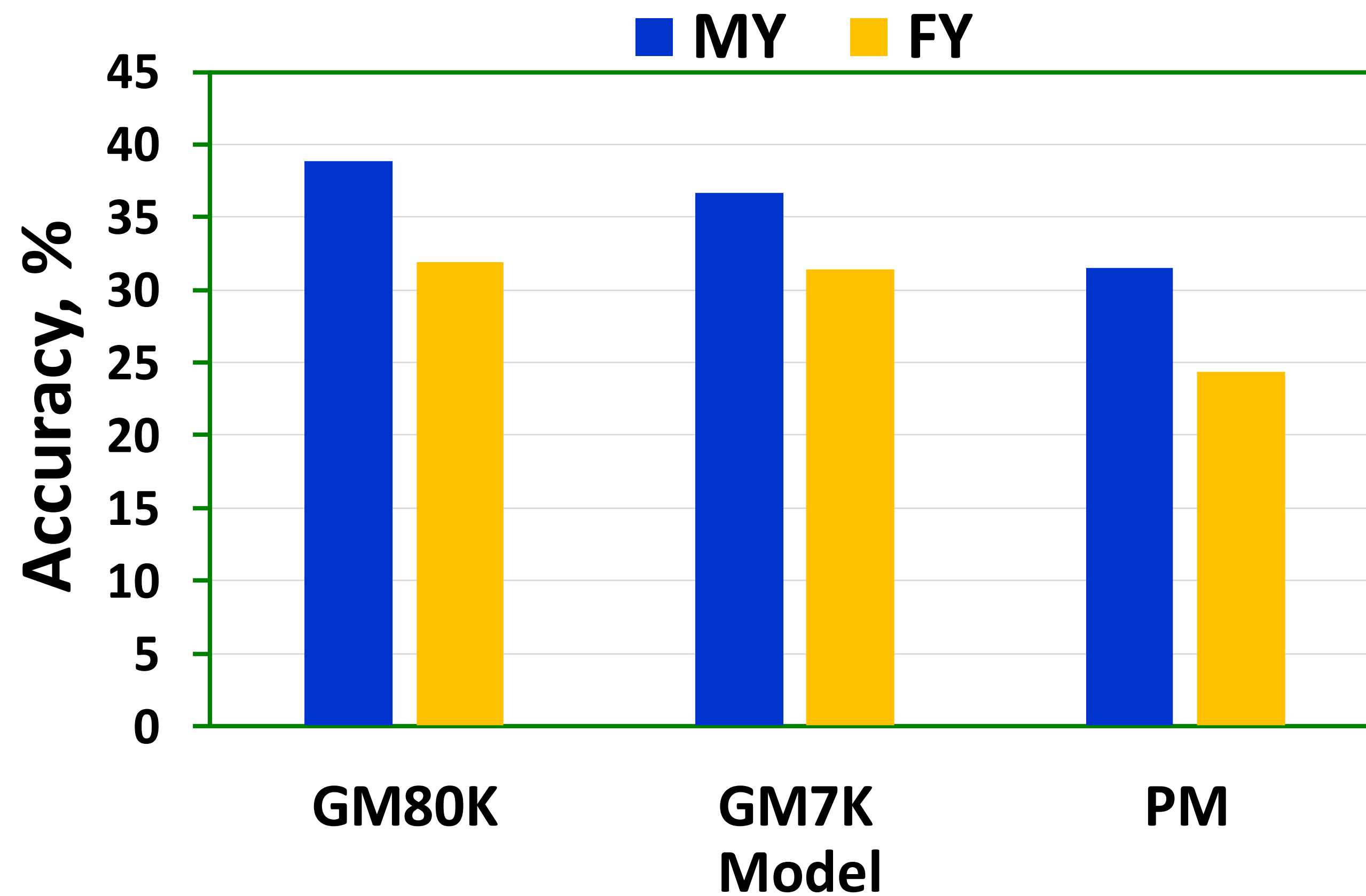
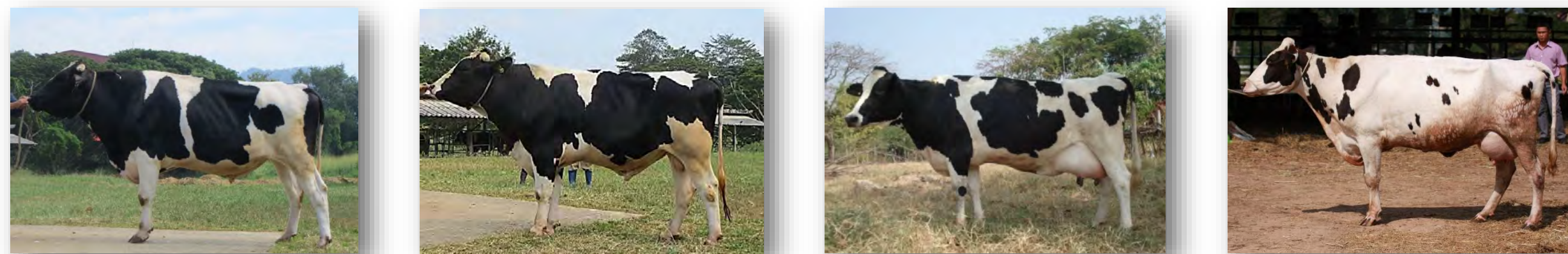


Fig. 1. Accuracy of estimated breeding values for 305-d milk yield and 305-d fat using a genomic model with actual and imputed 80K SNP markers (GM80K), a genomic model with actual 7K SNP markers (GM7K), and a polygenic model (PM)

## MATERIALS AND METHODS

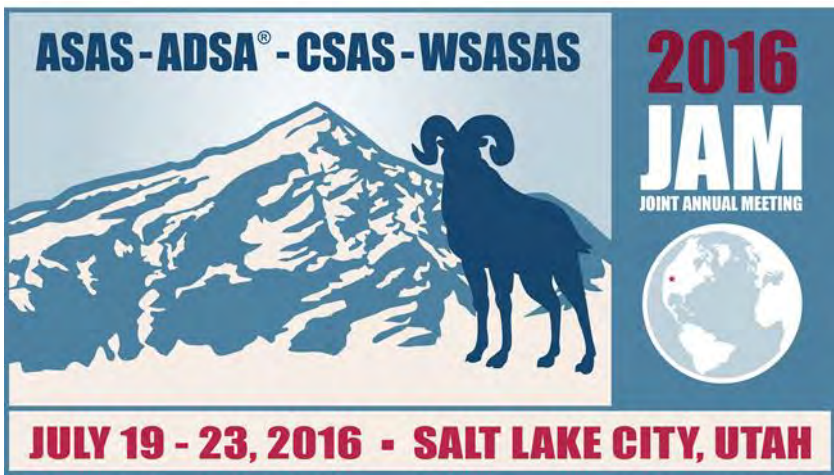
**Animals and traits:** Animals used in this research were 8,361 first-lactation cows from 810 farms located across regions in Thailand. These cows were the progeny of 1,210 sires and 6,992 dams, and had their first calving between 1989 and 2014. Traits were 305-day milk yield (MY) and 305-day fat yield (FY).

**Genotypic data:** Blood or semen samples were collected from 2,661 animals (89 sires and 2,572 cows). Genotypes were obtained at GeneSeek (Lincoln, NE, USA) with four GeneSeek Genomic Profiler (GGP) chips: 9K (n = 1,412), 20K (n = 570), 26K (n = 540), and 80K (n = 139) SNP marker chips. The actual SNP markers obtained from these chips were used to construct SNP set 1 and SNP set 2. The SNP markers in common among the GGP9K, GGP20K, GGP26K, and GGP80K chips were used to represent SNP set 1. Construction of SNP set 2 was obtained by imputation from GGP9K, GGP20K, and GGP26K to GGP80K using Flmpuete 2.2 (Sargolzaei et al., 2014). After SNP quality checks (MAF = 0.04; Call Rate > 0.9), 7,656 and 74,144 were kept to represent the genotypic information for SNP sets 1 and 2.

**Variance and covariance components:** Estimates of variance and covariance components for MY and FY were obtained using a bivariate polygenic model (**PM**) and two bivariate single-step genomic models (Aguilar et al., 2010), namely: 1) **GM7K**, a model that used pedigree, phenotypes, and SNP set 1 genotypes, and 2) **GM80K**, a model that used pedigree, phenotypes, and SNP set 2 genotypes. Fixed effects for the polygenic model and the two genomic models included contemporary group (herd-year-season), calving age, and heterosis. Random effects were animal additive genetic and residual. The BLUPF90 family of programs (Misztal et al., 2002) was utilized to estimate variance components and genetic parameters for MY and FY. Variance components were estimated using an average information restricted maximum likelihood algorithm using AIREMLF90 (Tsuruta, 2014).

**Prediction accuracies and animal rankings:** Animal EBV for MY and FY were computed for the polygenic and the two genomic models. Prediction accuracies for each EBV was obtained as  $\sqrt{1 - \frac{PEV}{\sigma_a^2}} * 100$ , where PEV was the prediction error variance and  $\sigma_a^2$  was the additive genetic variance. Rankings of animal EBV for MY and FY were compared using Spearman's rank correlations for each trait for all animals in the population.





# 13707: Genomic and polygenic evaluations for milk and fat yields in Holstein upgraded Thai dairy cattle

Danai Jattawa<sup>1</sup>, Mauricio A. Elzo<sup>2</sup>, Skorn Koonawootrittriron<sup>1</sup>, and Thanathip Suwanasopee<sup>1</sup>

Kasetsart University, Chatuchak, Bangkok 10900, Thailand<sup>1</sup>

University of Florida, Gainesville, FL 32611-0910, USA<sup>2</sup>



## RESULTS AND DISCUSSION

**Variance components and genetic parameters:** Estimates of variances and covariances for MY and FY from the two genomic models and the polygenic model are shown in **Table 1**. Estimates of phenotypic variances and covariances were similar across models. However, estimates of additive genetic variances and covariances for MY and FY were larger for **GM80K** (46%), and **GM7K** (12%) than corresponding values from **PM**. Conversely, estimates of environmental variances and covariances for the two genomic models were lower (7% for **GM80K** and 2% for **GM7K**) than those from the polygenic model. *These results indicated that inclusion of genotypes in addition to pedigree and phenotypes in genomic models accounted for substantially larger amounts of additive genetic variation than by using only pedigree and phenotypic information in the polygenic model.* Similarly, additive genetic variances and covariances were 29% larger for **GM80K** than for **GM7K**, whereas environmental variances and covariances were 6% lower for **GM80K** than for **GM7K**. This indicated that the additional SNP used by **GM80K** (74,144 SNP) explained nearly 30% more additive genetic variation for MY and FY than that accounted for by the 7,656 SNP in **GM7K**.

**Table 2** shows estimates of heritabilities and correlations for MY and FY obtained with the polygenic model and the two genomic models. Heritabilities from **GM80K** were the highest (MY = 0.26; FY = 0.18) and those from **PM** were the lowest (MY = 0.15; FY = 0.14). Heritabilities estimates for MY tended to increase with the number of SNP included in the model (from 0.19 for **GM7K** to 0.26 for **GM80K**). *This indicated that genomic models likely accounted for additive genetic relationships among animals in the Thai population more accurately resulting in higher additive genetic variances and heritabilities than those from the polygenic model.*

**Prediction accuracies:** **Fig. 1** shows the EBV accuracies for MY and FY computed from the polygenic model and the two genomic models. The **GM80K** had the highest EBV accuracies (38.8% for MY and 31.9% for FY), **GM7K** was second (36.7% for MY, and 31.4% for FY), and the least accurate was **PM** (31.5% for MY, and 24.4% for FY). *Gains in EBV accuracies were 7.2% between the **PM** and **GM80K** and 5.2% between the **PM** and **GM7K**. The increase in EBV accuracy between the low density **GM7K** and the high density **GM80K** was only 2%. This indicated that a strategy to keep genotyping costs reasonably low in this population would be to genotype animals that are highly connected in the population (i.e., most sires and some dams) with high density chips, and use low density chips with the rest of the population.*

**Table 1. Variance and covariance components for 305-d milk yield (MY) and 305-d fat yield (FY) estimated using two genomic models and a polygenic model**

| Variance component            | Model   |        |         |        |         |        |
|-------------------------------|---------|--------|---------|--------|---------|--------|
|                               | GM80K   | SE     | GM7K    | SE     | PM      | SE     |
| Additive genetic              |         |        |         |        |         |        |
| Var (MY), kg <sup>2</sup>     | 172,190 | 30,558 | 126,480 | 24,023 | 100,030 | 25,447 |
| Cov (MY, FY), kg <sup>2</sup> | 4,148   | 1,176  | 3,264   | 950    | 3,057   | 1,001  |
| Var (FY), kg <sup>2</sup>     | 230     | 65     | 184     | 54     | 177     | 55     |
| Environmental                 |         |        |         |        |         |        |
| Var (MY), kg <sup>2</sup>     | 500,420 | 27,936 | 542,750 | 22,912 | 565,990 | 25,164 |
| Cov (MY, FY), kg <sup>2</sup> | 16,125  | 1,117  | 16,955  | 942    | 17,099  | 1,011  |
| Var (FY), kg <sup>2</sup>     | 1,039   | 69     | 1,083   | 54     | 1,085   | 57     |
| Phenotypic                    |         |        |         |        |         |        |
| Var (MY), kg <sup>2</sup>     | 672,610 | 14,943 | 669,230 | 14,749 | 666,030 | 14,472 |
| Cov (MY, FY), kg <sup>2</sup> | 20,273  | 615    | 20,219  | 611    | 20,156  | 604    |
| Var (FY), kg <sup>2</sup>     | 1,269   | 35     | 1,266   | 35     | 1,262   | 34     |

**Table 2. Heritabilities and correlations for 305-d milk yield (MY) and 305-d fat yield (FY) estimated with two genomic models and a polygenic model**

| Parameter                          | Model |      |      |      |      |      |
|------------------------------------|-------|------|------|------|------|------|
|                                    | GM80K | SD   | GM7K | SD   | PM   | SD   |
| Heritability (MY)                  | 0.26  | 0.04 | 0.19 | 0.03 | 0.15 | 0.04 |
| Heritability (FY)                  | 0.18  | 0.05 | 0.15 | 0.04 | 0.14 | 0.04 |
| Genetic correlation (MY, FY)       | 0.66  | 0.11 | 0.68 | 0.11 | 0.73 | 0.13 |
| Environmental correlation (MY, FY) | 0.71  | 0.02 | 0.7  | 0.02 | 0.69 | 0.02 |
| Phenotypic correlation (MY, FY)    | 0.69  | 0.01 | 0.69 | 0.01 | 0.7  | 0.01 |

**Ranking of animals:** Spearman rank correlations between rankings of all animal EBV from the polygenic and the two genomic models are shown in **Table 3**. Rankings between animal EBV from **GM80K** and **GM7K** had the highest correlations (0.90 for MY, and 0.91 for FY). Rank correlations between animal EBV from **GM80K** and **PM** (0.84 for MY, and 0.83 for FY) were second, and the lowest rank correlations were between animal EBV from **GM7K** and **PM** (0.80 for MY, and 0.80 for FY). *Animal rankings from GM80K should be preferred because its EBV had higher accuracy than EBV from GM7K and PM. Faster selection responses for MY and FY would be expected from GM80K than from GM7K and PM in this Holstein upgraded population.*

**Table 3. Rank correlations between animal EBV for 305-d milk yield (MY) and fat yield (FY) evaluated using two genomic models and a polygenic model**

| Trait | Rank correlations |           |          |
|-------|-------------------|-----------|----------|
|       | GM80K, GM7K       | GM80K, PM | GM7K, PM |
| MY    | 0.90              | 0.84      | 0.80     |
| FY    | 0.91              | 0.83      | 0.80     |

## FINAL REMARKS

- ❖ *Estimates of additive genetic variances, heritabilities, and prediction accuracies for MY and FY from genomic models were higher than those from the polygenic model*
- ❖ *Additive genetic variances, heritabilities, and prediction accuracies tended to increase as the number of SNP increased*
- ❖ *Animal rankings from GM80K should be preferred to speed up selection response in the Thai Holstein upgraded population*

## REFERENCES

Aguilar et al., 2010. J. Dairy Sci. 93, 743-752; Goddard, 2009. Genetica 136, 245-257; Misztal, et al., 2002. Proc.7th World Congr. Genet. Appl. Livest. Prod., Montpellier, France; Mulder, et al., 2012. J. Dairy Sci. 95, 876-889; Sargolzaei, et al., 2014. BMC Genomics, 15, 478; Tsuruta, S., 2014. [http://nce.ads.uga.edu/wiki/doku.php?id=application\\_programs](http://nce.ads.uga.edu/wiki/doku.php?id=application_programs); VanRaden, et al., 2011. Genet. Sel. Evol. 43, 10; VanRaden, et al, 2009. J. Dairy Sci. 92, 16-24.