

## GENETIC VARIATION AND ESTIMATING BREEDING VALUES FOR SMALL-HOLDER CROSSBRED DAIRY CATTLE IN INDIA

M. Al kalaldehy<sup>1</sup>, Y. Gaundare<sup>2</sup>, M. Swaminathan<sup>2</sup>, S. Joshi<sup>2</sup>, H. Aliloo<sup>1</sup>, E.M. Strucken<sup>1</sup>, V. Ducrocq<sup>3</sup> and J.P. Gibson<sup>1</sup>

<sup>1</sup>University of New England, School of Environmental and Rural Science, Armidale, NSW, 2350 Australia

<sup>2</sup>BAIF Development Research Foundation, Central Research Station, Uruli-Kanchan, Pune, Maharashtra, India

<sup>3</sup>INRA 78352 Jouy-en-Josas Cedex, France

### SUMMARY

We report the results of the first large scale milk recording and genetic evaluation for crossbred cows in a smallholder dairy production system in India. Preliminary results represented 8,144 smallholder crossbred cows with a total of 140,214 daily milk records, of which 2,946 animals were genotyped with the GGP Bovine and Illumina SNP assays. Data were adjusted for fixed effects and analysed with a random regression (RR) model with the 1<sup>st</sup> degree Legendre polynomial and heterogeneous variance. Heritabilities of milk yield ranged from 0.14 to 0.22 throughout the lactation period, with an average value of 0.19. Genomic Estimated breeding values (GEBV) for the genotyped animals including the smallholder crossbred cows and the bulls and dams from the BAIF bull stud ranged from +1.9 to -1.4 kg/day. The moderate heritability of the milk yield found in our results together with the wide range of GEBV, indicate that a good response to genomic selection for milk yield can be expected for smallholder dairy farms in India.

### INTRODUCTION

Genomic selection is now widely applied in the dairy industry, which has resulted in a rapid rate of genetic gain for dairy cattle breeds in developed countries. In India, the world's largest milk producer, the majority of milk production is undertaken by smallholders, who mostly keep one to three cows. These cows are mainly multi-generation crosses between indigenous Indian cattle and exotic dairy breeds, which have no pedigree information. Consequently, implementing traditional genetic evaluation for these smallholder cows has not been possible. Use of genomic selection based on genome-wide single nucleotide polymorphisms (SNPs) and phenotype recording in smallholder herds offers a route to implementing genetic improvement in such populations. The objective of this study was to estimate genetic parameters for milk yield and the genomic estimated breeding values (GEBV) for a large genetic gains project initiated by BAIF, a large NGO serving smallholder farmers in India.

### MATERIALS AND METHODS

A mixed model analysis using random regression to estimate genetic parameters and GEBV was performed on animals with genotypes. These included 2,946 smallholder crossbred cows with milk records plus 653 bulls and cows from the BAIF bull stud. 2,389 smallholder crossbred cows were genotyped with the GGP Bovine 50K array and 557 crossbred cows were genotyped with the Illumina 780k BovineHD array. 496 animals from the BAIF bull stud animals were genotyped with the Illumina BovineSNP50 array and the rest of the bulls (157) were genotyped with the GGP Bovine 50K array. SNPs were excluded if they had a call rate (CR) < 0.90 and if they are located on the X and Y chromosomes. Individual samples with CR < 0.90 were excluded. Furthermore, an individual sample for which the correlation between its genotype and that of another sample > 0.98 was removed.

The genotypes of all animals across the different arrays were then combined to impute the medium density genotypes up to the high-density using 2,961 reference animals. The reference set for imputation included 1,309 indigenous plus crossbred cattle, 968 Holsteins and 684 Jerseys all genotyped on the Illumina BovineHD array (Illumina, San Diego, CA). The imputation was performed using FImpute v2.2 (Sargolzaei *et al* 2014) and resulted in 737,073 SNP genotypes across 29 *Bos taurus* autosomes. A prior analysis using data for all animals with phenotypes (8,144 cows with 140,214 milk records) corrected milk records for fixed effects. Cows were included in the analysis when they had at least five monthly records in at least one lactation and there was a minimum of 10 cows per AI distribution centre. Milk records were corrected for fixed effects, including parity, season, centre, the interaction of the season by centre, the average lactation curve of the population modelled by a 3<sup>rd</sup>-order Legendre polynomial, the lactation curves for parities modelled by the 3<sup>rd</sup> order Legendre polynomials, and the lactation curves for centres modelled by the 3<sup>rd</sup> order Legendre polynomials. In matrix notation, the fitted random regression model was:

$$y = Xb + Z_1a + Z_2p + Z_3h + e,$$

where  $y$  is a vector of corrected milk records,  $b$  is a vector of fixed effects. The genetic model included fixed effects for the three environments (high, medium and low production environments, based on the estimates for the centre and the herd in which each cow sits), five crossbreed groups (0-20%, 20-40%, 40-60%, 60-80% and 80-100% exotic dairy breed as estimated from SNP genotypes), and the interaction between environment and breed composition;  $a$ ,  $p$ , and  $m$  are vectors of animal additive genetic, permanent environmental, and household effects;  $X$ ,  $Z_1$ ,  $Z_2$ , and  $Z_3$  are incidence matrices of fixed, additive genetic, permanent environmental, and household effects; and  $e$  contains the residual effects. Legendre polynomials of the 1<sup>st</sup> order ( $\alpha_0$  and  $\alpha_1$ ) were fitted to the animal additive genetic, animal permanent environmental and farm random effects. Days of lactation used to estimate Legendre polynomials ranged from 8 to 340. Residual variance was assumed to be homogenous for test-day records within, but heterogeneous between eight lactation period classes: 8-49, 50-91, 92-133, 134-175, 176-217, 218-259, 260-301, and 302-340 days. Variance components and GEBV were estimated using restricted maximum likelihood in Wombat software (Meyer 2007).

## RESULTS AND DISCUSSION

Figures 1 and 2 show the estimated lactation curves for parities (the first 7 parities) and centres (81 centres) within the six states. The variation in lactation curves between Centres is much greater than that between lactations, perhaps reflecting the smaller number of cows per Centre (ranging 10 to 224) but likely also the substantial differences in true lactation shape between the different environments represented by different Centres. Some estimated lactation curves have implausible shapes. For example, those with extreme curvature and peak yields in mid lactation seen for several Centres in Punjab, Maharashtra and Jharkhand states. Unusual shapes of curves could result if cows are not evenly balanced across the lactation. However, examination of the most extreme curves did not indicate that this was a major factor.

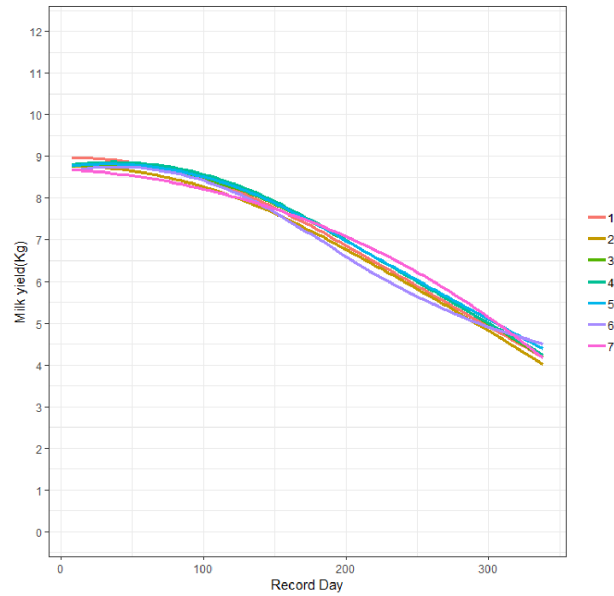


Figure 1. Estimated lactation curves for the first seven parities

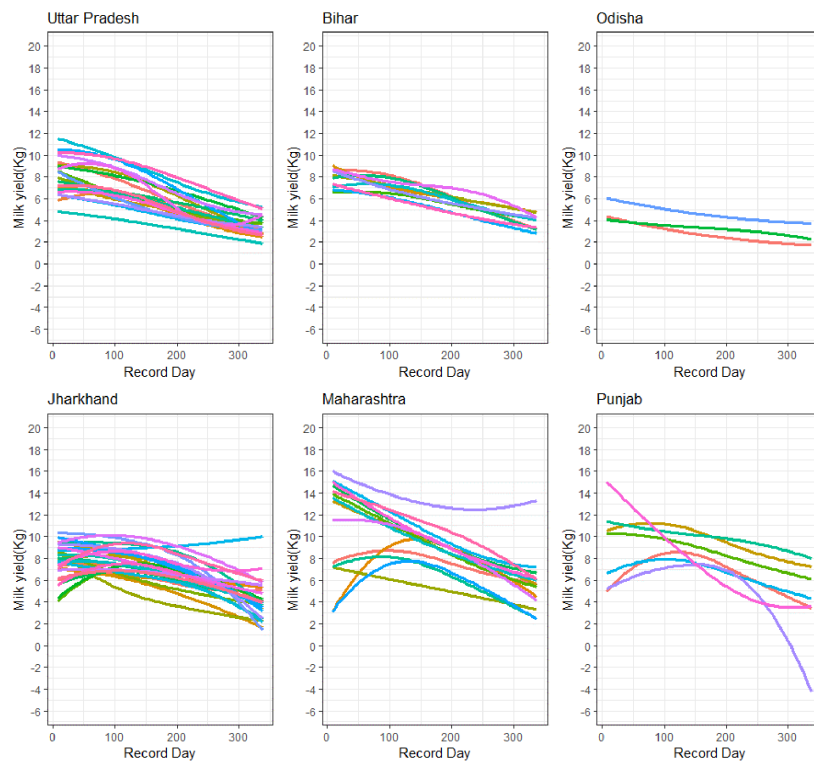
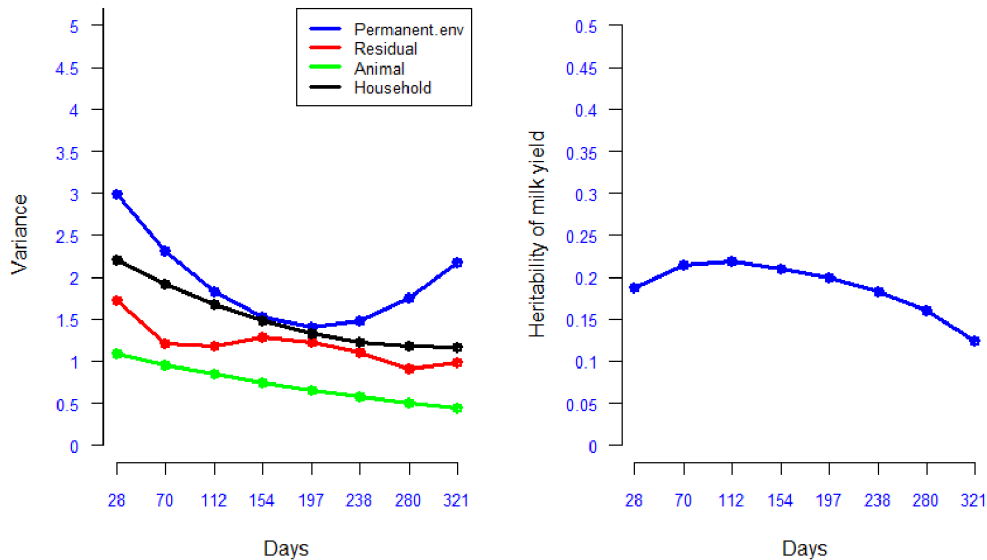


Figure 2. Estimated lactation curves for centres in each state



**Figure 3.** heritability estimates of milk yield over 28, 70, 112, 154, 197, 238, 280, and 321 days of lactation. The heritability was estimated as the proportion of additive genetic variance relative to the sum of the additive genetic, permanent environment, and residual variances

Heritability estimates of milk yield (Figure 3) ranged from 0.14 to 0.22 throughout the lactation period with an average value of 0.19. The highest heritability was observed at around three to four months of lactation. GEBVs for the genotyped animals including the bulls and cows from the BAIF bull stud and the cows from the smallholders were estimated. The range of GEBV of smallholder crossbred cows for average daily yield ranged from +1.9 to -1.35 kg/day, whereas the range of GEBV of bull stud animals ranged from +1.4 to -1.3kg/day. Since the average daily milk yield in this population is 7.5 kg/day, this indicates that there is good opportunity for substantial genetic gain through selection based on GEBV for crossbred performance.

#### ACKNOWLEDGEMENT

We acknowledge the Bill and Melinda Gates Foundation grant OPP1112185. The contributions of J. Khadse, V. Podtar, Y. Gaundare, K. Bhavne, S. Jhadav, S. Khade, V. Dhanikachalam & P. Deshpande in collecting and processing the samples is gratefully acknowledged, as is the support and guidance of G. Sohani and A. Pande.

#### REFERENCES

- Meyer K. (2007) *J Zhejiang Univ Sci B*. **8**:815.  
 Sargolzaei M., Chesnais JP. and , Schenkel FS. (2014) *BMC Genomics*. **15**:478.