A novel use of high density SNP assays to optimize choice of different crossbred dairy cattle genotypes in small-holder systems in East Africa.

J.M.K. Ojango[1], A. Marete[1], D. Mujibi[1], J.Rao[1], J. Poole[1], J.E.O. Rege[2], C. Gondro[3], W.M.S.P. Weerasinghe[3], J. P. Gibson[3] and A.M. Okeyo[1].
[1]International Livestock Research Institute, Kenya; [2]PICOTeam East Africa, Kenya;
[3]University of New England, Australia

ABSTRACT: We present use of SNP technologies to obtain for the first time, rapid, large-scale, *in situ* estimates of performance of crossbred cows in smallholder herds. Compared to historical approaches, our approach allowed optimum crossbreed choices to be determined more rapidly and without question about the relevance of the environment. High-density SNP assays were used to estimate breed composition for a monitored population comprising 1292 cows from 610 smallholder farms in Kenya. The achieved milk yields are much lower than generally assumed and lactation curves are very flat out to 400 days. In poor production environments, lower grade crosses (containing <60% exotic dairy alleles) had the same yield as high grade crosses. Allowing for the larger size and maintenance requirements of high grade exotics, lower grade exotics will be the most economically productive animals in these environments.
Keywords: SNP assay; breed composition; dairy

Introduction

Developing countries are faced with an increasing demand for livestock products mainly driven by an increasing human population, urbanization and growth in incomes (World Bank (2008); Thornton (2010)). In many developing countries there is rapidly increasing demand for milk and the majority of milk in most of these countries is being produced by an expanding smallholder dairy sector. In East Africa, it is generally estimated that 90% of milk is produced by smallholders, with the majority of milk traded through both formal and informal milk markets being produced by crossbreds between high production potential exotic dairy breeds, such as Holstein Friesian, Ayrshire, Guernsey, Jersey and highly adapted, low production potential indigenous breeds. The rapid expansion and large scale of the smallholder dairy sector is testament to the fact that most smallholders who keep dairy cows experience a substantial improvement in their livelihoods, despite the fact that the achieved yields are low compared to yields in intensive dairy systems. There is an implicit assumption in many dairy development programs that smallholder farmers should progressively move to higher grade exotic crossbreed genotypes. The only available semen in most of East Africa is from purebred-exotic breeds. Although only 5% to 20% of smallholder cows are bred using AI, the lack of crossbred semen means there is a steady pressure toward higher grade exotic genotypes. However, it remains unclear what are the most suitable grades of crossbred animal for different types of smallholders. To answer this question it needs to be recognized that smallholders operate in a wide variety of environments and have a wide range of management ability, hence the production environment experienced by the cow is expected to vary widely.

Historically, assessment of crossbreed performance has involved designed breeding programs with animals tested either in research herds or in some cases provided to farmers. Such approaches typically involve limited numbers of animals and take many years (often more than 10 years) to complete. Since smallholders already own a range of different crossbred genotypes, and operate across a range of different production environments, the ideal situation would be to test crossbreed performance, using smallholders' animals recorded *in situ.* Lack of deep and reliable pedigree information has prevented this approach until now. High density SNP assays now provide a route to solving this problem. We here report the use of high density SNP assays to estimate individual animal breed composition and combine that information with field recording of individual animal performance to obtain assessment of crossbreed performance in different herd environments. By undertaking this study using a cluster randomized sampling of farms we were able also to test assumptions about the range of production levels achieved by smallholder cattle. The widespread belief is that the average smallholder achieves yields averaging 8 to 15 kg per day (equivalent to 2400 to 4500kg per lactation) and this belief is an assumption behind most extension and development programs. But appropriate interventions will differ markedly for farmers operating above or below this range of achieved yield.

Materials and Methods

Farm and animal recording was undertaken with smallholder farmers from 7 sites in Kenya and Uganda, representing a range of climate and farming conditions typical for highland (>1000m) sub-tropical and temperate smallholder dairy systems. The sampling did not include peri-urban systems. Smallholders were defined as farms with <11 cows. Farms were only sampled for recording if they had two or more lactating cows or in-calf heifers, and animals included in recording were assessed to be crossbreds based on physical appearance (essentially coat colour patterns and absence of hump) and farmer statements about the breed of the animal. Data on physical characteristics, size and milk yield of 1976 cows were collected by enumerators. Extensive data on household and farm characteristics, farmer beliefs, practices and access to services was also collected. At the initiation of recording, hair samples were collected from 1820 crossbred animals and a total of 212 animals believed to be purebred indigenous. Genotypes for these animals were obtained using the Illumina BovineHD BeadChip.

Reported here are analyses of the performance data from Kenyan sites collected between January 2011 and December 2012. Analyses included the data from 1292 animals on 610 farms for which there was substantial data through the recording period. From an initial 38,085 test-day (TD) records, a final dataset meeting a range of quality control parameters comprised 31,657 TD records. Analyses of the SNP data and estimation of breed composition reported here included all the animals with hair samples in both Kenya and Uganda. After stringent quality assurance checks, data on 566,000 SNP were included in the analyses.

Data analyses. The SNP data were subject to principal component analyses (PCA) and analyses to obtain estimates of breed composition using the Admixture software (Alexander et al., (2009)). Data on reference breeds as controls were obtained variously from the Hapmap consortium and commercial operations. After excluding reference populations that PCA showed were not purebred, data were included for Nelore (as *Bos indicus* reference), N'dama (as a purebred African *Bos taurus*), and Friesian, Holstein, Ayrhsire, Guernsey, Jersey (as reference for *Bos taurus* dairy breeds known to have been used in the region). The Admixture analyses of the exotic dairy versus indigenous breed proportion were highly robust to the data and assumptions used in the analyses. Estimates of each animal's proportion of exotic dairy breed alleles were used to group animals into 5 classes, where 1= 0-20%, 2=21-35%, 3=36-60%, 4=61-87.5% and 5= >87.5% exotic dairy contribution. Estimates of individual dairy breed contributions to each animal's genotype (e.g., proportion of Friesian vs Holstein vs Ayrshire etc) were more dependent on the analyses performed and are not reported here.

A mixed-effects regression model in ASReml 3 (2009) was used to analyse the performance data as follows:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Qu} + \mathbf{Wsf} + \mathbf{e},$$

where **y** = individual test-day yields, and **b** = solutions for the fixed effects in the model, with the following fixed effects: year-month of TD j (j=1-28), parity k (k = 1, 2, ≥3), %dairyness class l (l=1 to 5); lactation stage in 100 day intervals, m (m= 1 to 4). **u** = random cow effects, **sf** = random herd effects, **e** = random residual effects, **X** is the incidence matrix relating records to fixed effects, and **Q** and **W** are incidence matrices that relate records to cows and farms. It is assumed that var(**u**) = σ^2_a, var(**sf**) = σ^2_{sf} and var(**e**) = $\mathbf{I}\sigma^2_e$**.** This model was used to obtain solutions for random herd effects and these solutions were ranked into the bottom, middle and top one third to create three herd level classes. Herd level (HeL) was then added to the above model as another fixed effect along with the interaction between HeL and %dairyness and the interaction between %dairyness and lactation stage. This approach was taken to let the data define what was the farm production environment experienced by the cow and then determine whether genotypes (%dairyness classes) performed differently at different herd levels of production. Additional analyses were also performed fitting lactation curves using Legendre polynomials, the results of which are being further evaluated due to peculiarities in resultant lactation curves of breed groups with small sample sizes. .

Results and Discussion

The raw average estimate TD milk production was 5.38 with s.d. of 3.31 kg, giving a coefficient of variation of >60%. All fixed effects in the model were statistically significant ($p<0.01$).

The estimates of herd, animal and residual variance were σ^2_{sf} = 2.62, σ^2_a = 1.84, and σ^2_e = 5.93, which translate to s.d. for milk yield of 1.62, 1.36 and 2.32 L/day. The implied 95% range for herd levels is approximately 2.2 to 8.6 L/d and for animals at an average herd level, approximately 2.7 to 8.0 L/d. Converting this range of animal yields to implied 305d yields, the range of animal yields would be from 823 to 2453L. These yields are less than half yield levels usually assumed for smallholder dairy in Kenya.

When herd effects were grouped into the three HeL classes and HeL was included in the analysis, the least squares means were 3.5, 4.9 and 7.3L/d for classes 1,2 and 3 respectively. The interactions between HeL and %dairyness and between %dairyness and stage of lactation were significant with p=0.004 and P<0.001 respectively. This indicates significant differences in performance of %dairyness classes at different levels of herd production and also that there were differences in shape of lactation curve between the %dairyness classes.

Figure 1 shows the fitted lactation shapes for the %dairyness classes 2 to 5 (for clarity of display, class 1 is omitted because the limited amounts of data caused erratic estimates for this class). The most striking feature is that yields are remarkably constant across the lactation, with no evidence of a major peak early in lactation and very slow decline out to 400 days of lactation. Data beyond 400 days was not included in these analyses but approximately 20% of cows had lactations of over 500days and these cows continued to maintain yields out to 500 days and beyond. With such flat lactation curves, and two often erratically timed rain seasons each year it is questionable whether the optimum income will be achieved through 365day calving intervals, which many extension programs still push farmers to achieve. With very flat lactation curves, the optimum genotype and management strategy may be to push for very long lactations and long calving intervals that maximize the proportion of time each cow is producing income and maximizes the probability that each cow is in lactation when the rains arrive and feed becomes plentiful.

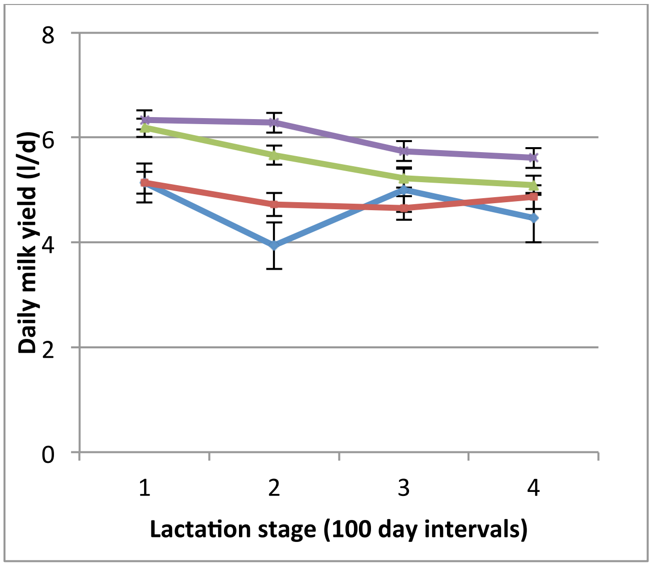

Figure 1. Daily milk yield at each lactation stage for %dairyness classes 2,3, 4 and 5 with standard error bars. *blue= class 2, Red= class 3, green = class 4, and purple = class 5*

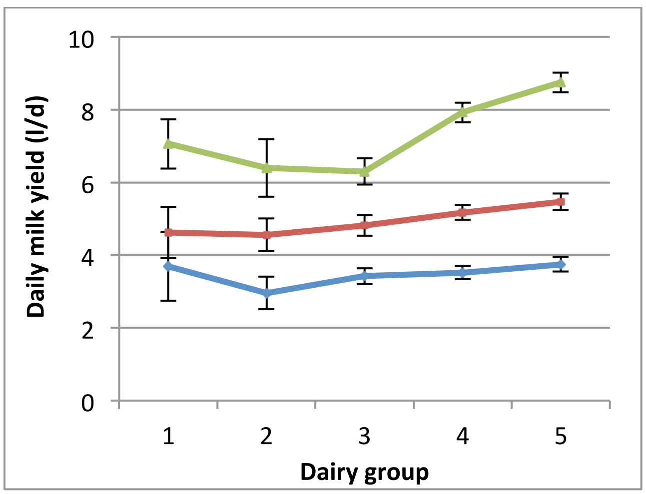

Figure 2. Daily milk yield of the five % dairyness classes with standard error bars for each of the three herd levels. *Blue = lowest, red = middle, and green = highest herd level.*

Figure 2 shows the fitted effects for the five %dairyness classes for each of three levels of herd production. A clear picture emerges that at the lowest level of herd production there is little difference in milk yield between the five %dairyness classes. As herd level of production increases the %dairyness classes 4 and 5 (61-87.5% and >87.5% dairy genes, respectively) give higher yields than lower grade cows with the clearest difference in yields being at the highest herd level. Although detailed analyses remain to be performed, prior prediction and field observation indicate that high grade cattle are much larger than lower grade cattle (the indigenous small east African zebu breed in this region is much smaller than the Holstein/Friesian and Ayrshire breeds to which they have predominantly been crossed). Also there is published evidence that *Bos taurus* dairy breeds have at least 40% higher maintenance requirements per kg body weight than zebu breeds. Extensive data on disease, survival and reproduction are not yet complete but preliminary results indicate a small trend to lower performance in high grade cows. This indicates that high grade cattle will be less economically viable than intermediate or low grade cattle in low herd production environments. The situation remains unclear in the better herd environments but suggests strongly that high grade exotics will become the most economically productive genotypes in the best herd environments. Although our samples did not capture very high herd production environments, there are a modest proportion of farmers, particularly in peri-urban cut-and-carry systems who achieve production levels well above those encountered in this study. These farmers will almost certainly benefit from the use of high grade animals.

Conclusions

The very low average yields observed in this large random sample of representative smallholder farming systems in Kenya has substantial implications for dairy extension and development programs and for businesses providing services, including breeding services to these farmers The current results clearly show that low to medium grade exotics will be the best option for farmers with low, and probably also for medium, herd production levels. More detailed analyses of the socio-economic implications of the various levels of productivity under the systems studied and additional data being generated in the project should improve the predictions of where use of higher grade exotics will begin to generate substantial extra value for smallholders.

Literature Cited

Alexander, D. H., Novembre, J., & Lange, K. (2009). Genome Research. doi: 10.1101/gr.094052.109

Thornton, P. K. (2010). Philosophical Transactions of the Royal Society: Biological Sciences 365(1554): 2853-2867.

World Bank (2008). FAP, PPLPI Working Paper RR Nr 08-07; August 2008