Use of SNP genotyping to determine pedigree and breed composition of dairy cattle in Kenya

D.M. Gorbach¹, M.L. Makgahlela¹,², J.M. Reecy¹, S.J. Kemp³,⁴, I. Baltenweck⁴, R. Ouma⁴, O. Mwai⁴, K. Marshall⁵, B. Murdoch⁵, S. Moore⁵ & M.F. Rothschild¹

¹ Department of Animal Science and Center for Integrated Animal Genomics, Iowa State University, Ames, IA, USA
² Agricultural Research Council, Private Bag X2, Irene, South Africa
³ School of Biological Sciences, Biosciences Building, Liverpool, UK
⁴ International Livestock Research Institute, Nairobi, Kenya
⁵ 410 AgFor Centre, University of Alberta, Edmonton, AB, Canada

Introduction

Coefficients of relationship between pairs of individuals play a very important role in many areas of quantitative genetics, conservation genetics and molecular ecology. Knowledge of the genetic relationships in different populations is used for the estimation of quantitative genetic parameters (e.g. heritabilities and genetic covariances) and breeding values (Lynch & Walsh 1998; Ritland 2000), is necessary for kin selection (Morin et al. 1994), and allows for the study of mating systems (Engh et al. 2002; Frankham et al. 2002). In the management of populations, availability of pedigree structure or the co-ancestries between the individuals that belong to it helps to avoid the loss of diversity and control inbreeding (Ballou & Lacy 1995; Meuwissen 1997; Caballero & Toro 2000; Frankham et al. 2002).

In many developing countries, such as those of East Africa, the necessary pedigree and performance data are often not reliably recorded or are unavailable (Rege et al. 2001). Furthermore, the relatively few exotic (i.e. non-indigenous) genotypes available for import as semen are typically selected on the basis of their genetic merit under European or North American production systems. Consequently, losses because of disease and other environmental demands may be high and this presumably further...
narrates the range of exotic genetics in the African dairy herds (McDermott & Arimi 2002; Mattioli et al. 2000). There is reason for concern that the herds are subject to inbreeding and subsequent depression of productivity (Rege et al. 2001).

Another risk with importation of exotic germplasm is the loss of species diversity because of elimination of native stock from the African breeding population. Centuries of natural selection have resulted in native African cattle which are adapted for the harsher conditions, and these genetic resources may be lost if too many matings occur to animals of European ancestry. Studies utilizing high-density markers enable researchers to assess the current levels of genetic diversity and determine the optimal method for conservation of genetic diversity (Oliehoek et al. 2006; Windig & Engelma 2009). Currently, genetic information about Kenyan cattle is missing, which makes determining the best method of conservation impractical.

Therefore, the objective of this study was to use large-scale SNP data to determine parentage and breed composition of each animal in an admixed population of dairy cattle in Kenya. The determination of breed composition of parents and offspring could provide information on how to improve population management as accurate pedigree records are not available for this assessment.

Materials and methods

The following 23 bulls were sampled for this study with all but the first three living in Kenya: three worldwide Holsteins, 10 from the Central Artificial Insemination Station (CAIS) and 10 from smaller farms. All bulls were chosen based on their heavy use in Kenyan breeding programmes for genetic improvement of dairy cattle populations and status as putative sires of replacement heifers (stock). An additional 71 putative cow-calf pairs from different small herds were identified and sampled, plus 25 cows across 4 generations from a large Kenyan ranch (with deeper pedigree reported) and 5 unmatched cows and calves presumed genetically similar to the cow-calf pairs (172 cows and calves total). Blood or semen samples were collected from all 195 animals. For each animal, DNA was extracted from blood samples and genotyped for SNP markers across the genome, using the Illumina 50K bovine SNP chip (San Diego, CA, USA). All markers with Illumina QC scores <0.8 in more than 10% of the population or minor allele frequency <0.015 were removed leaving 37 238 markers.

Maternity and paternity were checked for each calf using Sire-Match software (E. J. Pollak, Cornell University, Ithaca, NY, USA) with 100 randomly selected markers with minor allele frequency greater than 0.45.

Genomic inbreeding coefficients were estimated using the pair-wise kinship coefficients of Ritland (1996) implemented in Spatial Genetic Diversity (SPAGeDi 1.2; Hardy & Vekemans 2002). Coefficients were computed separately for each chromosome because of the software’s limitations and then averaged across the first 20 chromosomes.

A total of 1000 markers which were genotyped in both the HapMap study (http://bgl.anri.barc.usda.gov/cgi-bin/hapmap/affy2/viewMarkers) and the 195 Kenyan animals were inputted into Structure 2.1 (Pritchard et al. 2000) to detect breed composition of the animals. Runs were completed with \( k = 2 \) through \( k = 10 \) clusters using 10 000 reps during burn-in and 10 000 MCMC reps after burn-in with each run replicated twice. Each run predicts the proportion of the animals’ genetic composition that originates from each of the \( k \) clusters.

Results

A total of 8 of the 81 putative cow-calf pairs (which include 10 from the large ranch) showed incorrect maternity. The true dams of the eight calves, all from small farms, were not among the genotyped animals. Either some incorrect parentage identification occurred on the small farms or errors may have occurred during sampling and processing.

Sire-Match software predicted 24 sire-calf pairs, with all nine of the predicted sires coming from CAIS. One bull from CAIS was the predicted sire of eight genotyped calves. This sire was recorded under two different names within the large-scale ranch pedigree, which may have resulted in him being used more than owners realized.

Based on how much of the genetic make-up of the Kenyan cows clustered with the known Holsteins and Guernseys from the HapMap population, the Structure program gave evidence that the cows from the large-scale farm are primarily of Holstein descent, while the animals from smaller farms are a mix of Holstein and Guernsey (Figure 1). The bulls used for AI varied widely in their ancestry with anywhere from 30 to 98% Holstein genetics with the next most represented breed being Guernsey. None of the animals appeared to have significant native African or Bos indicus heritage.
Of 195 animals sampled, 41 had inbreeding coefficients greater than 2.5% including 14 originating from the large-scale ranch (56% of that population) and all 3 worldwide Holstein bulls (Figure 2).

Discussion

To infer genetic relationships and breed composition, we examined the population structure based on SNP data from an admixed population of dairy cattle in Kenya. Based on how well the Structure program clustered animals within breeds from the HapMap population and recognized breed composition for known admixed breeds, this software was used instead of assignation methods which would have assumed animals were purebred. The clustering results from the Structure program reflect the widespread use of Holstein semen in the Kenyan dairy breeding populations and perhaps indicate earlier access to Holstein genetics by the larger ranch owners compared to small-scale farmers. While the large ranch cows are mostly Holstein, the animals from smaller farms average about 50% Holstein blood, consistent with one generation of using Holstein semen. Surprisingly, the Kenyan animals showed little to no evidence of *Bos indicus* ancestry or other native African blood. The other major contributor to most of these cows was most likely the Guernsey breed. Further work should be completed to assess the suitability of these European breeds for milk production in East Africa.

Based on the pedigrees provided by small-scale farmers, ancestry records maintained in Kenya appear to be incomplete and contain a significant number of inaccuracies. While the large ranch pedigree was far more complete, it also contained a small number of errors and showed no evidence of attempts to minimize inbreeding. Because true base population allele frequencies were not available, the values of the genomic inbreeding coefficients are only accurate for comparisons of animals within this population and say nothing about the true inbreeding levels of each animal. Furthermore, insufficient pedigree structure and trio data existed to make other inbreeding estimation techniques viable (data not shown). Consequently, the main conclusion to be drawn is that large-scale ranch cows appear to exhibit more inbreeding than cattle from smaller farms. The cows from the large-scale ranch obviously cluster more with Holsteins than those from smaller farms. Common relatives with Guernsey also appear to have contributed significantly to the Kenyan populations (further resolved with larger cluster size).
farms, although the worldwide Holstein bulls have higher average inbreeding than either group of cows (Figure 2).

Overall, the Structure program clustering results and parentage information obtained in this study indicated that genetically cows from large-scale farms are more similar to Holstein dairy cattle and may be more inbred than local Kenyan animals. Bulls used for AI varied substantially in their genetic background. We also found that data on animal relationships were not always accurately recorded.

Acknowledgements

Funding was provided by the Ensminger Endowment, State of Iowa and Hatch funding. Support for D. M. Gorbach was provided by a USDA-CSREES National Needs fellowship under Grant no. 2007-38420-17767. We thank the two reviewers for their constructive feedback on this article.

References


