

Genetic diversity and breed proportions of Indian stud cattle

E.M. Strucken¹, D.C. Hidalgo², V. Ducrocq², A. Pande³, M. Swaminathan³ & J.P. Gibson¹

¹ *University of New England, Animal Science, Armidale, 2351 NSW, Australia*
estrucke@une.edu.au (Corresponding Author)

² *GABI, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France*

³ *BAIF Development Research Foundation, Warje, Pune 411058, India*

Summary

India is the world's leading milk producer, however, average milk yield per animal is low. Most of India's cattle milk is produced by indigenous breeds, followed by crossbred cows. Crossbreeding intensified in the 1960s and 70s with a large influx of Holstein Friesian and Jersey bulls, and an increased use of artificial insemination. Based on genotypes generated from the 50K v2 BovineHD BeadChip assay, we analysed the genetic diversity and breed composition of 242 breeding cattle sampled from the BAIF breed herds, including indigenous, exotic dairy, and crossbred animals. We compared these animals to indigenous and animals used in smallholder systems in East Africa. Key findings were: confirmation that Indian Zebu breeds are pure *Bos indicus* in contrast to African Zebu and Sanga breeds, which are admixtures of *Bos indicus* and African *Bos taurus*; as found elsewhere, many apparently pure animals contain some admixture from imported breeds; crossbred animals exhibit a wider range of breed composition than intended, reflecting the lack of pedigree when sampling bull mothers from smallholder farms; breed composition of purebred and crossbred animals can in future be tested using SNP data, and only those meeting the desired composition retained for breeding.

Keywords: India, indicine, taurine, crossbred, cattle, cow

Introduction

India is the world's leading milk producer with more than 199 million cattle (**livestock census 2007**), but the low productivity per animal (<1000kg per lactation) is the largest obstacle for India's dairy sector (**FAO**). Nation-wide efforts to increase milk production connected rural dairy sheds to urban markets and increased the penetration of artificial insemination (AI), including the import of exotic taurine breeds to improve productivity through crossbreeding. India has 37 acknowledged indigenous breeds, amongst others Gir, Sahiwal, Red Sindhi, Amritmahal, Hallikar, Khillari, Haryana, or Tharparkar, and a large number of non-descript Zebu cattle. The advent of AI resulted in a large scale crossbreeding of exotic dairy breeds (mainly Jersey and Holstein Friesian) with non-descript Zebu cows. Today, India has 76.7 million adult cattle of which 22.3 million are dairy crossbreds. Indigenous and non-descript dairy cows contribute 26%, crossbred cows 21%, and pure exotic cows 1% of the total milk production in India (**Annual Report 2016-17, Ministry of Agriculture & Farmers Welfare, Government of India**). The remaining milk is produced by buffalo and goats.

Crossbreeding pedigree records are sparse and hence breed proportions of the crossbred stud population in India are unreliably estimated or unknown. This study analysed 242 cattle

sampled from the BAIF breeding herds, comprising 8 indigenous breeds, a purebred Holstein and Jersey population, and a number of crossbred animals, to better understand the genetic diversity of India's breeding cattle.

Material and methods

Animals

Genetic and breed information of 254 Indian dairy cattle were available (BAIF Development Research Foundation, India). After an initial quality control (QC), 12 cattle were excluded due to low call-rates (<90%). The represented breeds and number of animals are listed in **Table 1**.

Table 1. Breeds, number of animals, indicine content, and average diagonal and off-diagonal elements of the GRM per breed for 242 Indian dairy cattle (\pm s.d.).

Type	Breed	# of Animals	Indicine content ¹	GRM diag ²	GRM off-diag ²
indigenous	Amrithmahal	2	1 (\pm 0.0004)	0.15 (\pm 0.003)	0.64
indigenous	Hallikar	1	1	0.13	0.13
indigenous	Haryana	4	0.99 (\pm 0.027)	1.13 (\pm 0.01)	0.60 (\pm 0.065)
indigenous	Khillar	3	1 (\pm 0.0004)	1.16 (\pm 0.014)	0.64 (\pm 0.01)
indigenous	Red Sindhi	1	1	1.12	1.12
indigenous	Tharparkar	1	1	1.22	1.22
indigenous	Gir	12	1 (\pm 0.0)	1.15 (\pm 0.016)	0.64 (\pm 0.022)
indigenous	Sahiwal	25	0.98 (\pm 0.025)	1.11 (\pm 0.024)	0.57 (\pm 0.048)
exotic	Holstein	55	0.00 (\pm 0.0)	1.28 (\pm 0.035)	0.34 (\pm 0.097)
exotic	Jersey	31	0.00 (\pm 0.0)	1.32 (\pm 0.029)	0.55 (\pm 0.09)
crossbred	75% Holstein	67	0.12 (\pm 0.092)	1.14 (\pm 0.065)	0.15 (\pm 0.109)
crossbred	62.5% Holstein	1	0.51	0.83	0.83
crossbred	50% Holstein	13	0.33 (\pm 0.116)	1.03 (\pm 0.068)	0.08 (\pm 0.125)
crossbred	75% Jersey	8	0.16 (\pm 0.103)	1.16 (\pm 0.12)	0.31 (\pm 0.126)
crossbred	62.5% Jersey	4	0.25 (\pm 0.03)	1.06 (\pm 0.015)	0.35 (\pm 0.154)
crossbred	50% Jersey	14	0.41 (\pm 0.14)	0.96 (\pm 0.136)	0.09 (\pm 0.072)

¹ indicine content represents Nelore content derived from ADMIXTURE with 7 ancestral breeds

² GRM was constructed including the reference and East African animals

Reference datasets for N'Dama (an African *Bos taurus* breed; n = 20), Nelore (a *Bos indicus* breed; n = 20), Guernsey (n = 20), Holstein (n = 20), and Jersey (n = 20) were sourced from the International Bovine HapMap consortium. British Friesian (n = 25) were sourced from the SRUC in Scotland and Canadian Ayrshire (n = 20) from the Canadian Dairy Network (CDN). East African Ankole (n = 40), Zebu (n = 21), Sahiwal (n = 38), and Gir (n = 30) from the HapMap consortium were included to analyze similarities of related breeds from other countries.

Genotypes

The Indian cattle were genotyped with the 50K v2 BovineHD BeadChip (Illumina Inc., San Diego, CA, USA). Quality control included a median GC score >0.6, leaving 46,849 SNPs on

autosomes for analysis. The reference breeds were genotyped with 777K BovineHD BeadChip (Illumina Inc., San Diego, CA, USA), and the same QC criteria as for the Indian data were applied when data was provided unfiltered. The datasets were merged to include only those SNPs in common between the Indian and reference data sets.

Data analyses

We constructed a genomic relationship matrix (GRM) according to **van Raden (2009)**, which was used to assess inbreeding and genomic relatedness between animals, and to perform principal component analyses (PCA). Breed compositions of the BAIF cattle were estimated in a supervised analysis with 7 ancestral reference breeds using the ADMIXTURE 1.23 program (**Alexander et al. 2009**).

Results and discussion

The GRM constructed to include the reference and East African breeds showed the highest diagonal elements for the BAIF pure Holstein and Jersey groups, followed by the indigenous pure breeds (**Table 1**). The crossbreds showed, as expected, lower diagonal elements compared to the pure breeds; however, the higher the content of pure exotic the higher the average diagonal element of the GRM (**Table 1**).

The PCA separated the taurine breeds from the indicine breeds on the first PC, and the Holstein and Jersey breeds on the second PC. PC1 explained by far the largest proportion of variance, being 82.1% when just the BAIF animals were included and 90.3% when the reference animals were also included. The crossbreds spread between the indigenous and their respective pure Holstein and Jersey ancestry (**Figure 1a**), with some of the Holstein crossbreds aligning with the Ayrshire and Friesian reference breeds (**Figure 1b**). This likely reflects that BAIF has imported Holstein-Friesian semen since the 1960's and some of the semen used in crossbreeding may be closer to the Friesian than to modern Holsteins.

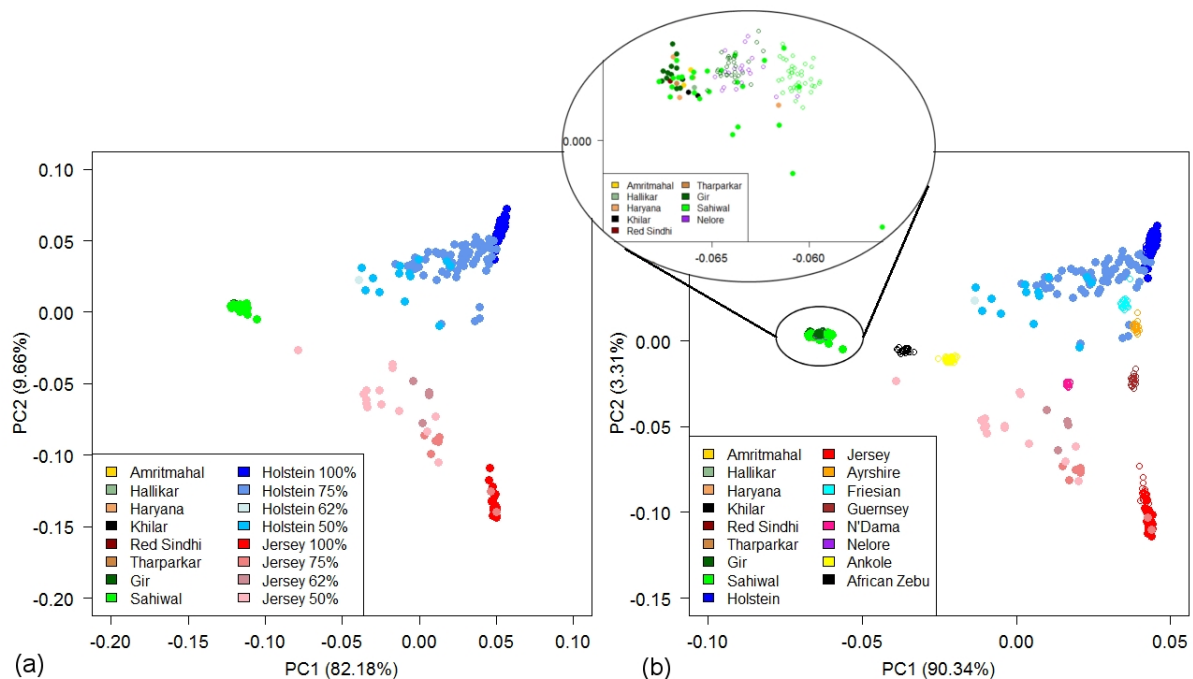


Figure 1. First and second principal components of pure and crossbred bulls from India (a),

and with reference breeds (open circles, b).

The African Zebu animals form a separate cluster to the Indian Zebu, lying a substantial distance along the line between the *Bos indicus* breed cluster and the African taurine (N'Dama), indicating that African Zebu are not pure *Bos indicus* (**Figure 1b**). When the cluster of *Bos indicus* animals in **Figure 1b** is expanded, it reveals that the East African Sahiwal, and the Gir and Nelore from the HapMap data (all shown as open circles) are also located slightly more towards the *Bos taurus* end of PC1 when compared to the Indian Sahiwal and Gir (**Figure 1b**).

The ADMIXTURE analysis with 7 assumed ancestral populations (**Figure 2**) found that most East African Sahiwal contained a small proportion of African *Bos taurus* (here, N'Dama). Many of the Indian Sahiwal showed small proportions of European *Bos taurus* ancestry. The likely explanation is that neither population is completely pure, with Kenyan Sahiwals having some traces of admixture with African Zebras while the Indian Sahiwal sample shows some history of admixture with European *Bos taurus* breeds. The East African Zebu animals average 34% African *Bos taurus* content, which clearly separates them from pure Indian Zebu breeds that are 100% *Bos indicus*. The BAIF pure Holstein group showed some ancestry of other *Bos taurus* breeds. The most likely explanation for this is that BAIF Holsteins result from imports of semen from several countries going back to the 1960's, with older imports of semen still being used. So the BAIF Holsteins are likely to be poorly represented by the HapMap sample of modern Holsteins. ADMIXTURE deals with this discrepancy by attributing differences from ancestral reference (HapMap Holsteins) to other possible ancestral sources of *Bos taurus* genes (here Friesian, Jersey, Ayrshire, Guernsey).

The results for average and s.d. of *Bos taurus* proportion of crossbred cattle in **Table 1**, which come from the ADMIXTURE analysis, show that: A) all crosses have an average *Bos taurus* proportion that is higher than the target proportion for that group; B) that most groups have high s.d. indicating large variation around the target proportions. This is not surprising given the historical lack of pedigree recording in smallholder farms from which many crossbred bull dams are sourced.

The approaches used here to understand the genetic diversity among pure and crossbred stud animals contributing to smallholder dairy improvement can be used in future to improve breeding programs. SNP assays can be used to screen putative purebred animals to ensure that purebred populations used for improvement and conservation are indeed maintained pure. SNP data can also be used to screen crossbred cattle to ensure that bulls entering AI conform to the desired range of breed composition. Similarly, bull dams of crossbred sires can be screened to maximise the chance that progeny bulls will conform to the desired range of breed composition.

List of References

- Alexander, D.H., Novembre J. & K. Lange, 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19:1655-1664.
- Van Raden, P.M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414-4423.

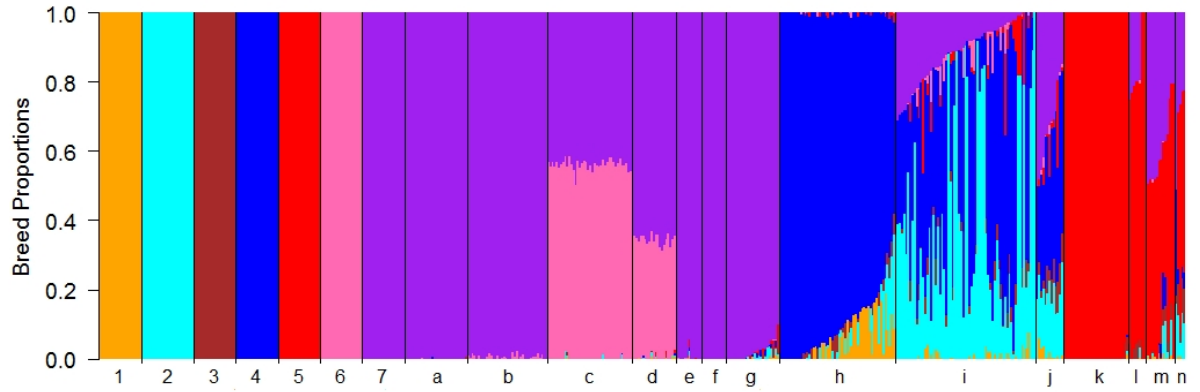


Figure 2. Breed proportions estimated in ADMIXTURE with 7 ancestral breeds.

1: Ayrshire, 2: Friesian, 3: Guernsey, 4: Holstein, 5: Jersey, 6: N'Dama, 7: Nelore
a: Gir HapMap, b: Sahiwal East Africa, c: Ankole, d: Zebu East Africa, e: Pure Indian indigenous breeds, f: Gir India, g: Sahiwal India, h: pure Holstein India, i: 75% Holstein cross, j: 50% Holstein cross, k: pure Jersey India, l: 75% Jersey cross, m: 50% Jersey cross, n: 62.5% Holstein and Jersey cross