Proc. Assoc. Advmt. Anim. Breed. Genet. 23:187-190

GENETIC CHARACTERIZATION OF INDIAN INDIGENOUS CATTLE BREEDS

E.M. Strucken[1], M. Swaminathan[2], S. Joshi[2], and J.P. Gibson[1]

[1]School of Environmental and Rural Science, University of New England, Armidale, NSW, 2351 Australia
[2]BAIF Development Research Foundation, Pune, India

SUMMARY

Using the Illumina 770k bovine SNP assay, we genetically characterized 15 Indian indigenous cattle breeds and 1 non-descript indigenous population, which is the largest sample of Indian breeds yet studied. 15.6% of the animals were found to have more than 1% recent *Bos taurus* admixture and were removed or separately analysed. Inbreeding levels for the Indian indigenous breeds, based on F_{IS} and diagonal elements of the GRM, were similar compared to European taurine breeds. We did not find evidence for historical admixture with *Bos taurus* during breed formation. Only 1.4% of the genetic variance in allele frequencies was between breeds, compared to about 42.4% for European taurine breeds. Consequently, Indian breeds can be treated as a single population for some purposes, such as SNP assay design.

INTRODUCTION

Present day India is accepted to be one of the centres of cattle domestication, in particular where *Bos indicus* cattle have developed from its supposed ancestor *Bos primigenius nomadicus* some 100,000-850,000 years ago (e.g. MacHugh *et al.* 1997; Verkaar *et al.* 2004). Furthermore, archaeological evidence suggest that there might have been several centres of domestication within India, as phenotypic differences between cattle from the North and South were already described as early as the Neolithic time period (Naik 1978). Today, the National Bureau of Animal Genetic Resources in India (http://www.nbagr.res.in/nbagr.html) lists 43 registered Indian cattle breeds, however, the large majority of cattle used as milk, draught or dual purpose cattle are raised by smallholders and are of no descriptive breed (e.g. Sharma *et al.* (2015)).

Bos indicus cattle are well adapted to high temperatures and resistance to some prevailing parasites of tropical regions, and have therefore been exported, bred, and adapted in other parts of the world. Zebu cattle are believed to have entered Africa between 3,500 and 700 BCE through present day Egypt (Marshall 1989), and contributed to the formation of African indigenous Sanga and Zenga type breeds (Rege & Tawan 1999). Others, such as the Brazilian Nellore and Guzerat or Australian Brahman and Droughtmaster have been imported to these countries and crossed with other breeds during the last 200 years (Porter *et al.* 2016).

Despite the importance of indicine cattle breeds globally and their wide use especially for cross-breeding with taurine breeds, knowledge of the genetic diversity of the pure *Bos indicus* breeds in India itself is scarce. Many studies focussed on limited numbers of microsatellite or single nucleotide markers, on breeds outside India, or limited sample sizes (e.g. Dash *et al.* 2018; Nayee *et al.* 2018). Here, we have assembled and analysed the largest dataset on Indian indigenous breeds for genetic diversity and relationship, and compared these breeds with taurine and other indicine reference breeds. Lastly, we draw conclusions with regards to the requirement of genomic tools designed specifically for indicine cattle populations.

MATERIALS AND METHODS

Data. A total of 702 Indian indigenous cattle from 15 registered breeds and one non-descript (ND)

population were sampled by the BAIF Development Research Foundation (Table 1). All animals were genotyped with the 777k-SNP BovineHD Beadchip (Illumina Inc., San Diego). Genotypes and animals were quality controlled (QC) based on a median GC score >0.6 and a call rate >0.9. Further, animals with more than 1% *Bos taurus* content (based on a preliminary Admixture analysis) were excluded, leaving 588 animals and 716,599 SNPs for subsequent analyses. Reference breeds included 6 *Bos taurus* breeds (each N=20), and 16 *Bos indicus* breeds (N=10 to 20), sourced from the bovine HapMap consortium, Canadian Dairy Network, SRUC, and Decker *et al.* (2014). Reference data were either previously quality controlled or subjected to the same QC criteria as the Indian indigenous breeds. The 770k Illumina assay has close to 300,000 SNPs that are at high minor allele frequency in *Bos indicus* breeds, so that it has much lower ascertainment bias than earlier versions of the 50k assay.

Analyses. Analyses included calculation of Pearson's correlation coefficient between observed (R_{obs}) and expected (R_{exp}) allele frequencies for each breed-pair. These calculations only included SNPs with frequencies $0.05<p<0.95$ to reduce bias due to limited numbers of SNPs with small frequencies. R_{exp} was calculated as follows:

$$R_{exp} = V_p/[V_p+V_{e1}+V_{e2}],$$

where V_p is the variance of p in the meta-population (i.e. all Indian indigenous animals or all *Bos taurus* animals), V_{e1} and V_{e2} are the error variances of the estimates of p in the two breeds. V_{e1} and V_{e2} were estimated as the average across all loci of $p(1-p)/2n$, where n is the number of animals in the given breed and p is the meta-population value of p for each SNP. V_p was not corrected for the sampling error of p, which in all cases was less than 1% of the estimate of V_p. The variance of true SNP allele frequencies in one breed that was explained by the true SNP allele frequencies in another breed was estimated as the ratio of R_{obs}^2/R_{exp}^2.

Principal components were estimated using a GRM based on Van Raden (2008). Further analyses included supervised Admixture models including reference breeds as potential ancestors (Alexander *et al.* 2009). Genetic differentiation between and within breeds were estimated using F_{ST} (Weir & Cockerham 1984), and F_{IS} (Nei 1972), respectively (Table 1).

Table 1. Data information on Indian indigenous breeds and inbreeding levels (F_{IS})

Breed	N	Sampling location	# excluded / reason	F_{IS}
Dangi	68	Maharashtra	3 / taurine>0.01	-0.015 (±0.124)
Gaolao	20	Maharashtra	1 / taurine>0.01	0.022 (±0.226)
Gir	121	Gujarat	3 / taurine>0.01	0.012 (±0.101)
Hallikar	28	Karnataka	1 / taurine>0.01	0.002 (±0.179)
Hariana	17	Haryana	4 / taurine>0.01	0.0002 (±0.262)
Khillar	25	Maharashtra	1 / taurine>0.01	0.015 (±0.2)
Krishna Valley	22	Karnataka	5 / taurine>0.01	0.004 (±0.218)
Red Kandhari	35	Maharashtra		0.008 (±0.168)
Malnad Gidda	19	Karnataka	5 / taurine>0.01	0.001 (±0.274)
Ongole	50	Andhra Pradesh	4 / low call rates	0.028 (±0.153)
Rathi	1	Rajasthan		NA
Red Sindhi	63	Odisha	1 / low call rates 20 / taurine>0.01	-0.034 (±0.154)
Sahiwal	140	Punjab	36 / taurine>0.01	0.015 (±0.108)
Tharparkar	48	Rajasthan	3 / taurine>0.01	-0.024 (±0.144)
Vechur	1	Kerala		NA
Non-descript	43	Maharashtra, Odisha, Uttar Pradesh	27 / taurine>0.01	0.029 (±0.236)

RESULTS AND DISCUSSION

Principal components analysis and Admixture showed a clear separation between *Bos indicus* and *Bos taurus* breeds. The majority of the 109 Indian indigenous animals with more than 1% *Bos taurus* content belonged to Sahiwal, Red Sindhi and ND. Nayee *et al.* (2018) also found some of their Red Sindhi sample to have some taurine admixture. The otherwise tight clustering of the Indian indigenous breeds indicates that the taurine admixture is recent and not, as some sources speculate, a result of crossing *Bos indicus* with *Bos taurus* animals during the early history of breed formation.

Observed allele frequency correlations between Indian indigenous breeds were on average 0.92 (±0.02). In comparison, R_{obs} between the exotic taurine breeds was 0.65 (±0.04). The estimated proportion of variance of true allele frequency explained by the true frequency in another breed was on average 0.986 in the Indian indigenous breeds; i.e. most or all of the genetic variance at the SNP level is within breeds. The estimated proportion of variance that is within-breeds for the *Bos taurus* breeds was 0.576. These results suggest that, in contrast to *Bos taurus* breeds, Indian indigenous breeds can be treated as a single population for some purposes, such as SNP assay design.

Figure 1a) shows the estimated breed proportions of the Indian indigenous breeds based on the indicine reference breeds as a heatmap. Red Sindhi and Gir were both best represented by the Red Sindhi and Gir reference breeds. Hallikar and Khillar showed a strong Ongole signal, whilst Ongole were best represented by the Nelore reference, which confirms the connection that Brazilian Nelore were bred from imported Indian Ongole (Porter *et al.* 2016). Tharparkar were, however, not best represented by the Tharparkar reference but by Kankraj and Dhanni; and Sahiwal were represented as an admixture of Tharparkar, Sahiwal and Hissar, which stands in contrast to Nayee *et al.* (2018) and Gajjar *et al.* (2018) who reported their Sahiwal sample to have the least evidence for admixture. These and other analyses indicate that it is difficult to trace history and relationships among Indian indigenous breeds which is not unexpected given the low level of between-breed variation estimated for these populations.

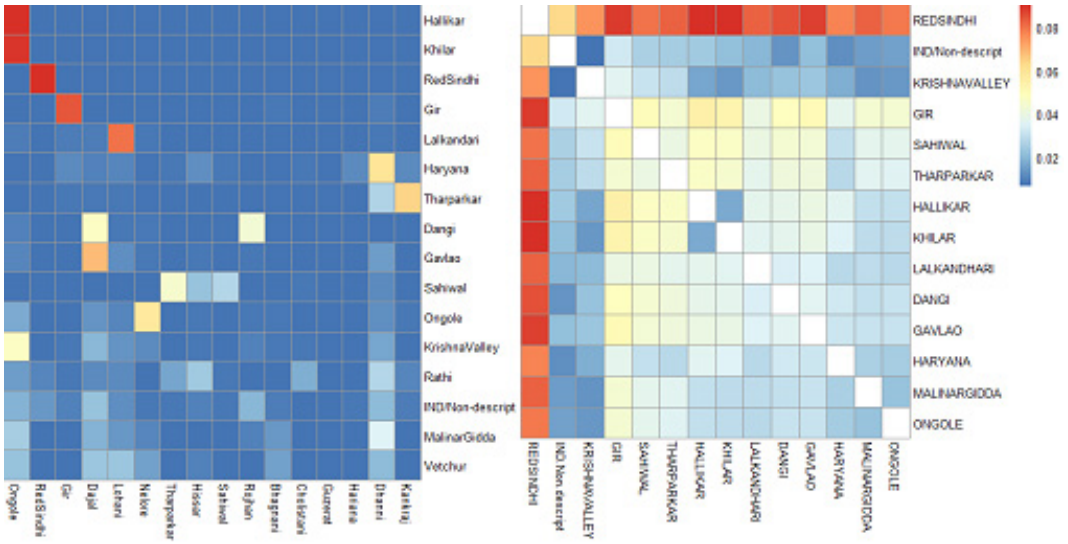

Figure 1. a) Heatmap of estimated breed proportions for 16 Indian indigenous populations (vertical) from a supervised Admixture analyses with 16 indicine reference populations (horizontal); b) heatmap of pair-wise F_{ST} values between 14 Indian indigenous populations

Figure 1b) shows pairwise F_{ST} values as a heat map with a phylogenetic tree based on hierarchical clustering, when only the Indian indigenous breeds were considered. This clearly shows Red Sindhi as an outgroup to the other indigenous breeds, whilst the ND followed by Krishna Valley are the least genetically distinct groups. The genetic distinction of Red Sindhi might reflect their sampling from a single central breeding farm. However, Nayee *et al.* (2018) also found Red Sindhi to be genetically different from other Indian indigenous breeds.

Levels of inbreeding as measured by F_{IS} are similar (-0.034 to 0.029) compared to the taurine reference breeds (-0.026 to 0.023). Studies based on microsatellite data found increased F_{IS} values (e.g. Sharma et al. 2015). Whilst exact F_{IS} values cannot be directly compared between these studies, we can confirm that higher inbreeding levels were found for Gaolao and Ongole and comparatively lower values for Hariana (Table 1).

CONCLUSION

Indian indigenous breeds show remarkably little between-breed variation, and therefore can be treated as a single population when developing genomic tools such as SNP assays.

ACKNOWLEDGEMENT

We acknowledge the Bill and Melinda Gates Foundation grant OPP1112185.

REFERENCES

Alexander D.H., Novembre J. and Lange K. (2009) *Genome Res.* **19**: 1655.

Danecek P., Auton A., Abecasis G., Albers C.A., Banks E., DePristo M.A., Handsaker R.E., Lunter G., Marth G.T., Sherry S.T., McVean G. and Durbin R. (2011) *Bioinfor.* **27**: 2156.

Dash S., Singh A., Bhatia A.K., Jayakumar S., Sharma A., Singh S., Ganguly I. and Dixit S.P. (2018) *Anim. Biotechnol.* **29**: 129.

Decker J.E., McKay S.D., Rolf M.M., Kim J., Molina Alcala A. and Sonstegard T.S. (2014) *PLoS Genet.* **10**.

Gajjar S.G., Guldbrandtsen B., Nayee N., Sudhakar A., Trivedi K., Lund M.S. and Sahana G. (2018) *Proc. World Cong. Genet. App.Livest.Prod. (WCGALP)*, Auckland.

Hill W.G. and Robertson A. (1968) *Theor. Appl. Genet.* **38**: 226.

MacHugh D.E., Shriver M.D., Loftus R.T., Cunningham P. and Bradley D.G. (1997) *Genet.* **146**: 1071.

Marshall F. (1989) *Current Anthropology* **30**: 235.

Naik S.N. (1978) *J. Hum. Evol.* **7**: 23.

Nayee N., Sahana G., Gajjar S., Sudhakar A., Trivedi K., Lund M.S. and Guldbrandtsen B. (2018) *J. Anim. Breed. Genet.* **135**: 432.

Nei M. (1972) *The Am. Naturalist* **106**: 283.

Porter V., Alderson L., Hall S.J.G. and Sponenberg D.P. (2016) 'Mason's World Encyclopedia of Livestock Breeds and Breeding', 2 Volume Pack. CABI.

Rege J.E.O. and Tawan C.L. (1999) *Anim. Genet. Res. Information Bulletin* **26**: 1.

Sharma R., Kishore A., Mukesh M., Ahlawat S., Maitra A., Pandey A.K. and Tantia M.S. (2015) *BMC Genet.* **16**: 73.

Sved J.A. (1971) *Theor. Popul. Biol.* **2**: 125.

Van Raden P.M. (2008) *J. Dairy Sci.* **91**: 4414.

Verkaar E.L., Nijman I.J., Beeke M., Hanekamp E. and Lenstra J.A. (2004) *Mol. Biol. Evol.* **21**: 1165.

Weir B.S. and Cockerham C.C. (1984) *Evol.* **38**: 1358.