Genetic diversity of the indigenous cattle of Kenya, Uganda, Ethiopia and Tanzania using high-density SNP data

W.M.S.P Weerasinghe1, 2, C. Gondro1, A.M. Okeyo3, J. Ojango4, J. Rao5, T. Dessie6, F. D. Mujibi4, J. E. O. Rege4 & J. P. Gibson1

1Centre for Genetic Analysis and Applications, University of New England, Armidale, NSW, 2351, Australia
shalanee.weerasinghe@abri.une.edu.au (Corresponding Author)
2Agricultural Business Research Institute, University of New England, Armidale, Australia
3International Livestock Research Institute, Nairobi, Kenya
4PICO Eastern Africa, Nairobi, Kenya

Summary

Indigenous cattle make a significant contribution to the livelihood of many communities in Ethiopia, Tanzania and other countries in eastern Africa. Here, we identify the genetic structure and the admixture levels of several East African indigenous cattle breeds in Ethiopia and Tanzania. Two ‘groups’ were studied: Indigenous cattle consisting of a number of breeds; and Mpwapwa cattle – a composite breed. A total of 386 individual animals from the two groups were genotyped using the Illumina high-density Bovine SNP chip (778k Panel). Principal component analysis was used to study the genetic structure and admixture levels of the indigenous cattle, and Mpwapwa were estimated using the ADMIXTURE program. All East African indigenous breeds other than the Ankole appear genetically closely related to each other and consist of a mixture of African taurine and indicine signals. Ethiopian indigenous breeds, Fogera, Danakil Harar and Ethiopian Boran show high purity, whereas Ethiopian Central Highland Breed and the Begait samples show significant amounts of European Bos taurus admixture. Tanzanian indigenous cattle, Singida White and Iringa Red, have a high degree of purity while the TALIRI Boran shows some European Bos taurus genetic background. The synthetic Mpwapwa breed had estimated breed proportions of Bos indicus, African Bos taurus and European Bos taurus of 0.82, 0.05 and 0.13 respectively. These results are useful for genetic conservation and genetic improvement programs.

Keywords: genetic structure, indigenous cattle, East Africa, admixture

Introduction

The genetic diversity of the East African cattle is high and they have distinctive adaptations to local biophysical environments (Mbole-Kariuki et al., 2014). These cattle breeds have a complex admixture, resulting from ancient crossings of African Bos taurus cattle and Bos indicus cattle. Some breeds have more recent, deliberate or unplanned contributions of European Bos taurus breeds introduced to Africa (Gifford-Gonzalez and Hanotte, 2011). There is a general concern that the genetic variation within East African indigenous cattle is being eroded through breed substitution and indiscriminate crossbreeding. Having a deeper understanding of the admixture levels and population structures of these breeds is essential for the genetic design and implementation of conservation programs and better utilisation of the cattle genetic resources.
In Tanzania, the composite dual-purpose Mpwapwa cattle breed was developed from 1958 onwards to increase productivity while retaining a high degree of adaptation to harsh environments (Syrstad, 1990). The use of high density molecular markers provides an opportunity to determine the current breed composition of Mpwapwa. The objectives of this study were to assess the genetic structure of East African indigenous cattle breeds and Mpwapwa cattle using high-density SNP markers.

Material and Methods

The genotypes from the BovineHD (780k) Beadchip array (Illumina Inc) were collected on indigenous in Ethiopia and Tanzania, and Mpwapwa cattle in Tanzania by the Dairy Genetic East Africa project. Samples were collected from a series of indigenous breeds: the TALIRI (developed by the TANZANIA LIVESTOCK RESEARCH INSTITUTE) Boran herd, the TALIRI Mpwapwa herd, the Kenya Sahiwal stud and purebred Kenyan Orma Boran as summarised in Table 1. As reference samples, Ankole (ANK), Kenyan Small East African Zebu (SEAZ), Friesian (FRI), Holstein (HOL), Canadian and New Zealand Ayrshire (AYRCDN and AYRNZ), Jersey (JER) and Guernsey (GUE), N’Dama (NDA), and Nelore (NEL) were included.

A principal component analysis (PCA) based on the genomic relationship matrix was applied to illustrate the genetic diversity among the indigenous and the Mpwapwa cattle. The extent of the admixture was investigated using unsupervised runs of the ADMIXTURE program.

Table 1. The data structure of the reference breeds.

<table>
<thead>
<tr>
<th>East African indigenous breeds</th>
<th>Abbreviation</th>
<th>Number of animals</th>
<th>Reference breeds</th>
<th>Abbreviation</th>
<th>Number of animals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ankole</td>
<td>ANK</td>
<td>43</td>
<td>Mwapwa</td>
<td>MPW</td>
<td>20</td>
</tr>
<tr>
<td>Small East African Zebu</td>
<td>SEAZ</td>
<td>58</td>
<td>Friesian</td>
<td>FRI</td>
<td>26</td>
</tr>
<tr>
<td>Ethiopia Central Highlands</td>
<td>ECHB</td>
<td>28</td>
<td>Ayrshire_New Zealand</td>
<td>AYRNZ</td>
<td>20</td>
</tr>
<tr>
<td>Ethiopia Begait Barka</td>
<td>EBB</td>
<td>30</td>
<td>Ayrshire_Canada</td>
<td>AYRCAN</td>
<td>20</td>
</tr>
<tr>
<td>Ethiopia Boran</td>
<td>EBR</td>
<td>28</td>
<td>Nelore</td>
<td>NEL</td>
<td>20</td>
</tr>
<tr>
<td>Ethiopia Danakil Harar</td>
<td>EDH</td>
<td>30</td>
<td>Guernsey</td>
<td>GUE</td>
<td>20</td>
</tr>
<tr>
<td>Ethiopia Fogera Orma Boran</td>
<td>EFO</td>
<td>28</td>
<td>Holstein</td>
<td>HOL</td>
<td>20</td>
</tr>
<tr>
<td>Kenya Sahiwal</td>
<td>KBR</td>
<td>28</td>
<td>Jersey</td>
<td>JER</td>
<td>20</td>
</tr>
<tr>
<td>TALIRI Boran (from Tanzania)</td>
<td>TBR</td>
<td>20</td>
<td>N’Dama</td>
<td>NDA</td>
<td>20</td>
</tr>
<tr>
<td>Tanzania Singida</td>
<td>TIR</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tanzania Iringa Red</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tanzanian White</td>
<td>TSW</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Results and Discussion
The plot of PC1 versus PC2 is shown in Figure 1. PC1 explains 89.3% of the variance and differentiates *Bos indicus* (represented here by Nelore) from *Bos taurus* cattle. PC2 explains 3.7% of the variance and differentiates African *Bos taurus* (represented here by N’dama) from *Bos indicus* and from European *Bos taurus*. (The next PC explains less than 1% of the variance and is therefore not plotted here.) The two reference East African breeds, Ankole and SEAZ, lie on the axis between *Bos indicus* and African *Bos taurus*, indicating that they are likely an admixture of these two types of cattle. Their relative positions indicate that SEAZ (a zebu breed) have a higher proportion of *Bos indicus* ancestry, while Ankole (a Sanga breed) appears to be approximately half African *Bos taurus* and half *Bos indicus*. Figure 1 shows that indigenous animals from Ethiopia and Tanzania closely cluster to SEAZ. The indigenous reference samples for Fogera, Danakil Harar and Ethiopian Boran very tightly clusters together, indicating that these breeds have not experienced recent admixture with imported *Bos taurus* breeds. Such animals could be used directly for genetic conservation and pure-bred genetic improvement programs. A proportion of the Ethiopian Central Highland Zebu Breed and the Begait samples lie to the right of the main indigenous cluster and spread towards the European *Bos Taurus*. This indicates that they have significant European *Bos taurus* content. The ADMIXTURE results for K=3 and K=5 (Figure 2) confirm that SEAZ, Ethiopian Central Highland Breed, Begait Barka and the TALIRI Boran have significant proportions of European *Bos taurus* ancestry. The Tanzania Iringa Red included a couple of animals that appeared to have low proportions of European *Bos taurus* genes, but the Singida White samples showed a high degree of purity. The TALIRI Boran population is very diverse and the plot is consistent with a substantial admixture of this population with European *Bos taurus* genes. The Kenyan Orma Boran forms a very tight cluster, and clusters with the Ethiopian and Tanzanian indigenous samples, including the Ethiopian Boran. The Kenyan Sahiwal population was established from animals imported to Kenya over 50 years ago from India and Pakistan. These animals cluster tightly and are very close to Nelore, indicating that this population of Sahiwal samples remain substantially pure *Bos indicus* despite the many generations of isolation from their origins and are therefore good candidates for both conservation and pure-breeding programs.

Admixture analyses attempt to identify the most likely ancestral composition of animals on the assumption that all animals are derived from a given number, K, of ancestors. For K=3 the ADMIXTURE assigns as the major ancestors *Bos indicus* (Nelore and Sahiwal), European *Bos taurus* and African *Bos taurus* (here N’dama). At K = 4, ADMIXTURE separates Jersey from other European *Bos taurus* breeds. At K=5 a new genetic signal appears which can be interpreted as an African ancestral indigenous breed signal. This finding is consistent with a study of East African Shorthorn Zebu cattle sampled from Western Kenya (Mbole-Kariuki et al. 2014), where at K=4 a common signal was observed between EASZ and Sheko, which was absent in both the Nelore and N’Dama breeds. A plausible explanation for this observation is that when admixtures of *Bos indicus* and African *Bos taurus* cattle first appeared in Africa the ancestral *Bos indicus* and *Bos taurus* breed(s) from which they descended were genetically distinct from the *Bos indicus* and African *Bos taurus* that are currently available to be used as reference breeds. In Figure 2, K=5, the Ankole (Sanga) appear as an admixture of an ancestral indigenous and African *Bos taurus*, whereas all the African zebu breeds appear as admixtures of an ancestral indigenous breed and *Bos indicus*. Compared to results with K<5, this suggests that the Sanga and zebu breeds may have shared a common indigenous ancestor that was itself an admixture of African *Bos taurus* and *Bos indicus* rather than arising from two completely separate admixture events.

Results become more difficult to interpret as K is further increased because the
evolutionary origin of breeds become increasingly likely to deviate from the assumption that all breeds arose as an admixture of K ancestor breeds. This is particularly apparent at K= 9 when Mpwapwa emerges as an ancestral breed, whereas it is known to have been formed as a recent synthetic from *Bos indicus*, African zebu and European *Bos taurus* breeds.

Some of the indigenous breeds were sampled from institutional herds and others from farmers but all animals were sampled on basis that their physical appearance was 100% pure indigenous breed. The results here show that animals that physically appear pure can contain quite high proportions of European *Bos taurus* ancestry. Luckily the same SNP assays that reveal this substantial contamination of many of the indigenous breeds can also be used to identify which animals are pure indigenous.

Figure 1. Plots of PC2 vs PC1 for individuals of East African indigenous cattle colour-coded by breed (See Table 1 for breed names).
Of interest, are the estimates of ancestral breed proportions of Mpwapwa from a supervised ADMIXTURE analysis with K=7 (5 European breeds plus N’dama and Nelore as ancestors) were 0.817 Bos indicus, 0.053 African Bos taurus and 0.129 European Bos taurus. Very similar estimates were produced by other models. These estimates are almost exactly what would be predicted from the recorded history of Mpwapwa creation (MacFarlane, 1971). The PC plot and also the admixture estimates show that, as expected for a synthetic breed, the variation among Mpwapwa animals is greater than that among pure breeds. PC and admixture analyses could be used to identify and breed from animals that conform most closely to the desired composition to achieve a more uniform population in the long term.

Figure 2. Estimation of admixture proportion of East African indigenous animals using ADMIXTURE program with K=2 to K=5 and K=11.
Conclusions

The term “zebu” has been applied to most humped cattle globally but is usually also considered synonymous with *Bos indicus*. As shown here and previously (Mbole-Kariuki *et al.*, 2014), East African humped (zebu) cattle are admixtures of *Bos indicus* and African *Bos taurus* and as such are genetically very different from pure *Bos indicus* cattle. This is important to understand because it means that genetic tests and phenotypic characteristics found in *Bos indicus* cattle cannot be assumed to apply to African zebu cattle, and *vice versa*.

The clear discrimination achieved between indigenous and non-indigenous breeds based on SNP genotypes means that existing herds and new conservation and genetic improvement programs can now ensure that they are working with populations of desired genetic background. Similarly, SNP analyses can be used to achieve desired breed compositions for synthetic breeds such as the Mpwapwa.

Although the analyses here clearly separate the Sanga from Zebu breeds, they do not provide a clear discrimination between breeds within the Zebu group. This lack of discrimination would likely be true within the Sanga group but because we only had one Sanga breed here, this still needs to be tested. It is possible, however, that alternative analyses could be developed that would provide some degree of discrimination among indigenous breeds within the Zebu and Sanga groups.

List of References


