

GENETIC CHARACTERIZATION OF LIVESTOCK POPULATIONS AND ITS USE IN CONSERVATION DECISION-MAKING

O. Hanotte and H. Jianlin

International Livestock Research Institute (ILRI), P.O. Box 30709, Nairobi 00100, Kenya
o.hanotte@cgiar.org; h.jianlin@cgiar.org

Summary

This paper reviews the importance and putative impacts of molecular markers on decision-making for livestock genetic resource conservation. Livestock diversity is shrinking rapidly and there is an urgent need to define strategies to prioritize breed conservation. For most livestock species a large number of genetic markers that show different Mendelian pattern of inheritances (maternal, paternal, bi-parental) are now available. Applied at large geographic scale to the study of livestock populations, they provide information on centers of origins and on migration routes as well as they identify geographic areas of admixture amongst populations of different genetic origins. Such information can guide the choice of breeds and geographic areas for conservation actions. Calculations of within and between diversity parameters allow selection of priority breeds for conservation to maximize diversity conserved for the benefit of future human generations.

Keywords

Genetic diversity, molecular markers, livestock breeds, conservation

Introduction

The ‘World Watch List for Domestic Animal Diversity’ report (3rd ed.) documents more than 6300 breeds of breeds of livestock belonging to 30 domesticated species [1]. These breeds were developed following domestication and natural and human selection over the past 12,000 years. The current number of breeds is likely an underestimation as a large proportion of indigenous livestock populations of the developing world, where most animal genetic resources are today found, have yet to be described at phenotypic and genetic levels. Livestock populations have evolved unique adaptation to their agricultural production system and agro-ecological environments. Their genetic diversity has provided the material for the very successful breeding improvement programs of the developed world in the 19th and 20th century. It represents a unique resource to respond to the present and future needs of livestock production both in developed and developing countries.

However, livestock diversity is shrinking rapidly. With the exception of the wild boar - the ancestor of the domestic pig - and wild red junglefowl - the ancestor of the domestic chicken - the putative wild ancestors of our major livestock species, the reservoir of genetic diversity, are now either extinct (e.g. the auroch the wild ancestor of cattle or the ancestral species of the Old World camelids) or low in numbers and threatened by extinction (e.g. wild goat populations of the Near East, vicuña from Andean plateau, wild donkey in Africa). Among the domesticated populations, it is estimated than 1 to 2 breeds are lost every week [1]. However the impact of these losses on the global or the local diversity remains undocumented. While it is already too late for many breeds in Europe, the situation is also particularly worrying in the developing world where rapid changes in production systems are leading to the replacement of breeds or at best crossbreeding. There is therefore an urgent

need to document the diversity of our livestock genetic resources and to design strategies for their sustainable conservation.

The task is enormous. It has prompted the Food and Agricultural Organization (FAO) and other international organizations to develop domestic animal diversity information systems and databases [2, 3]. More recently, FAO has initiated a major country-driven documentation exercise, the 'State of the World's Animal of Animal Genetic Resources' (AnGR), which with the publication of a companion 'Strategic Priority Action Report' document, will hopefully lead to immediate actions for conservation on the ground at the country or regional level [4].

To put this plan into practice, effective conservation of AnGR, whether *in situ* or *ex-situ*, will require the mobilisation of substantial economic resources over long period of time. Financial resources are limited and they will always be. Methods to identify priority decisions, to maximize the diversity conserved - both at the local and global level - or by focusing on unique genetic resources of global significance, are required. Genetic characterization through the use of molecular markers associated to powerful statistical approaches is providing new avenues for decision-making choices for the conservation and rational management of AnGR.

Genetic characterization tools: molecular markers

Protein polymorphisms were the first molecular markers used in livestock. A large number of studies, particularly during the 1970's, have documented the characterization of blood group and allozyme systems of livestock [e.g. 5]. However, the level of polymorphism observed in proteins is often low which has reduced the general applicability of protein typing in diversity studies. With the development of Polymerase Chain Reaction (PCR) and sequencing technologies associated with automatic and/or semi-automatic large scale screening system, DNA-based polymorphisms are now the markers of choice for molecular-based surveys of genetic variation. Importantly, polymorphic DNA markers showing different patterns of Mendelian inheritances can now be studied in nearly all of our major livestock species. Typically, they include *D-loop* and cytochrome *B* mitochondrial DNA (mtDNA) sequences (maternal inheritance), *Y* chromosome specific single nucleotide polymorphism (SNP) and microsatellites (paternal inheritance), autosomal microsatellite (bi-parental inheritance) [6]. Interestingly, while recent developments in cytogenetic technologies should facilitate the isolation of *Y* chromosomes specific markers [7], for most livestock species there are still few *Y* polymorphic markers. This is possibly a consequence of the demographic history of domestication and breed formation. In polygenous species, like most livestock, we expect indeed that a small number of male lineages would have contributed to the genetic pool of the species. Polymorphic *Y* microsatellite markers are currently only available for cattle [8], yak [9] and to some extent small ruminants (J. A. Lenstra workshop presentation). They have not been yet isolated, to the best of our knowledge, in some major livestock species, e.g. the Old and New World camelids or the domestic pig. On the other hand autosomal microsatellites have now been isolated in large numbers from most livestock species and recommended FAO/ISAG lists of autosomal microsatellite markers for genetic characterization studies are publicly available [10].

Important assumptions on the use of genetic markers include: (i) that the polymorphisms observed at the molecular markers are neutral; (ii) that the use of a relatively small number of independently segregating marker loci will be a good predictor of the overall genomic diversity of a population; in other words that variation in allele frequencies between populations will reflect the distribution of genetic diversity within and amongst populations.

A diversity of genetic characterization information

Genetic markers will provide different types of information relevant for conservation-making decisions for livestock [11]. Autosomal microsatellite loci will be commonly used for individual genetic identification and parentage analysis, e.g. for the successful implementation and monitoring of *ex-situ* conservation program, population diversity estimations, differentiation of populations, calculation of genetic distances, genetic relationships and population genetic admixture estimation. Microsatellite loci are also highly sensitive to genetic bottlenecks and they are commonly used for inbreeding estimation. MtDNA sequences will be the markers of choice for domestication studies as the segregation of a mitochondrial DNA lineage within a livestock population will only have occurred through the domestication of a wild female or through the incorporation of a female into the domestic stock. More particularly, mtDNA sequences will be used to identify the putative wild progenitors, the number of maternal lineages and their geographic origins. To some extent it may provide important information on the geographic distribution of diversity within livestock species although the usefulness of mtDNA sequences data will vary between species, depending of the demographic history of the migration from the center(s) of domestication. Last but not least the study of a diagnostic Y chromosome polymorphism is an easy and rapid way to detect and to quantify male-mediated admixture.

A surprising result of the application of molecular markers in genetic characterization of livestock has been the discovery that several ancestral species, subspecies or maternal lineages have contributed to today's genetic pool of our major livestock species [9, 12, 13, ILRI unpublished data]. It is clear from these recent results that multiple domestications and/or maternal introgression are the rule not the exception (Table 1).

Domestic species	Number of maternal lineages	Geographic centers of origin
Cattle		
<i>Bos taurus</i>	2	Near East, northeast Africa
<i>Bos indicus</i>	1	South Asia (Indus Valley)
Yak		
<i>Bos grunniens</i>	3	Hindu-Kush Himalayan region
Sheep		
<i>Ovis aries</i>	3	Near East, south Asia
Goat		
<i>Capra hircus</i>	A least 3	Near East, south Asia
Horse		
<i>Equus caballus</i>	Multiple	North and Central Asia
Donkey		
<i>Equus asinus</i>	2	Northeast Africa
Pig		
<i>Sus scrofa domestica</i>	2	Near East, east Asia
Water Buffalo		
<i>Bubalus bubalis bubalis</i>	1	South Asia
<i>Bubalus bubalis carabensis</i>	1	East Asia
Llama		
<i>Llama glama</i>	4	Northern Andean chain
Alpaca		
<i>Vicugna pacos</i>	4	Northern Andean chain

Table 1. Number and putative centers of origin of major maternal lineage in livestock
 From references [12] and [9 -yak, 13-Donkey]

More particularly, mtDNA information supports the conclusion that there were at least five major centers of livestock domestication: the northern Andean chain (New World camelids), the northeast African region (donkey and likely taurine cattle), the Near East (taurine cattle, sheep, goat, pigs), south Asia (Indus Valley, indicine cattle and chicken) and East Asia (pigs, chicken, horse, buffalo) to which should be added the Hindu-Kush Himalayan region (yak) and North and Central Asia (horse). Similarly, we could expect to some extent multiple male-mediated introgression lineages. It is the case in yak [9] with two distinct male lineages but not in the domestic horse where screening for SNP in 52 stallions from 15 different breeds did not identify a single polymorphic site [14].

Taking into account the history of human migration and trading, it is expected that our indigenous breeds of today will often have multiple genetic signatures of origin and admixture. Available molecular data indicate that ancient genetic admixtures between livestock populations from different domestication events are common on the Asian continent and to some extent also in Africa, but mostly absent from Europe. This has been shown in cattle [e.g. 15] in pigs and small ruminants [12].

Genetic characterization and conservation decision-making

As illustrated above genetic characterization is providing new information to guide and prioritize conservation decisions for livestock. Possibly, the most urgently required action is the effective protection of all remaining wild ancestral populations and closely related species of livestock, most of them now endangered. They are the only remaining sources of putative alleles of economic values that might have been lost during domestication events. Coordination with international wildlife conservation institutions (such as IUCN) is required. It is equally important to ensure that the breeds selected for conservation include populations from the geographic areas representing the different domestication centers where we would expect to find large genetic diversity and genetically differentiated populations. Animals and populations present at the geographic area of a center of domestication will also be expected to be very distinct from the ones found at other centers of domestication. Also, the understanding of the geographic pattern of livestock migration from a center of origin will allow the identification of populations present at the end of a migration route. It is expected that these populations will be genetically distinct from the populations present at the ends of other migration routes as a result of random genetic drift and/or the effect of local selection pressures. Importantly, knowledge of both the global diversity of the breeds and admixture events will be needed in order to be able to make sound priority decisions.

The next challenge is to make priority decisions for conservation amongst today's thousands of domestic breeds or populations. The primary objective is to maximize the conservation of the genetic diversity available for potential future use. At the ideal extreme, this would be achieved through the conservation of all breeds of livestock. Such a comprehensive approach would ensure complete conservation of diversity. In practice, it is unrealistic and prioritization of actions will have to be made. Two criteria (perhaps to be eventually combined) have been proposed [16]: priority breeds for conservation should be the ones with the largest within-breed diversity and/or should maximise the conservation of between breed diversity. Both within and between breed diversity parameters are classically measured using molecular genetic markers. In both cases soundly-based priority decisions for conservation at the global level will require the availability of large datasets.

The mean number of alleles (*MNA*) and observed (*Ho*) and expected (*He*) heterozygosity are the most commonly calculated population genetic parameters for assessing within breed diversity. For example, in a recent study, three distinct sets of microsatellite diversity cattle data were merged to provide for the first time within breed diversity (*He* and *MNA*) and admixture information combined for Europe, Africa, the Near East and South Asia [17]. The geographic region with the highest diversity is found between the two likely Asian centers of cattle domestication in a broad geographic area corresponding to what are today Iran, Iraq and the Caucasian region. Global geographic analysis of admixture suggests that the region corresponds to a geographic area of around 50% admixture between taurine and indicine cattle. Genetic diversity and admixture information from more indigenous breeds are needed to confirm the results. If it is confirmed, this geographic area will undoubtedly represent a major livestock diversity hotspot, a priority region for a global plan for the conservation of the diversity of domestic cattle.

The simplest parameters for assessing the distribution of diversity between breeds using genetic markers are the genetic differentiation or fixation indices (e.g. F_{ST} , G_{ST} , θ). The most widely used is F_{ST} which measures the degree of genetic variation between subpopulations through the calculation of the standardised variances in allele frequencies amongst populations [18, 19]. Any set of genetic distances can also be analyzed in terms of between breed genetic diversity and more particularly in terms of individual breed contributions to the total diversity of a set of breeds. The most commonly approach used so far is a method proposed by Weitzman [20]. It involves calculation of a matrix of genetic distances and generation of dendrograms. Individual breed contributions are calculated by comparing the total length of the dendrogram including all breeds with the dendrogram including all breeds less the individual breed. Priority breeds for conservation would therefore be the breeds contributing most to the diversity of the set. The method can be extended further to estimate the impact of conservation decisions on the diversity of a set of breeds in the future with the calculation of the extinction probability of each breed and the marginal diversity which reflect the relative lost (gain) in expected diversity of a set of breeds following a decrease (increase) in the probability of survival of a breed by one unit.

The largest dataset to which the Weitzman approach has been applied in livestock is on 49 African cattle breeds [21]. The breeds were divided into two groups corresponding to the 'taurine' and 'indicine' division and extinction probabilities were calculated for each breed. In both groups, the results clearly indicate that the optimal conservation strategy is to give priority to the breeds with the highest marginal diversity rather than the most endangered ones.

Future challenges and opportunities

Major challenges remain for livestock conservationists. Documentation of genetic diversity is still all but lacking for some livestock species and incomplete for others (e.g. Old World camelids, chicken, buffalo, Asian small ruminants and cattle etc). With a few exceptions [e.g. 15] molecular datasets will include in a single study only a limited number of breeds or populations, and combined analysis of molecular datasets obtained in different studies will often be impossible. For example, we are still lacking statistical approaches that allow the combination of genetic-distancing information obtained in separate studies. Last but not least, the molecular markers used to characterise diversity have little to do with the genes under selection for economically important traits [16].

But there are promising and exciting new avenues. We can expect that with the increased adoption of common sets of markers and common breeds of references [17], the combination of microsatellite datasets will be facilitated. The publication of the entire genome sequences of several livestock species will allow the easy identification of thousands of neutral and selected genetic markers. It will open the way to the detection of signatures of selection allowing researchers to trace the presence and the spread of economically important alleles [22]. A recent study in cattle milk protein genes has indirectly, yet nicely, illustrated the putative application of such selected markers in the identification of breeds and geographic areas as priorities for the conservation of specific economically important traits [23]. The new field of livestock 'landscape genetics' is emerging [24]. It will combined geo-referencing of breed distributions, spatial global genetic diversity, climatic, ecological, epidemiological and production system information which will facilitate and direct priority decisions for *in situ* breed conservation.

THE ROLE OF BIOTECHNOLOGY
Villa Gualino, Turin, Italy – 5-7 March, 2005

REFERENCE LIST

- [1] Schearf, B. (ed.) World Watch List for Domestic Animal Diversity. FAO 2003.
- [2] ILRI Domestic Animal Genetic Information System (DAGRIS). [Http://Dagris.ILRI.cgiar.org/dagris/](http://Dagris.ILRI.cgiar.org/dagris/)
- [3] DAD-IS 2.0 <http://www.FAO.org/dad-is>
- [4] [Http://dad.fao.org/cgi-dad/\\$cgi_dad.dll/sow?eng](http://dad.fao.org/cgi-dad/$cgi_dad.dll/sow?eng)
- [5] Baker, C. M. A and C. Manwell 1980. Chemical classification of cattle. I. Breed group. *Animal Blood Groups and Biochemical Genetics* 11, 127-150.
- [6] Avise, J. C. 1994. *Molecular markers, natural history and evolution*. Chapman and Hall, New York.
- [7] Petit, E. et al. 2002. Mammalian population genetics. Why not Y. *Trends in Ecology and Evolution* 17, 28-33.
- [8] Hanotte et al. 2000. Geographic distribution and frequency of a taurine *Bos taurus* and an indicine *Bos indicus* Y specific allele amongst sub-Saharan Africa cattle breeds. *Molecular Ecology* 9, 387-396.
- [9] Xuebin, Q. 2004. Genetic diversity, differentiation and relationship of domestic yak populations: a microsatellite and mitochondrial DNA study. PhD Thesis. 262 Pages.
- [10] <http://dad.fao.org/en/refer/library/guidelin/marker.pdf>
- [11] Sunnucks, P. 2000. Efficient genetic markers from population biology. *Trends Ecology and Evolution* 15, 199-203.
- [12] Bruford, W. et al. 2003. DNA markers reveal the complexity of livestock domestication. *Nature Review Genetics* 4, 900-910.
- [13] Beja-Pereira, A. 2004. African origins of the domestic donkey. *Science* 304, 1781.
- [14] Lindgren G. et al. 2004. Limited number of patriline in horse domestication. *Nature Genetics* 36, 335-336.
- [15] Hanotte, O. et al. 2002. African pastoralism: genetic imprints of origins and migrations. *Science* 296, 336-339.
- [16] Gibson, J. et al. (in press). Measures of diversity as inputs for decisions in conservation of livestock genetic resources. CABI
- [17] Freeman et al. (in preparation). Combination of multiple microsatellite datasets to investigate genetic diversity and admixture.
- [18] Weir, B. S. and C. J. Basten 1990. Sampling strategies for distances between DNA sequences. *Biometrics* 46, 551-582.
- [19] Mburu, D. N. et al. 2003 Genetic diversity and relationships of indigenous Kenyan camel (*Camelus dromedarius*) populations: implications for their classification. *Animal Genetics* 34, 26-32.
- [20] Weitzman, M.L. 1993. What to preserve? An application of diversity theory to crane conservation. *Quarterly Journal of Economics* 108, 157-183.
- [21] Reist-Marti, S. et al. 2003. Weitzman's approach and conservation of breed diversity: an application to African cattle breeds. *Conservation Biology* 17, 1299 – 1311.
- [22] Luikart, G. et al. 2003. The power and promise of population genomics: from genotyping to genome typing. *Nature Review Genetics* 4, 981-994.
- [23] Beja-Pereira, A. et al. 2003. Gene-culture coevolution between cattle milk protein genes and human lactase genes. *Nature Genetics* 35, 311-313.
- [24] Manel, S. 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution* 18, 189-197.