ENGAGEMENT PAPER

The Draft Genome of Extinct European Aurochs and its Implications for De-Extinction

Mikkel-Holger S. Sinding*† and M. Thomas P. Gilbert*‡§

Whether as a cave painting, a mounted skeleton in a museum, or as described in ancient texts, the extinct aurochs has long mesmerized humans. In the context of genetics, aurochs have been targeted since the early days of the establishment of ancient DNA techniques, and for two decades analyses of its mitochondrial genome have considerably deepened our knowledge of this animal. These studies have produced major discoveries, such as how cattle were domesticated from aurochs through at least two separate events. However, answers to many other aspects of its evolutionary history require more than the sequence from a single non-recombining marker such as the mitochondrial genome. Of these questions, perhaps one of the most fascinating is whether domestic cattle and wild aurochs continued to cross-breed following the initial domestication event, and if so, to what extent? Resolving this question would provide valuable new insights into how our ancestors domesticated cattle and subsequently manipulated their gene pool. In addition, it will become increasingly relevant as we enter the de-extinction debate. Are we witnessing the recovery of information that might, in the not too distant future, allow the re-creation of aurochs through selective back breeding of carefully chosen modern cattle lineages?

Keywords: De-extinction; back breeding; ancient genomics; aurochs-cattle hybridization

Introduction – aurochs genetics in the pre-genomic era

The geographic range of the aurochs (Bos primigenius) was once considerable (Figure 1). Its Holocene distribution spanned North Africa, the Indian subcontinent, and over the vast area encompassed by Europe, West and Central Asia (north of the Himalayas) as far as the Asian Pacific coast (Van Vuure, 2005). Although once widespread their range steadily collapsed, likely as a consequence of human pressure, with their last stronghold thought to have been in a royal hunting reserve in an area that is today Poland (Rokosz, 1995). Here aurochs went extinct in the 1600s, with the last known bull recorded to have died in 1620 and the last cow in 1627 (Rokosz, 1995).

Given this relatively late persistence and extensive zooarchaeological material (Wright and Viner-Daniels, 2015), ancient DNA (aDNA) studies of aurochs have been heavily biased towards the European population. Indeed, a focus of many of the aurochs aDNA papers published to date has been whether there was introgression from European wild aurochs into domestic cattle (see below for a short summary of this discussion). For technological reasons, until very recently such studies were limited to the study of mitochondrial and Y-chromosome markers. While they provided some tantalizing glimpses into the question, these single-locus studies left many questions unanswered. For example, analyses that used Y-chromosomes to document possible introgression of male aurochs into domestic cattle have yielded inconclusive results (Bollongino et al., 2008; Götherström et al., 2005; Pérez-Pardal et al., 2010; Svensson and Götherström, 2008). In contrast, a clearer picture has been obtained through the analysis of mitochondrial data, which provide not only insight into the phylogeography of aurochs as well as the timing and number of cattle domestication events, but also potential reproduction events between wild aurochs and domestic cattle.

In particular, while mtDNA evidence suggests that the domestication of aurochs originally occurred through at least two separate events, central European, Italian, and Balkan aurochs likely also contributed to some modern cattle lineages (Achilli et al., 2009, 2008; Anderung et al., 2005; Hristov et al., 2015a; Schibler et al., 2014). Despite these observations, questions remain relating to the levels, timing, and location of subsequent aurochs introgression into domesticated cattle. Answering these questions will
Mitochondrial studies on both modern and ancient cattle, as well as ancient aurochs remains, have revealed seven mitochondrial clades, C, E, I, P, Q, R and T (Figure 1). Overall, mitochondrial evidence is consistent with analyses undertaken on modern cattle samples using nuclear DNA markers (Gibbs et al., 2009; Troy et al., 2001), and points towards a domestication of aurochs through at least two separate events, one resulting in taurine cattle (Bos taurus) that derive from Near Eastern aurochs, while a second resulting in zebu cattle (Bos indicus) that derive from South Asian aurochs (Loftus et al., 1994; Machugh et al., 1997). When one examines the mitochondrial data in more detail, three major trends can be observed. Firstly, the majority of modern taurine cattle carry lineage T, believed to derive from the first domesticated Near Eastern aurochs. Secondly, all zebu cattle carry lineage I. The distinct sister relationship of this clade to all other lineages is as expected, given zebu's independent domestication. Thirdly, with the exception of one German aurochs that carried its own unique lineage (E), all remaining northern, central and eastern European aurochs studied to date carried lineage P (Achilli et al., 2009; Edwards et al., 2007; Lari et al., 2011). The P lineage has also been identified in a small number of modern taurine breeds and ancient European domestic cattle, suggesting introgression of northern, central and/or eastern European aurochs (Achilli et al., 2008; Anderung et al., 2005; Schibler et al., 2014).

Intriguingly, analyses of Italian cattle breeds and aurochs have suggested that Italy conceivably was home to an additional domestication, or at least introgression, event. While the above mentioned T lineage was present in Italian aurochs pre-farming (Lari et al., 2011), Italian cattle also carry lineage R, which while still unidentified in any aurochs population studied to date, could logically derive from Italian aurochs. Furthermore, the rare Q lineage is principally today found in several modern Italian and Egyptian breeds, but has also been found in both a small number of other taurine breeds, and Neolithic European cattle, indicating a origin in the Near East along with the T lineage (Achilli et al., 2009; Bonfiglio et al., 2010; Olivieri et al., 2015). A specific lineage of T has also been observed in South Eastern European aurochs (Hristov et al., 2015), and is known to have survived into modern Balkan cattle (Hristov et al., 2015b; Kantanen et al., 2009) thus suggesting an important role in the region as either an additional domestication centre, or a location in which introgression was happening. Finally, there is lineage C, which to date has only been found in China, recovered from a 10,660 year-old presumably domestic cattle specimen. The uniqueness of this lineage may point to a separate domestication of cattle from aurochs in the Far East (Zhang et al., 2013).

In contrast to the wealth of results derived from mainland European remains, very few genetic results have been reported for Near Eastern aurochs. Indeed, the total dataset published to date is represented by the mitochondrial d-loop sequence of only a single Syrian Bos specimen from the early Neolithic that is believed to be either a true aurochs or early domesticated cattle (Edwards et al., 2007). Furthermore, at this time the genetic background of aurochs from Iberia, and for its entire North African range is completely absent. With the exception of the insight gained from the single Chinese sample, this is also the case for our understanding of the ancient populations in Central, South and East Asia.

**Introgression from European wild aurochs into domestic cattle**

Mitochondrial studies on both modern and ancient cattle, as well as ancient aurochs remains, have revealed seven mitoinelineages, C, E, I, P, Q, R and T (Figure 2A). Overall, mitochondrial evidence is consistent with analyses undertaken on modern cattle samples using nuclear DNA markers (Gibbs et al., 2009; Troy et al., 2001), and points towards a domestication of aurochs through at least two separate events, one resulting in taurine cattle (Bos taurus) that derive from Near Eastern aurochs, while a second resulting in zebu cattle (Bos indicus) that derive from South Asian aurochs (Loftus et al., 1994; Machugh et al., 1997). When one examines the mitochondrial data in more detail, three major trends can be observed. First, the majority of modern taurine cattle carry lineage T, believed to derive from the first domesticated Near Eastern aurochs. Second, all zebu cattle carry lineage I. The distinct sister relationship of this clade to all other lineages is as expected, given the independent domestication of zebu. Third, with the exception of one German aurochs that carried its own unique lineage (E), all remaining northern, central and eastern European aurochs studied to date carried lineage P (Achilli et al., 2009; Edwards et al., 2007; Lari et al., 2011). The P lineage has also been identified in a small number of modern taurine breeds and ancient European domestic cattle, suggesting introgression of northern, central and/or eastern European aurochs (Achilli et al., 2008; Anderung et al., 2005; Schibler et al., 2014).

Intriguingly, analyses of Italian cattle breeds and aurochs have suggested that Italy conceivably was home to an additional domestication, or at least introgression, event. While the above mentioned T lineage was present in Italian aurochs pre-farming (Lari et al., 2011), Italian cattle also carry lineage R, which while still unidentified in any aurochs population studied to date, could logically derive from Italian aurochs. Furthermore, the rare Q lineage is principally today found in several modern Italian and Egyptian breeds, but has also been found in both a small number of other taurine breeds, and Neolithic European cattle, indicating a origin in the Near East along with the T lineage (Achilli et al., 2009; Bonfiglio et al., 2010; Olivieri et al., 2015). A specific lineage of T has also
Figure 1: Map of the original wild aurochs populations. Letters refer to mitochondrial haplogroups confirmed by ancient DNA, squares refer to unique variation observed at specific location, and haplogroups followed by '?' refer to haplogroups where the locality of origin is not based on sequencing of aurochs remains, but representation in either modern or ancient domestic cattle. Areas shaded in grey, labelled with '?' designate unknown population affiliation and unknown original mitochondrial linages. The "C*" implicates C is not found in any genome sequenced cattle or aurochs.

Figure 2: A. Nuclear genome-based phylogeny of aurochs and cattle, simplified from Park et al. (2015) and Decker et al. (2014). Embedded colours correspond to assumed ancient source populations shown in figure 1. B. Phylogeny of the full mitochondrial lineages in aurochs and cattle with colours in branches corresponding to figure 1, simplified from Zhang et al., (2013). Mitochondrial lineage E, which is only known from a short fragment of the control region (Edwards et al., 2007), is not shown in the figure, the "C*" implicates C is not found in any genome sequenced cattle or aurochs.
been observed in South-Eastern Balkan aurochs (Hristov et al., 2015a), and is known to have survived in modern Balkan cattle (Hristov et al., 2015b; Kantanen et al., 2009) thus suggesting an important role in the region as either an additional domestication centre, or a location in which introgression was happening. Finally, there is lineage C, which to date has only been found in China, recovered from a 10,660 year-old, presumably domestic cattle specimen. The uniqueness of this lineage may point to a separate domestication of cattle from aurochs in the Far East (Zhang et al., 2013).

In contrast to the wealth of results derived from mainland European remains, very few genetic results have been reported for Near Eastern aurochs. Indeed, the total dataset published to date is represented by the mitochondrial d-loop sequence of only a single Syrian Bos specimen from the early Neolithic that is believed to be either a true aurochs or early domesticated cattle (Edwards et al., 2007). Furthermore, at this time the genetic background of aurochs from Iberia, and for its entire North African range is completely absent. With the exception of the insight gained from the single Chinese sample, this is also the case for our understanding of the ancient populations in Central, South and East Asia.

**Introgression in light of a British aurochs genome**

Within the last decade, the field of ancient DNA has evolved into that of ancient genomics. Following the release of the first de novo sequenced taurine cattle genome (Elsik et al., 2009), the possibility of generating a near-complete aurochs genome sequence suddenly became plausible. In particular, researchers began to hope they would be able to recover sufficient numbers of aDNA fragments from ancient aurochs remains, which would enable subsequent mapping to this reference genome. And this is exactly what Park and colleagues (2015) have achieved – they obtained sufficient levels of mappable aDNA to yield a 6.23x coverage genome from a 6,750 year-old British aurochs. Although aurochs persisted until much more recently, the age of this individual is notable given that it died approximately one thousand years before Neolithisation in Britain (Stevens and Fuller, 2012). Thus, this particular aurochs can be confidently assumed to represent a “pure” sample, free from possible genomic contamination through later cross-breeding with domestic cattle.

With this data in hand, the authors were subsequently able to analyse it against a genome-wide dataset drawn from 73 modern cattle populations, and clearly demonstrate nuclear introgression of European aurochs into taurine cattle. Their powerful nuDNA-based analyses show that British breeds such as Highland, Dexter, Kerry, Welsh Black and White Park carry ancestry from this aurochs population, while the non-British breeds studied do not.

This finding of multiple local introgression events has at least three key implications. First, it settles a decade-long debate about the significance of mixing between domestic cattle and aurochs. Second, evidence of local admixture suggests that several different subpopulations of aurochs may have contributed to the genetic background of different Old World cattle breeds. And third, in the genomic context at least, the aurochs remains at least partially with us, embedded within the broader genetic background of modern cattle. In this regard, a major outstanding question is therefore how many local modern cow breeds contain genetic remnants of the aurochs nested within their own genome; and thus, what total fraction of the aurochs’ genome may still exist? While there are many possible answers, one that is particularly topical given the current climate in which ‘de-extinction’ projects are starting to receive considerable attention (Barnett, 2016; Richmond et al., 2016), is that if aurochs genetic material still exists preserved in these genomes, it may support the possibility of one day bringing back this species.

**The dawn of de-extinction**

The aims of de-extinction projects are to literally re-create biological entities that are extinct. Although currently there are several targeted species, including the woolly mammoth (Mammuthus primigenius) and the passenger pigeon (Ectopistes migratorius) (Sherkow and Greeley, 2013), the aurochs was probably the original species to be chosen for de-extinction efforts. Specifically, the idea that aurochs could be back-bred was initiated by two German brothers, the Hecks, in the 1920s, through what they described as “new breeding”. Specifically, they argued that as cattle are the direct descendants of aurochs, aurochs are not actually extinct, just physically modified. Thus, they argued, it might be possible to reverse the changes incurred through selective crossing of modern cattle breeds that retain partial aurochs-like phenotypes in a way that gradually collects all the phenotypes in one resulting breed (Heck, 1951; Van Vuure, 2005).

At this point we should highlight that the exact definition of a successful “de-extinction” is a controversial subject (Richmond et al., 2016). For example, one could attempt to create a modified animal that fulfils the same ecosystem function as a now extinct form, while not exhibiting the fully accurate phenotype. Alternatively, one could focus on recreating the lost phenotype, while ignoring its behaviour or ecosystem role (Zimov, 2005; Sherkow and Greeley, 2013; Swart, 2014; Shapiro, 2015a). In the case of the aurochs, the Heck brothers took the latter approach, attempting to selectively cross cattle to recreate what they believed to be the original aurochs phenotype (as based on historical descriptions). A dozen years later the resulting animal, the so-called ‘Heck cattle breed’, bore some similarities to what experts believe the aurochs may have looked like. These similarities were sufficient that the Heck brothers even announced publically that their “new breeding” attempt had been successful (Van Vuure, 2005), although today few regard the achievement as any more than partial. Indeed, there are presently several modern projects attempting to recreate phenotypically accurate aurochs in this way, for example the Tauros and Uruz projects (Stokstad, 2015; Tauros Project, 2016).
The obvious weakness of the Heck brothers’ approach is that aurochs-like phenotype does not necessarily equate to aurochs-like genotype, something which would be critical under a pure interpretation of de-extinction. An attractive solution to this issue is to exploit the complete genome sequences of extinct species in their de-extinction attempts (Shapiro, 2015b). Such an approach is, in fact, already in use in the aurochs context. The True Nature Foundation, which is driving the Uruz aurochs back-breeding project, has stated that it will be using the recently published aurochs genome (Park et al., 2015) as a template for breeding (Stokstad, 2015), or for genome editing of hand picked individuals in the breeding program (True Nature Foundation, 2016). If one ignores the significant question of whether genome editing techniques such as CRISPR-CAS9 (Baker, 2012; Hsu et al., 2014; Ran et al., 2013a; 2013b) are sufficiently mature yet to enable this, there are two other key questions that require addressing. First, to what degree is the genome of a single British aurochs the ideal representative of the species? And second, to what extent can we resolve the complexity of the domestication history that we can be sure of maximising the lost aurochs diversity?

With regards to the first point, zooarchaeological evidence suggests that the extinction of aurochs involved a decline spanning roughly 10,000 years, over a geographic range that encompassed East to West Eurasia and North Africa (Van Vuure, 2005). Thus considerable diversity may have been lost even prior to domestication. As for the Park et al. (2015) genome in particular, it is worth noting that as it comes from Britain, it represents an island population that would have been cut off from the population of continental Europe and thus may possibly be quite distinct. We highlight this fact not to diminish the aims of the Uruz project, but to highlight the importance of sequencing additional aurochs genomes drawn from the considerable morphological and geographic range of variation within aurochs (Wright and Viner-Daniels, 2015).

With regards to the second point, Park et al.’s (2015) analyses contribute to our insights into the larger scale relationships between cattle and aurochs. Previous analyses on modern cattle datasets have demonstrated how modern cattle cluster into three major groups: the African breeds, zebu cattle and European cattle (Decker et al., 2014, 2009; Gautier et al., 2010; Gibbs et al., 2009) (Figure 2A). Whether Park et al. (2015) compare the aurochs data to between ca. 7000 and 13000 high quality SNPs drawn from 278 and 1225 cattle individuals respectively, or 81 genomes from 11 different B. taurus and B. indicus breeds, their findings remain consistent with the prior analyses. Specifically, domestic cattle remain subdivided into the three principal clusters of diversity and the aurochs is a fourth lineage, the zebu is the most divergent group to this the aurochs is sister to all taurine breeds, and African taurine cattle are still a highly divergent group to the European taurine cattle.

Local domestication hypotheses for the origin of African cattle have previously been proposed so as to explain traits unique in African cattle (e.g. Bradley et al., 1996; Gautier et al., 2010; Grigson, 1991; Hanotte et al., 2002; Linseele et al., 2009; Marshall and Weissbrod, 2011; Pérez-Pardal et al., 2010; Stock and Gifford-Gonzalez, 2013; Troy et al., 2001). A recent paper using enormous amounts of cattle SNP data deserves to be highlighted in relation to both the question of African domestication and Park and colleague’s (2015) results. In 2014, Decker and co-authors published the largest population sampling of any non-human mammalian species, containing data across 43,043 SNPs typed in 1,543 individuals spanning 134 domesticated bovid breeds. In their paper, they argue that the high amount of observed distinct genetic diversity in African taurine cattle is due to as much as a 26% African aurochs ancestry component derived from introgression, rather than a distinct third, African domestication event. However, while the authors highlight the importance of African aurochs skeletal material in ultimately settling the question of amount of cattle and aurochs admixture in Africa, the result exemplifies how the “extinct” African aurochs also partially remain within modern cattle breeds.

In Asia, several bovines have been domesticated, namely water buffalo (Bubalus bubalis), yak (Bos grunniens), gaur (Bos gaurus), and banteng (Bos javanicus) (Guo et al., 2006; Kumar et al., 2007; Machugh, 1996). Yak, gaur and banteng are closely related to B. taurus and B. indicus (Decker et al., 2009), and mitochondrial data indicates some admixture between Asian cattle and yaks (Kikkawa et al., 2003; Ward et al., 1999). Other molecular data indicates mixing between Asian cattle and banteng. Specifically, introgression of banteng into zebu has been shown using mitochondrial, Y-chromosomal, amplified fragment length polymorphism, satellite fragment length polymorphisms, restriction fragment length polymorphism of satellite DNA and microsatellite genotyping (Kikkawa et al., 2003; Mohamad et al., 2009; Nijman et al., 2003; Verkaar et al., 2003). Decker and colleagues (2014) and Hartati and coauthors (2015) also showed that the banteng is the partial ancestor to some zebu breeds, thus providing an excellent example of how modern breeds can be carrying genetic material that falls outside of the classical taurine and zebu domestication events.

A further relevant study that builds on the data released by Decker et al. (2014) is that of Wangkumhang et al. (2015). Through the addition of extra SNP data generated from 28 cattle individuals spanning 4 Thai breeds, these authors find that South East Asian zebu carry a distinct, previously unknown ancestral genetic component that appears restricted to local breeds. Given this, the authors argue for a source of variation deriving from an ancestor outside of the original zebu and taurine domestication events. Given Park et al.’s (2015) observations, one tantalising possibility could be that the source was an as yet unanalysed Asian aurochs population – something that we can hope future studies on Asian aurochs material might address. Ultimately, a clear lesson from the Park et al. study is that to resolve cattle domestication further, more ancient genomes are a
must. While once a daunting challenge, the possibility of ancient population genomic level studies is rapidly becoming a reality in evolutionary biology (e.g. Allentoft et al., 2015; da Fonseca et al., 2015; Orlando et al., 2015; Raghavan et al., 2014; Skoglund et al., 2012), thus more aurochs genomes will almost certainly be released in the near future.

Conclusion – several aurochs are only partially extinct

In summary, it is incontrovertible that (i) at least two geographically distinct populations of aurochs gave rise to the two extant domestic cattle groups (Gibbs et al., 2009; Loftus et al., 1994; Machugh et al., 1997; Troy et al., 2001), that (ii) these were subsequently dispersed across the full former range of aurochs, that (iii) in many places aurochs and cattle coexisted, and finally (iv) they could potentially have interbred. Park et al. (2015) clearly demonstrate introgression of local aurochs into British cattle, and should this not have been an isolated event but reflect the bigger picture across the aurochs/cattle range, perhaps several sup-populations of aurochs are not extinct at all. With an improved understanding of the aurochs’ phylogeny we are entering an era in which it will be possible to consider breeding back Bos that are genetically akin to specific original aurochs populations, through selective cross-breeding of local cattle breeds bearing local aurochs-genome ancestry.

Acknowledgments

We thank Mia Mottelson and Gustav Pontoppidan for the illustrations. ‘The authors acknowledge ERC Consolidator Grant 681396-Extinction Genomics for funding their work’.

Competing Interests

The authors declare that they have no competing interests.

References


Shapiro, B 2015b How to clone a mammoth: the science of de-extinction. Princeton University Press. DOI: http://dx.doi.org/10.1515/9781400865482.


Stevens, C J and Fuller, D Q 2012 Did Neolithic farming fail? The case for a Bronze Age agricultural revolution in the British Isles. *Antiquity,*


How to cite this article: Sinding, M-H S and Gilbert, M T P 2016 The Draft Genome of Extinct European Aurochs and its Implications for De-Extinction. Open Quaternary, 2: 7, pp. 1–9, DOI: http://dx.doi.org/10.5334/oq.25

Submitted: 18 April 2016  Accepted: 13 September 2016  Published: 27 September 2016

Copyright: © 2016 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.

Open Quaternary is a peer-reviewed open access journal published by Ubiquity Press.