Genomics: Captive Breeding and Wildlife Conservation

Alfred L. Roca
Lawrence B. Schook
Department of Animal Sciences, University of Illinois, Urbana, Illinois, U.S.A.

Abstract
The genomes of many domestic and wild mammals are being sequenced in order to enhance agricultural productivity, provide insights into biomedical model organisms, and identify conserved functional regions of the human genome. Sequencing of these species "genome-enables" other species within the same genus, family, or order by facilitating conservation genetic studies on wild, and often endangered, relatives of the sequenced species. Sequenced genomes have enabled researchers to examine whether interbreeding with domestic species has affected the germline of related wild species. Genomic sequences improve the ability of researchers to make inferences about neutral genetic variation within species; to find deleterious alleles, which can be purged through selective breeding programs; and to identify adaptive alleles. The sharply dropping cost of sequencing technology may allow the genomes of thousands of species to be sequenced, assisting conservation and captive breeding efforts on their behalf.

INTRODUCTION

Sequencing the human genome has greatly assisted researchers seeking to comprehend the genetic basis of health and disease. In order to properly understand, annotate, and interpret human genomic information, a comparative approach using the genomic sequences of other mammals is being used to elucidate the structure, function, and regulation of genes and other genomic elements.[1] For example, genomic or phylogenetic "shadowing" has been proposed as a means of identifying functional elements within the human genome.[2] In phylogenetic shadowing, DNA sequences from multiple species are aligned and regions that are conserved as well as non-conserved regions are identified (Fig. 1).[2] Regions conserved across species tend to have a functional role, such as the exons that code for the amino acids of proteins, or other elements that play an important role in regulating the transcription or expression of genes. Thus, genomic or phylogenetic shadowing has been used to discover and characterize functional components of the human genome, such as novel gene regulatory elements.[2]

The initial selection of mammalian species recommended for sequencing (Fig. 2) largely relied on the role of the species as important agricultural animals (livestock); as model organisms for biomedical studies (rodents, primates, livestock); as veterinary models for human diseases (cat, dog, horse); or as species important for comparative studies of recent human evolution (primates and their relatives) or of evolutionary events affecting all placental mammals (afrotherians such as the savanna elephant or xenarthrans such as the sloth).[1] In addition, some species were recommended for genomic sequencing in order to further studies on basic evolutionary questions, such as bats (reduced genome size) or insectivores (retention of primitive morphological traits[1]). Fig. 2 lists the species of mammals identified by the National Center for Biotechnology Information as having genome sequencing projects that are complete, under assembly, or in progress. Although many species were selected based on agricultural, biomedical, or evolutionary criteria, the genome sequences of each species will have a profound impact on conservation and captive breeding research, by allowing the development of markers for use in conservation genetics, and by permitting immediate access to sequences of candidate genes for hereditary traits or diseases that may affect the fitness of free-ranging or captive species.

GENOME-ENABLED SPECIES

Genomic sequencing of a species does not only facilitate genetic research within the sequenced species. It also facilitates genetic studies of closely related species within the same genus, family, or even order of mammals.[3] Many species of mammals being sequenced are either themselves endangered or part of the same genus or family as other species that are endangered.[3] Even mammals chosen because they are common species of livestock, companion

Encyclopedia of Biotechnology in Agriculture and Food DOI: 10.1081/E-EBAF-120045214

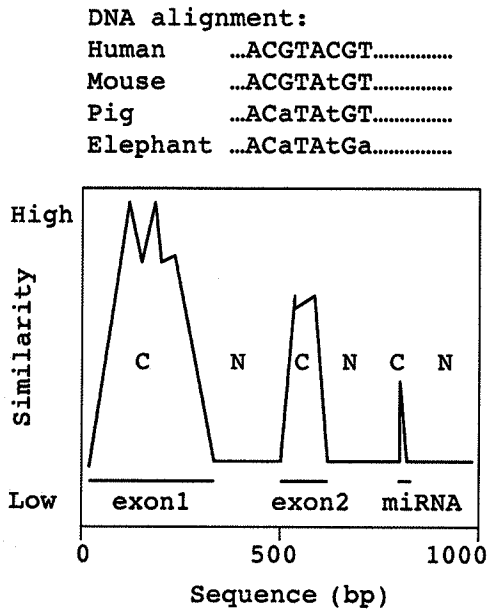

Fig. 1 Phylogenetic or genomic shadowing provided a major justification for sequencing non-human mammalian genomes.[2] Alignment of genomic sequences from a diverse set of organisms (top) will reveal regions (bottom) where the DNA sequence is conserved (C) or non-conserved (N). Functional regions of the genome tend to show greater sequence conservation across species. Thus non-human genomes can be used to annotate the human genome by identifying conserved regions of known or unknown functional importance for human health and disease studies. In the illustrative example depicted, conserved regions identify functional exons (coding regions of proteins) and a microRNA (which regulates gene expression).

animals, or model organisms in biomedical research (Fig. 2) can "genome-enable" a broad range of related wild or endangered species.[3] For example, the genome of the domestic dog was recently sequenced[4] since the species is a very common companion animal, and there are an estimated four hundred million dogs worldwide. Yet in addition to its impact on dog veterinary research, sequencing of the domestic dog genome will also enable the development of genetic markers and facilitate genetic studies among other members of the dog family Canidae, which includes some 35 species, several of which are endangered or threatened (Fig. 3). Thus, genetic studies on endangered African wild dogs or on the threatened maned wolf of South America (Fig. 3) will be greatly facilitated by the domestic dog genome project.[4]

Like the canids, other domesticated species will "genome-enable" their wild relatives. At the family level of classification, the genome sequence of the pig will enable studies on wild suids such as pygmy hogs and babirusas;[5] the domestic cat genome will enable studies on cheetahs, lions, tigers, and other wild felids.[6] Sequencing of the horse genome will enable genetic studies of wild horses, wild asses, and zebras; in one study of 20 short tandem repeat (STR or microsatellite) markers developed in the domestic horse, 15 proved useful in two species of zebras.[7] In all, the current mammalian genome sequencing projects will "genome-enable" almost a thousand endangered species of mammals.[3]

INTROGRESSIVE HYBRIDIZATION

Introgressive hybridization, in which the germ line of a wild species is threatened due to interbreeding with a domestic species, is a threat to the conservation of many wild species, notably those that are closely related to common domestic animals. For example, introgressive hybridization by domestic cats is a threat to African and European wild cats; hybridization with domestic dogs is a major threat to the highly endangered Simian wolf of Ethiopia; while hybridization with domestic horses threatens the Przewalski's or Mongolian wild horse. Genetic methods can assess the extent of hybridization present among wild populations, and genomic sequencing of domestic animal genomes has greatly facilitated such studies. For example, the genome sequence of domestic cattle was used to generate closely linked short tandem repeat (STR) markers for 14 chromosomal segments, examined in 14 populations of American bison (Fig. 4).[8] Although cattle and bison will preferentially mate with members of their own species, male bison were deliberately crossed with female domestic cattle in many of the private ranches that saved the bison from extinction in the late 1800s.[8] These small herds were the sources later used to stock protected populations across North America, and recent genetic tests have detected the introgression of cattle alleles in many bison populations. However, the plains bison population in Yellowstone National Park in the United States and the wood bison population in Wood Buffalo National Park in Canada are known from historical records to have been continuously free-ranging.[8] Thus bison populations from these two parks were used to identify STR loci with species-specific alleles that distinguish bison from cattle.[8] Using this information, along with previous studies of mtDNA and nuclear markers, six additional bison populations were identified for which no evidence was detected of introgression of either mitochondrial or nuclear markers, although several other bison populations carried high frequencies of domestic cattle alleles.[8] This effective use of genomic markers to detect interspecies hybridization has important consequences for conservation. For example, translocations of bison to establish new populations can rely on source herds that are free of introgression by domestic cattle, while hybrid source populations can be avoided in captive breeding of plains or wood bison.

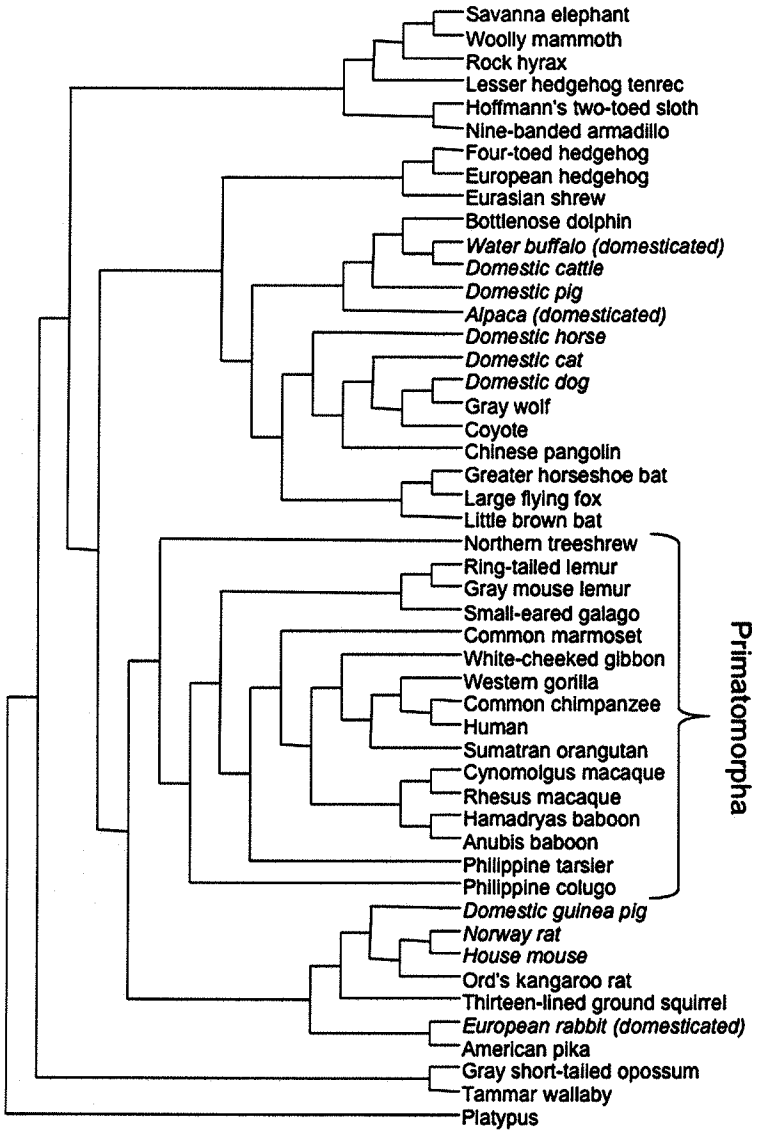

Fig. 2 Mammalian species for which genome sequencing projects are complete, under assembly, or in progress. Many primates and their close relatives (Primatomorpha) were chosen for sequencing due to their close evolutionary relationship to humans, while many other species are domestic or laboratory animals (italicized). Other criteria for selecting species included primitive or derived morphology, unique genomic attributes, or great evolutionarily distance from human.[1] The phylogenetic relationships among the species are shown. Branch lengths are not proportional to genetic distances or chronological age.

ADAPTIVE AND DETRIMENTAL VARIATION

Many endangered species have gone through population bottlenecks that led to reduced genetic variation within the species, while zoo populations often relied on a small number of founder individuals, with subsequent inbreeding of individuals in their collections. Zoos have instituted captive breeding programs that attempt to adjust the contributions of individual founder animals to the captive gene pool in order to maximize genetic diversity. Most conservation genetic studies of endangered species have largely relied on genetic markers presumed to be neutral, i.e., that do not affect the fitness of the species. For example, both mitochondrial DNA and short tandem repeats (STRs) are commonly used in analyses of population structure because these markers are assumed to be free of selective pressure. While genomic data enables more comprehensive surveys of neutral variation, genomic methods should also facilitate identification of alleles that contribute either to positive adaptive variation within a species or to negative deleterious effects on organismal health,[3] and may reveal how different taxa have adapted to different climatic or other conditions. Adaptive alleles could be targeted by selective breeding efforts, to increase the fitness of zoo or of managed

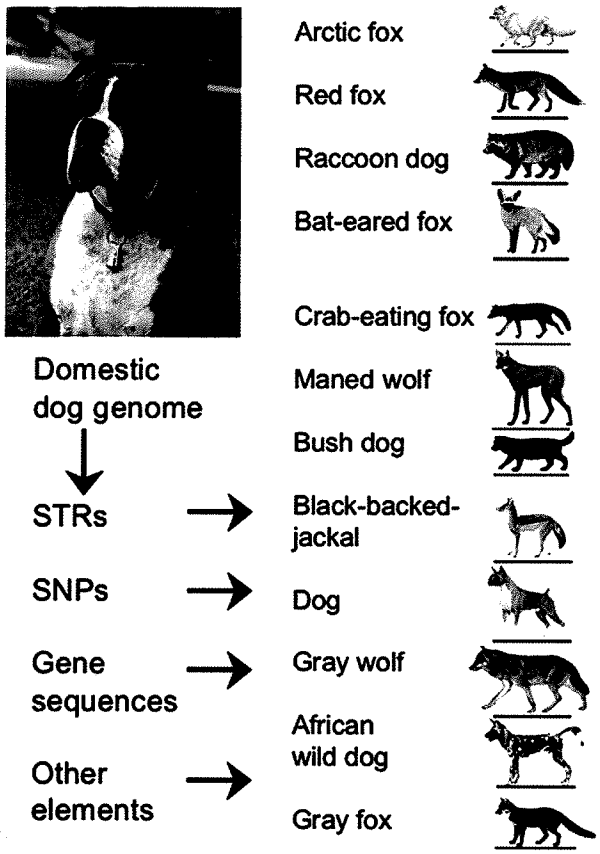

Fig. 3 Genomic sequences from one species will allow many related species to become "genome-enabled." For example, molecular markers derived from genomic sequencing of the domestic dog (left)[4] can be tested and used for research studies on wild species of the dog family Canidae (right). These include endangered or threatened species such as the African wild dog and the maned wolf. Thus although the domestic dog was selected for genomic sequencing based on the utility of its DNA sequences for veterinary and biomedical research, the sequence of the dog genome will also facilitate studies on related wild species, including genetic studies involving short tandem repeats (STRs), single nucleotide polymorphism (SNPs), gene sequences or other genomic elements. **Source:** Images courtesy of the NHGRI and The Broad Institute of MIT and Harvard; and of Elsevier.

wild populations. A number of diseases in zoo animals have been proposed for investigation using genomic approaches, including iron overload disorders in lemurs, cardiovascular diseases in bonobos, and susceptibility to pathogens in a broad range of species.[9] Susceptibility to many diseases will involve the major histocompatibility complex (MHC), a gene-dense region of the genome that plays a major role in immune function. The generation of genomic sequences has already enabled the comparative study of MHC structure across mammals,[6] and genomic data may reveal immunogenetic networks involved in host-pathogen interactions in this era of zoonotic diseases.

In wild populations, genomics may enable more extensive studies of the effects of fishing or hunting on heritable traits,[10] including quantitative traits. For example, the sequence of the elephant genome[11] could be used to examine how hunting for the ivory trade has impacted loci involved in determining tusk growth; or used to develop better genetic markers for determining, using DNA from confiscated tusks, the source population of illegally poached ivory. A wild population that may be especially amenable to genomic identification of loci that determine fitness is the Florida panther (Fig. 4), a subspecies of the puma that suffered the deleterious effects of substantial inbreeding following its isolation and decline in South Florida.[12] An increasing number of wild panthers were found to suffer from heart defects and cryptorchidism that resulted from substantial inbreeding, as the population numbered fewer than 50 individuals for many generations. In an attempt to reverse the increase in hereditary diseases due to inbreeding, several Texas pumas were introduced into Florida to increase the genetic diversity of the panthers. With the recent sequencing of a domestic cat genome (Fig. 4),[6] the Florida panther has become a "genome-enabled" taxon.[3] DNA has been sampled from generations of panthers, and pedigrees are well established, so the Florida panther could become one of the first wild species in which alleles affecting fitness can be identified. Likewise, since newly developed sequencing platforms permit genomic sequencing of ancient DNA from museum samples, current allele frequencies could be compared to historical patterns within the subspecies.[12] Thus, new genomic technologies will permit fine-tuning of the genetic restoration efforts for subspecies, guided by genetic patterns found among outbred specimens collected by museums before any recent genetic bottlenecks.

CONCLUSION

One of the most striking technological advances of recent years is the decline in the cost of DNA sequencing (Fig. 5), analogous in speed to the decline in the price of computing power or the rise of the internet.[10] This trend has been accelerated by the development of next generation sequencing platforms (Fig. 5).[10] The decline in sequencing costs will make genomic studies more feasible in wildlife conservation and captive breeding research. The limiting factor will no longer be sequencing costs but the cost of personnel and ability to collect samples.[10] The price decline has led to discussions on possibly sequencing the entire genomes of 10,000 different vertebrate species. Declining sequencing costs may also expand the scope of zoo research into new areas, such as measuring gene expression in individual animals to determine the effects of stress or other factors. It will also be possible to compare large numbers of individuals in captive collections to those in natural or historic populations, to determine whether, to what degree, and at which

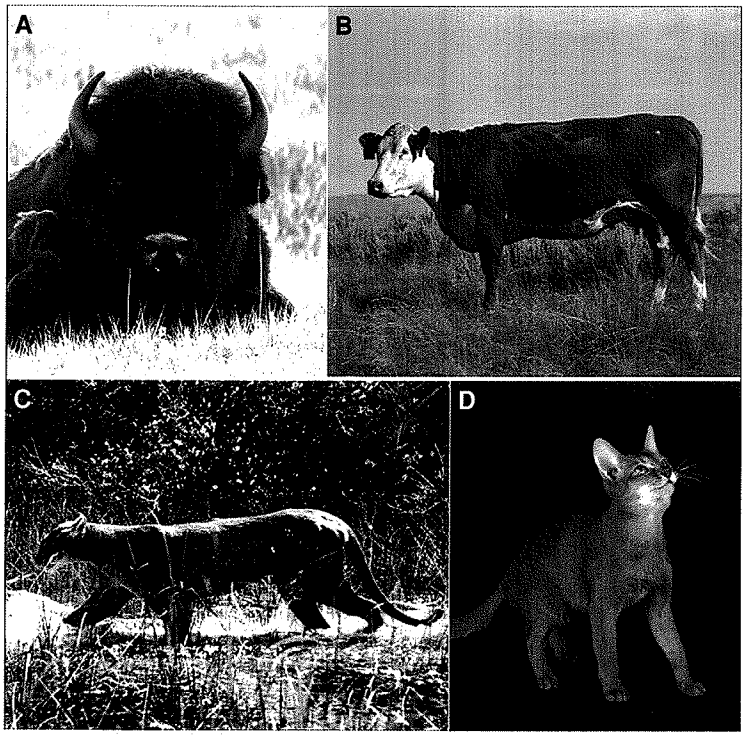

Fig. 4 Genome sequencing projects enable the study of introgressive hybridization and genetic fitness in wild species. An intensive study[8] of the degree to which surviving populations of the American bison (A) have experienced introgression of cattle alleles due to hybridization was enabled after the genome of a domestic cow (B) was sequenced. Bison populations that have remained genetically intact will be especially important for conserving the species. (C) The Florida panther is a wild puma subspecies that has suffered the deleterious hereditary effects of substantial inbreeding.[12] It has been intensively studied with well-characterized pedigrees. Sequencing of the domestic cat (D) genome[6] may permit identification of deleterious alleles in the genome of the panther. **Source:** Photos courtesy of Melissa Dowland; Michael MacNeil, USDA; Robert C. Garrison; and Dr. Kristina Narfstrom, University of Missouri–Columbia.

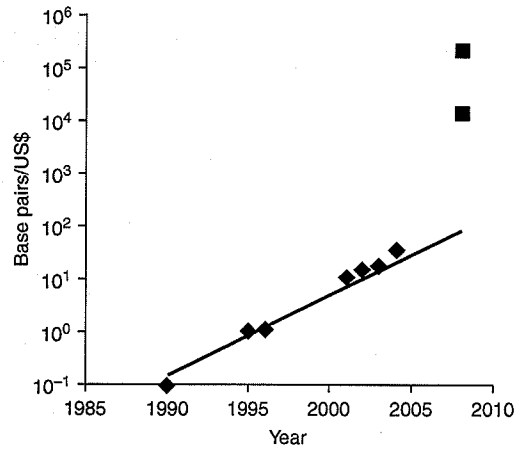

Fig. 5 The number of base pairs that can be sequenced for $1 has consistently doubled every 2 years (diamonds). Moreover, the recent development of next generation sequencing platforms has enabled an even faster increase in efficiency (squares). The plummeting cost suggests that genomic sequencing may soon become possible for thousands of wild and endangered species. **Source:** From Fish Fish.[10]

loci captive breeding may have inadvertently changed allele frequencies. Thus genomic technologies may revolutionize the conservation and captive breeding management of species.

REFERENCES

1. O'Brien, S.J.; Eizirik, E.; Murphy, W.J. Genomics. On choosing mammalian genomes for sequencing. Science **2001**, *292* (5525), 2264–2266.
2. Boffelli, D.; McAuliffe, J.; Ovcharenko, D.; Lewis, K.D.; Ovcharenko, I.; Pachter, L.; Rubin, E.M. Phylogenetic shadowing of primate sequences to find functional regions of the human genome. Science **2003**, *299* (5611), 1391–1394.
3. Kohn, M.H.; Murphy, W.J.; Ostrander, E.A.; Wayne, R.K. Genomics and conservation genetics. Trends Ecol. Evol. **2006**, *21* (11), 629–637.
4. Wayne, R.K.; Ostrander, E.A. Lessons learned from the dog genome. Trends Genet. **2007**, *23* (11), 557–567.
5. Schook, L.B.; Beever, J.E.; Rogers, J.; Humphray, S.; Archibald, A.; Chardon, P.; Milan, D.; Rohrer, G.;

Eversole, K. Swine genome sequencing consortium (SGSC): a strategic roadmap for sequencing the pig genome. Comp. Funct. Genomics **2005**, *6* (4), 251–255.

6. Pontius, J.U.; Mullikin, J.C.; Smith, D.R.; Agencourt Sequencing Team; Lindblad-Toh, K.; Gnerre, S.; Clamp, M.; Chang, J.; Stephens, R.; Neelam, B.; Volfovsky, N.; Schäffer, A.A.; Agarwala, R.; Narfström, K.; Murphy, W.J.; Giger, U.; Roca, A.R.; Antunes, A.; Menotti-Raymond, M.; Yuhki, N.; Pecon-Slattery, J.; Johnson, W.E.; Bourque, G.; Tesler, G.; NISC Comparative Sequencing Program, O'Brien, S.J. Initial sequence and comparative analysis of the cat genome. Genome Res. **2007**, *17* (11), 1675–1689.
7. Moodley, Y.; Baumgarten, I.; Harley, E.H. Horse microsatellites and their amenability to comparative equid genetics. Anim. Genet. **2006**, *37* (3), 258–261.
8. Halbert, N.D.; Ward, T.J.; Schnabel, R.D.; Taylor, J.F.; Derr, J.N. Conservation genomics: disequilibrium mapping of domestic cattle chromosomal segments in North American bison populations. Mol. Ecol. **2005**, *14* (8), 2343–2362.
9. Ryder, O.A. Conservation genomics: applying whole genome studies to species conservation efforts. Cytogenet. Genome Res. **2005**, *108* (1–3), 6–15.
10. Hauser, L.; Seeb, J.E. Advances in molecular technology and their impact on fisheries genetics. Fish Fish. **2008**, *9* (4), 473–486.
11. Roca, A.L.; O'Brien, S.J. Genomic inferences from Afrotheria and the evolution of elephants. Curr. Opin. Genet. Dev. **2005**, *15* (6), 652–659.
12. Culver, M.; Johnson, W.E.; Pecon-Slattery, J.; O'Brien, S.J. Genomic ancestry of the American puma (*Puma concolor*). J. Hered. **2000**, *91* (3), 186–197.