

Discoveries of new mammal species and their implications for conservation and ecosystem services

Gerardo Ceballos^{a,1} and Paul R. Ehrlich^{b,1}

^aDepartamento de Ecología de la Biodiversidad, Instituto de Ecología, Universidad Nacional Autónoma de México, AP 70-275, México, D.F. 04510, México; and ^bDepartment of Biological Sciences, Center for Conservation Biology, Stanford University, Stanford, CA, 94305

Contributed by Paul R. Ehrlich, December 21, 2008 (sent for review November 3, 2008)

In light of recent discoveries of many new species of poorly-studied organisms, we examine the biodiversity of mammals, a well known “charismatic” group. Many assume that nearly all mammal species are known to scientists. We demonstrate that this assumption is incorrect. Since 1993, 408 new mammalian species have been described, ≈10% of the previously known fauna. Some 60% of these are “cryptic” species, but 40% are large and distinctive. A substantial number persist only in areas undergoing rapid habitat destruction. Our findings suggest global animal and plant species diversity is badly underestimated even in well studied taxa. This implies even greater threats to ecosystem services and human well-being than previously assumed, and an increased need to explore, understand, and conserve Earth’s living resources.

biodiversity | extinction | new mammals

Today biology is in “a new age of discovery” (1). That age is characterized by the uncovering of vast new elements of biodiversity, which are the fundamental building blocks of ecosystems, and thus the provision of ecosystem goods and services. There are thousands of examples of unexpected discoveries of new taxa across broad taxonomic and geographic spectra, from extremophile bacteria in Yellowstone geysers to whole new ecosystems in the Pacific Ocean hydrothermal vents (2, 3). For example, the Census of Marine Life program has uncovered hundreds of new species (4). Similarly, recent work has shown that a “species” of skipper butterfly, *Astraptes fulgerator* was actually a complex of 10 species with distinct life histories, and that 16 species of “generalist” tropical parasitoid tachinid flies were actually 73 evolutionary lineages (as indicated by mitochondrial DNA barcoding) including many lineages specialized to attack different hosts (5, 6).

These findings are of much more than academic interest. Most of the focus in conservation has been on trying to preserve as much of species diversity as possible (7, 8). Although the equally critical need for population preservation is now recognized (9, 10), the diversity of species remains crucial as a source of populations that can assume more distinct ecological roles (e.g., as generalist or specialist predators) in a rapidly changing world. Previously unrecognized genetic diversity must therefore be evaluated so that biologists have some idea of what they must strive to preserve, and how to deploy their limited resources to reduce biodiversity loss.

Here, we evaluate discoveries of new species of mammals, an especially well-studied group. We first give the methods by which new mammalian diversity has been discovered. Then we review the taxonomic affiliations, range size, and patterns of geographic distribution of mammal species described since a comprehensive 1993 checklist (11). Finally, we discuss the significance of these findings for the status of biodiversity in general, the problems of maintaining it, and thus of the ecosystem services that depend upon that diversity.

What are the ways in which additional mammal diversity has been uncovered? We started with a thorough search for new species of mammals and created maps for all new species except for marine ones, from the literature (*SI Appendix*). Global

patterns of species distribution were done using 10,000-km² (2) grid cells, similar to our previous studies (10, 12, 13). The new mammal species we found were of three types. The first was morphologically distinct species found in previously poorly surveyed areas. The second, the result of using molecular genetic techniques, was discoveries that the geographic range of a well-known organism was actually the combined ranges of two or more cryptic species—one’s not easily recognized by morphological features. The third type consists of species that had been considered subspecies and were newly elevated to specific status (again, often as the result of molecular genetic discoveries). Two of the most prominent recent cases involved giving specific status to populations of forest elephants in central Africa and orangutans in Borneo (14).

In this article we will deal only with the first two cases—if the third were considered we would be dealing with >1000 “new” species. We did not map new species of marine mammals, which include whales and dolphins. Even 250 years after taxonomists started formally naming new mammals, 408 new species (excluding those elevated subspecies), have been documented in the last 15 years, a surprisingly large number considering <4,800 mammal species had been described at the beginning of that period. The discoveries include 18 new genera such as a large bovid (*Pseudoryx*), a rodent (*Cuscomys*), a bat (*Xeronycteris*), and a primate (*Rungwecebus*), and a living representative of Diatomyidae, a family considered extinct for 11 million years (Fig. 1 and *SI Appendix*). The new species belong to 18 mammalian orders (Table 1). The newly-discovered species varied in size from a 3-g shrew-tenrec (*Microgale jobihely*) to the 100-kg soala antelope (*Pseudoryx nghetinhensis*), and include some remarkable creatures such as a pygmy sloth (*Bradypus pygmaeus*) from a Panamanian island, a “giant” muntjac (*Megamuntiacus vuquangensis*) from Vietnam, a white titi monkey (*Callithrix mauesi*) from a river near Manaus in Brazil, and the Solomons islands monkey-faced bat (*Pteralopex taki*). The number of new species among orders was not random, i.e., related to the order’s total species richness. It was higher than expected for Primates, Chiroptera, Rodentia, and all orders that used to belong to marsupials; in contrast, it was less than expected in Soricomorpha, Artiodactyla, and Carnivora (χ^2 goodness of fit between expected and observed species richness order; $X(2) = 40.32$, $df = 12$, $P < 0.001$).

The discovery of some of these species has generated considerable interest within the scientific community. For example, both the recently described rodent species from the family Diatomyidae and genus *Cuscomys* were already known from paleontological and prehistoric remains, respectively. This is an instance of the “Lazarus effect” (15)—in which an organism

Author contributions: G.C. and P.R.E. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

¹To whom correspondence may be addressed. E-mail: gceballo@ecologia.unam.mx or pre@stanford.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0812419106/DCSupplemental.



Fig. 1. Examples of new species of mammals discovered since 1993. From top left to bottom right, *Rungwecebus kipunji* (Copyright 2006, Tim Davenport/World Conservation Society). *Cuscomys ashanika* [Reproduced with permission from Emmons (*SI Appendix*) (Copyright 1999, American Museum of Natural History)]. *Bradypus pygmaeus* (Copyright 2007, Bill Haycher/National Geographic Society). *Mirza zaza* (Copyright 2006, David Haring/Duke Lemur Center). *Cebus queirozi* [Reproduced with permission from Pontes et al. (*SI Appendix*) (Copyright 2006, Magnolia Press)]. *Rhyncocyon udzunwensis* [Reproduced with permission from Rovero et al. (ref. 17) (Copyright 2007, The Zoological Society of London)]. *Macrotarsomys petteri* [Reproduced with permission from Goodman and Saorimalala (*SI Appendix*) (Copyright 2005, Biological Society of Washington)]. *Laonastes aenigmamus* (Copyright 2007, David Redfield/Florida State University). *Scotophilus marovaza* [Reproduced with permission from Goldman et al. (*SI Appendix*) (Copyright 2006, Polish Academy of Sciences)]. *Microgale jenkinsae* [Reproduced with permission from Goldman et al. (ref. 18) (Copyright 2006, The Zoological Society of London)].

Table 1. Taxonomic composition of the new species of mammals (excluding marine species) discovered since 1993

Order	Families with new species	Genera with new species	New species	New species with restricted distribution	New species probably at risk of extinction
Afrosoricida	2	2	12	8	2
Artiodactyla	5	9	11**	7	1
Carnivora	1	2	2*	2	2
Macroscelidae	1	1	1	1	1
Chiroptera	8	44	94*	75	6
Cingulata	1	1	1	1	0
Dasyuomorpha	1	4	6*	2	0
Didelphimorphia	2	5	8*	8	0
Diprodontia	2	6	11*	11	2
Erinaceomorpha	1	1	1	1	0
Lagomorpha	2	3	5	3	0
Monotremata	1	1	1	1	0
Paucituberculata	1	1	1*	1	1
Peramelemorphia	1	1	2*	2	0
Pilosa	1	1	1	1	0
Primates	9	16	55*	51	10
Rodentia	16	87	174*	29	4
Soricomorpha	2	9	22**	17	2
TOTAL	57	195	408	221	34

There are new taxa up to the family levels. Some orders have either more (*) or fewer (**) new species than expected by their species richness.

known only from fossils is discovered alive. Remarkably, the diatomid species (*Laonastes aenigmamus*) and a new rabbit species (*Nesolagus timminsi*) were first discovered being sold as food in a market in a Laotian village (15, 16). It appears that exploration of new regions has been the main factor for the discovery of as much as 40% of the new species, such as the pygmy deer (*Muntiacus putaoensis*) in Bhutan, the Arunachal macaque (*Macaca muzala*) from the Himalaya foothills of north-east India, the Amazonian basin monkeys, and most of the new Philippines species (*SI Appendix*). The exploration of new regions has been based on both the use of either new techniques such as camera-traps, which were the first indication that there was a new giant elephant shrew (*Rhynchocyon udewensis*) in Tanzania (17), or traditional techniques, such as pitfall traps, which have yielded specimens of 8 new species of shrew-tenrecs from Madagascar since 1988 (18). Molecular techniques have revealed cryptic species across many orders. For bats and galago monkeys, discriminating among echolocation signals and vocalizations respectively have been key to identifying cryptic species (*SI Appendix*).

The patterns of distribution of new species are shown in Fig. 2, based on a global grid of some 17,000 10,000-km² (2) terrestrial cells. The number of new species in a single cell varied from 1 to 10. New species have been discovered on all continents except Antarctica, with the majority in South America and Asia (*SI Appendix*). In the Americas, cells with one or two new species occur in temperate regions of Alaska, the eastern U.S., Chile, and Argentina, whereas cells with two species or more have been found throughout tropical and semitropical regions in Mexico and Central America, eastern Colombia, Peru and Ecuador, the Amazon basin, and the Atlantic forests of Brazil. Most new species on this continent are bats and primates.

In Africa, most new species have been discovered in tropical regions, but some species have been found in arid regions in Western Sahara and Namibia; discoveries have been concentrated in eastern tropical forests of west Africa and the Congo Basin, from Liberia to Angola, the eastern mountains of Somalia, Kenya, and Tanzania, and Madagascar, where up to 3 new species have been discovered in some cells. Surprisingly, several

new species have been discovered in Europe, mostly around the Mediterranean basin. New species in Asia are concentrated in the Malayan Peninsula, Indonesia, and New Guinea. The number of new species discovered in Philippines is rather remarkable.

On average these species had ranges of ≈87,000 km² (2), significantly smaller compared with an average land mammal range of 400,000 km² (2) ($P < 0.0001$). Indeed, 81% of the new species have very restricted ranges [i.e., <10,000 km² (2)] (Fig. 2), which make them more prone to extinction. Interestingly, the distribution of newly discovered mammals often includes large areas not considered biodiversity hotspots (Fig. 3), which further indicates that conservation strategies to supplement the focus on hotspots are required (13, 19). Also interesting, and unexpected, is that the new mammal species were larger than average ($P < 0.0001$). This is primarily because few of the newly discovered species were either bats or rodents.

Although most (61%, 1640) of the cells where new species have been found have relatively little anthropogenic threat, measured as both the area of the cell under agriculture and human population, 24% of the cells are located in cells with >10% of their land area under agriculture, including 12% of cells with >50% of agriculture (Fig. 3A). In contrast, most (46%) cells are in regions with low human population density [< 10 individuals per square kilometer (2)]; however, >20% are found in regions with relatively high human populations (Fig. 3B), indicating higher vulnerability. A very interesting example is the mammalian fauna discovered in a limestone karst outcrop in the the Kammaouan province, in the Lao People's Democratic Republic, which included a new family and 6 species, in a region completely isolated by agriculture (15).

The discoveries of new mammals are hardly unique (20, 21). Our analysis supports the anecdotal conclusions from butterflies, flies, and other organisms mentioned above. It suggests that other prominent taxa (e.g., birds and reptiles), and more obscure groups, likely contain many more species than are currently described. This could amount to millions of species and other distinct entities, greatly expanding estimates of the diversity of the living elements of Earth's natural capital (22), to even perhaps hundreds of millions of species. In addition, because

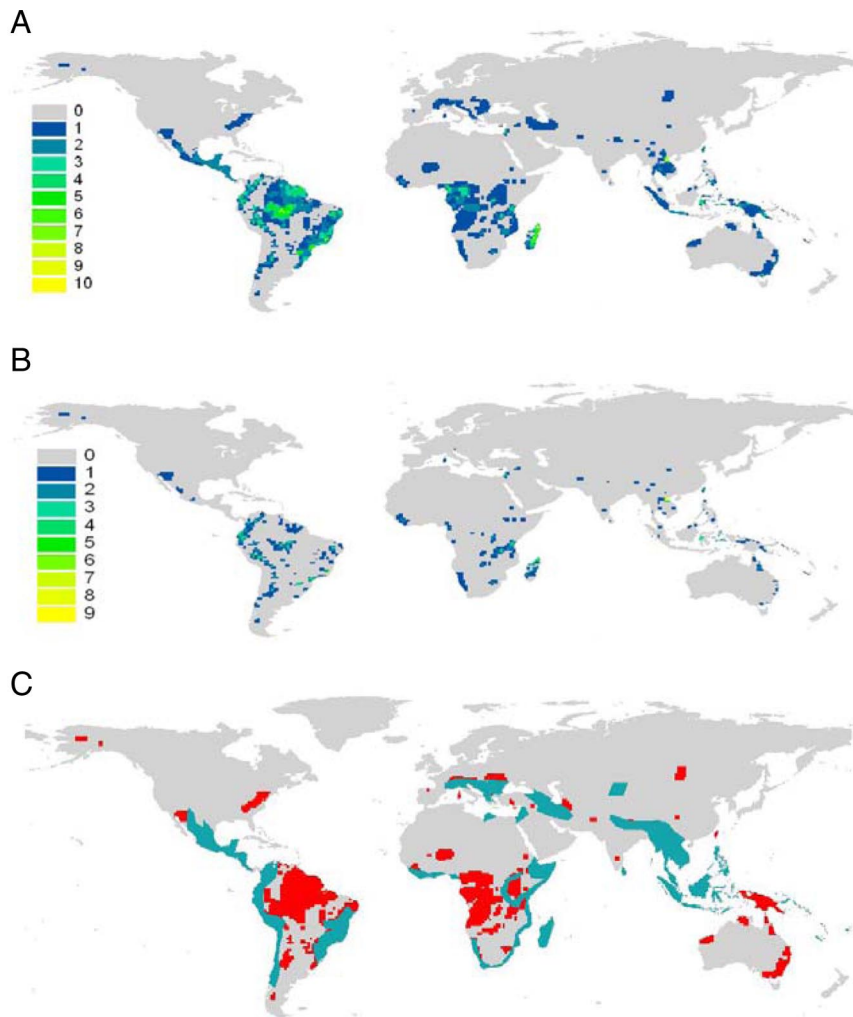


Fig. 2. Patterns of distribution in new species of mammals. (A) Species richness, $n = 408$. (B) Restricted-range species, $n = 221$. (C) Cells (in red) with new species located outside hotspots [in blue, *sensu* Myers (13)].

12% of Earth's land surface is used for crop agriculture, 25% is grazed by livestock, 2% has been paved or built on, 30% is exploited in other ways (23), our results suggest that many more unheralded organisms in all groups have likely recently gone extinct without being noticed. That implies that the levels of species extinction overall have been grossly underestimated. Thus, the situation is likely even worse than indicated by the steady rise of endangerment in the IUCN mammal statistics (8). Although it is common for estimates of total current plant and animal biodiversity to be in the tens of millions (24), those estimates are largely based on rates of discovery of morphologically defined species found in traditional surveys.

The problem of cryptic biodiversity, and the incompleteness of inventories of even charismatic organisms, is not usually considered. This is especially likely because the species now being discovered, as illustrated by mammals, tend to have limited distributions. For instance, the golden capuchin monkey (*Cebus queirozi*) was described in 2006, and is known to occur in a 200 ha remnant forest patch, isolated by sugar cane plantations (25). Similarly, the Solomon Islands flying fox (*Pteralopex taki*) was described in 2002 from 3 islands, and was already extinct on one of them (26). The lemur genus *Microcebus*, thought to consist of two species in 1982, has now been shown to comprise ≈ 13 cryptic species (27). It, of course, may have once contained many other cryptic species, all of which went extinct unheralded. This seems

likely, considering the massive deforestation that has occurred on Madagascar and the inconspicuous character of many lemurs.

Population loss is also largely unrecorded, except when a well-defined subspecies goes extinct, as in the case of the satyrine butterfly *Cercyonis sthenele sthenele* that famously disappeared in the 1880s from San Francisco sand dune habitats (28) or the more recent loss of the Caspian, Balinese, and Javan tiger subspecies (*Panthera tigris virgata*, *P. t. balica*, *P. t. sondaica*) and the well-publicized near extinctions of the Asian cheetah (*Acinonyx jubatus venaticus*) and Florida panther (*Puma concolor coryi*). In short, there has probably been substantial cryptic loss of population biodiversity over much of the planet even in well-studied groups (10).

Several commentators have suggested that the discovery of "new species" is problematic for conservation—especially "taxonomic inflation" (raising of subspecies to specific status and uncovering of cryptic species) (29). We and others disagree (30). There is little need to focus on taxonomic rank when what needs to be preserved are the numbers and diversity of biological entities. For example, it is important to know that most tachinid flies in Costa Rica are host specialists. Whether they are counted as "good species" or "mitochondrial lineages" makes no scientific difference. Conserving one of those tachinid lineages, for instance, may preserve a crucial biological control agent. The key thing is that in an ideal world we should conserve all such units,

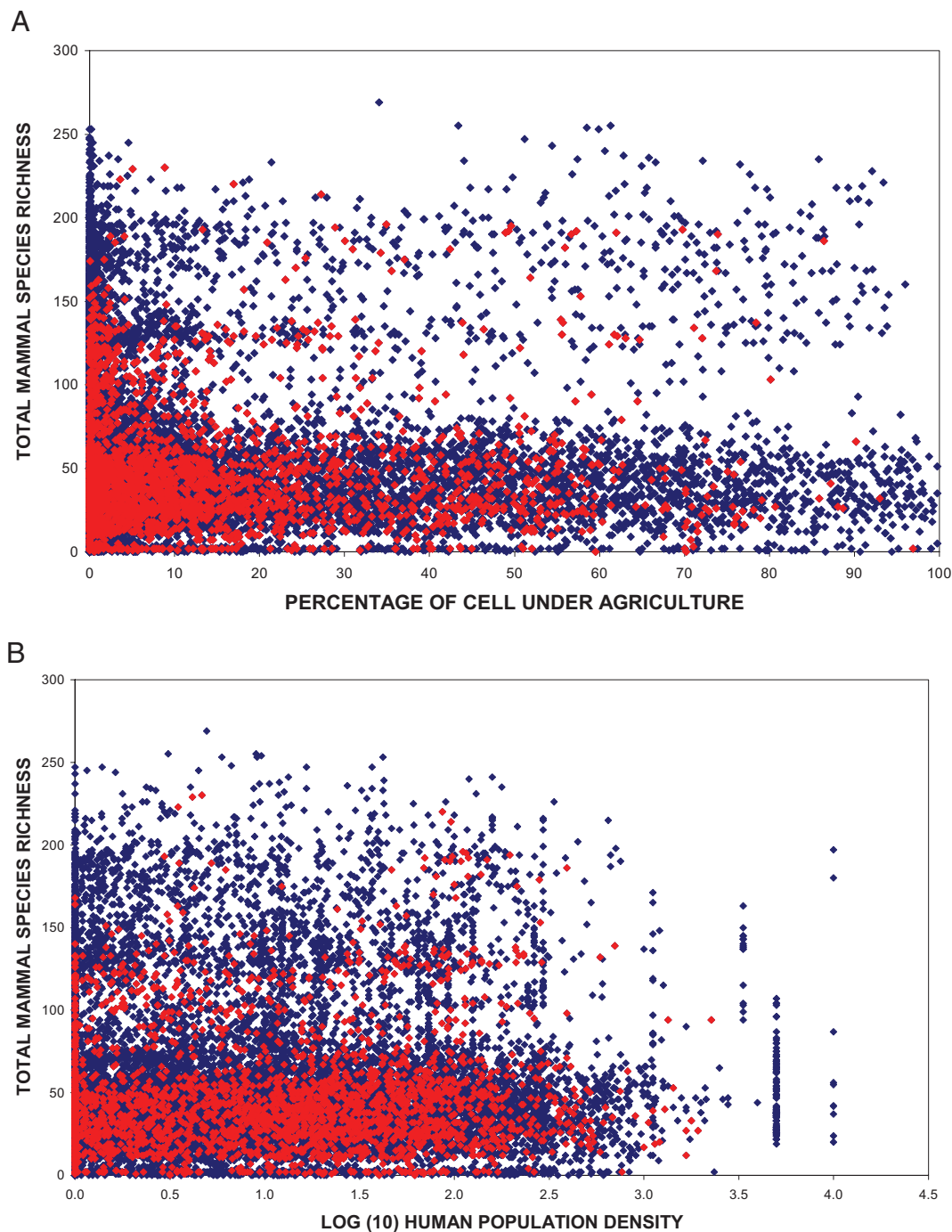


Fig. 3. Anthropogenic threat in cells either with (red) or without (blue) new species of mammals measured by the percentage of the cell under agriculture (A) and its human population density (B).

regardless of appellation, keeping the loss rate not significantly above the “background” rate.

Many newly discovered entities may supply previously unrecognized ecosystem services. For example, a recent study has shown that the abundance of a hantavirus-prone rodent species and hantavirus infection rates are negatively correlated with the number of native rodent species in Panamanian tropical forests (31). Loss of such native taxa can thus potentially have negative effects on human health and welfare. Furthermore, the role of large mammals in regulating the trophic and architectural properties of ecosystems has become even clearer with the recent

investigations of the impacts of large herbivores (32). Such results underscore the often-neglected point that conserving biodiversity over broad areas is essential to maintaining ecological function and critical ecosystem services (7, 9, 10).

However, no one is in a position to decide the full conservation value of any species, charismatic or not, let alone the other more or less distinct entities now being revealed. This moves the “rivet popper hypothesis” to a new level (33). Scientists know that there is some functional redundancy in the species composition of most ecosystems (34). However, the level of that redundancy may be generally overrated, as research on the buffering of ecosystem processes by diversity demonstrates (35).

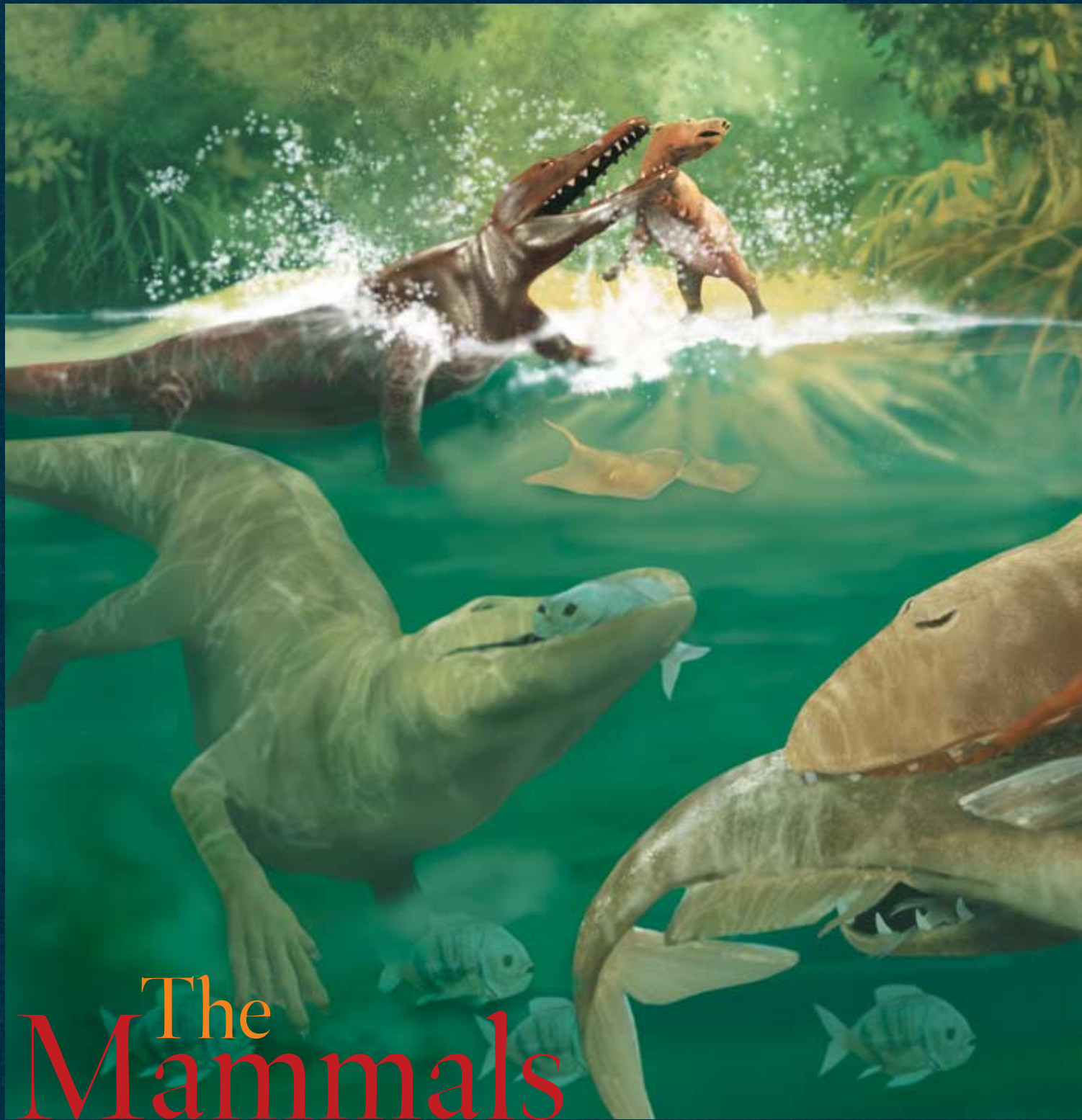
In response to these problems, what should be the strategy of conservation biologists? It goes without saying that they should try to preserve as many genetically distinct species as possible. It is also crucial that the number and diversity of populations—many of which are clearly more genetically and ecologically differentiated than previously thought—and the ecosystem services they provide, also be preserved and, where possible, restored. The whole issue of triage needs to be revisited—triage decisions may be required, but they will involve vast scientific, socioeconomic, and political uncertainties. Also vexed will be issues of “where to draw the line” (because most individuals are genetically distinct and we can not preserve everything) (36). The more diversity that is discovered the more urgent becomes putting additional resources into understanding and finding ways to conserving it. The insufficiency to date of ethical and esthetic arguments for preserving biodiversity means that ecosystem service based approaches, typified by countryside biogeography and the Natural Capital Project, must be expanded (37). This is especially the case in the face of increasing threats to virtually all organisms, which are experiencing rapid climate, land conver-

sion, and extensive toxic pollution—threats that now extend to areas previously considered protected, of marginal value, or remote.

Finding the political will to attain such goals will not be easy, but the survival of civilization may well hang on a cultural evolutionary sea change, and how much of societies resources get allocated to the task. Considering the complexity and uncertainty of the relationships between biodiversity and the delivery of ecosystem services, conservation decisions should include a very large precautionary principle bias toward protection of as many of our living companions as possible.

ACKNOWLEDGMENTS. We thank our friends Peter Raven, Rob Pringle, and Watt Ward for helpful comments on a previous draft of the manuscript; Ana Davidson and Navjot Sodhi for reviewing and helping us to improve the manuscript; Irma Salazar, Silvia Surhig, Pablo Ortega, and Regina Ceballos for kindly helping us with data and GIS analyses; and Tim Davenport, Louise Emmons, Bill Haycher, David Haringm, David Redfield, Francisco Rovero, Achille Raselimanana, and Steve M. Goodman for lending us their photographs. This work was supported by grants from Dirección General de Asuntos del Personal Académico of the Universidad Nacional Autónoma de México, Comisión nacional para el conocimiento y uso de la biodiversidad (Mexico), and the Mertz Gilmore Foundation.

- Donoghue MJ, Alverson WS (2000) A new age of discovery. *Ann Missouri Botanical Gardens* 87:110–126.
- Coriss JB (1979) Submarine thermal springs on the Galápagos Rift. *Science* 203:1073.
- Yooseph S, et al. (2007). The Sorcerer II: Global ocean sampling expedition: Expanding the universe of protein families. *PLoS Biol* 5:e16.
- Census of Marine Life Portal (2008) www.coml.org/medres/highlights2006/AnnHigh06.low.res.revised.pdf. Accessed August 20, 2008.
- Hebert PDN, et al. (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc Natl Acad Sci USA* 101:14812–14817.
- Smith MA, Wood DM, Janzen DH, Hallwachs W, Hebert PDN (2007) DNA barcodes affirm that 16 species of apparently generalist tropical parasitoid flies (Diptera, Tachinidae) are not all generalists. *Proc Natl Acad Sci USA* 104:4967–4972.
- Ricketts TH, et al. (2005) Pinpointing and preventing imminent extinctions. *Proc Natl Acad Sci USA* 51:18497–18501.
- International Union for the Conservation of Nature and Natural Resources (2008) *The 2008 IUCN Red List of Threatened Species* (IUCN, Gland, Switzerland).
- Hughes JB, Daily GC, Ehrlich PR (1997) Population diversity: Its extent and extinction. *Science* 278:689–692.
- Ceballos G, Ehrlich PR (2002) Mammal population losses and the extinction crisis. *Science* 296:904–907.
- Wilson DE, Reeder DAM (1993) *Mammal Species of the World: A Taxonomic and Geographic Perspective* (Johns Hopkins Univ Press, Baltimore) 2nd and 3rd Eds.
- Ceballos G, Ehrlich PR, Soberón J, Salazar I, Fay JP (2005) Global mammal conservation: What must we manage? *Science* 309:603–607.
- Ceballos G, Ehrlich PR (2006) Global biodiversity hotspots and conservation: Insights from mammal distributions. *Proc Natl Acad Sci USA* 103:19374–19379.
- Roca AL, Georgiadis N, Pecon-Slattey J, O'Brien SJ (2001) Genetic evidence for two species of elephant in Africa. *Science* 293:1473–1477.
- Dawson MR, Marivaux L, Li C, Beard KC, Metais G (2006) Laonastes and the “Lazarus Effect” in recent mammals. *Science* 311:1456–1458.
- Averianov AO, Abramov AV, Tikhonov AN (2000) A new species of *Nesolagus* (Lagomorpha, Leporidae) from Vietnam with osteological description. *Contributions Zool Inst St. Pettersburg* 3:1–22.
- Rovero F, et al. (2008) A new species of giant sengi or elephant-shrew (genus *Rhynchocyon*) highlights the exceptional biodiversity of the Udzungwa Mountains of Tanzania. *J Zool* 274:126–133.
- Goodman SM, Raxworthy CJ, Maminirima CP, Olson LE (2006) A new species of shrew tenrec (*Microgale jobihely*) from northern Madagascar. *J Zool* 270:384–398.
- Orme CD et al (2005) Global hotspots of species richness are not congruent with endemism or threat. *Nature* 436:1016–1019.
- Reeder DAM, Helgen KM, Wilson DE (2007) Global trends and biases in new mammal species discoveries. *Occas Pap Mus Texas Tech Univ* 269:1–35.
- Kholer J, Vieites DR, Bonett RM (2005) New amphibians and global conservation: A boost in species discoveries in a highly endangered vertebrate group. *Bioscience* 55:693–696.
- Ehrlich PR, Wilson EO (1991) Biodiversity studies: Science and policy. *Science* 253:758–762.
- Ehrlich PR, Ehrlich AH (2008) *The Dominant Animal: Human Evolution and the Environment* (Island, Washington, D.C.).
- Erwin TR (1991) How many species are there? *Conservation Biol* 5:330–333.
- Mendes Pontes AR, Malta A, Asfora PH (2006) A new species of capuchin monkey, genus *Cebus* Erxleben (Cebidae, Primates): Found at the very brink of extinction in the Pernambuco Endemism Centre. *Zootaxa* 1200:1–12.
- Helgen KM (2005) Systematics of the Pacific monkey-faced bats (Chiroptera: Pteropodidae), with a new species of *Pteralopex* and a new Fijian genus. *Systemat Biodiversity* 3:433–453.
- Tattersall I (2000) Paleoanthropology: The last half-century. *Evol Anthropol* 16:12–23.
- Scott JA (1986) *The Butterflies of North America: A Natural History and Field Guide* (Stanford Univ Press, Stanford, CA).
- Isaac NJB, Mallet J, Mace GM (2004) Taxonomic inflation: Its influence on macroecology and conservation. *Trends Ecol Evol* 19:464–469.
- Agapow PM, Sluys R (2005) The reality of taxonomic change. *Trends Ecol Evol* 20:278–280.
- Suzan G, et al. (2008) Epidemiological considerations of rodent community composition in fragmented landscapes in Panama. *J Mammal* 89:684–690.
- Pringle RM, Young TP, Rubenstein DI, McCauley DJ (2007) Herbivore-initiated interaction cascades and their modulation by productivity in an African savanna. *Proc Natl Acad Sci USA* 104:193–197.
- Ehrlich PR, Ehrlich AH (1981) *Extinction: The Causes and Consequences of the Disappearance of Species* (Random House, New York).
- Ehrlich PR, Walker B (1998) Rivets and redundancy. *BioScience* 48:387.
- Hobbs RJ, Yates, Mooney HA (2007) Long-term data reveal complex dynamics in grassland in relation to climate and disturbance. *Ecol Monogr* 77:545–568.
- Kareiva P, Levin S (2002) *The Importance of Species: Perspectives on Expendability and Triage* (Princeton Univ Press, Princeton, NJ).
- Turner RK, Daily GC (2008) The ecosystem services framework and natural capital conservation. *Environ Res Econ* 39:25–35.



The Mammals That Conquered the

New fossils and DNA analyses elucidate the remarkable



“They say the sea is cold,
but the sea contains
the hottest blood of all,
and the wildest, the most urgent.”

—D. H. Lawrence,
“Whales Weep Not!”

Dawn breaks over the Tethys Sea, 48 million years ago, and the blue-green water sparkles with the day’s first light. But for one small mammal, this new day will end almost as soon as it has started.

ANCIENT WHALE *Rodhocetus* (right and left front) feasts on the bounty of the sea, while *Ambulocetus* (rear) attacks a small land mammal some 48 million years ago in what is now Pakistan.

Seas

evolutionary history of whales

By Kate Wong

Tapir-like *Eotitanops* has wandered perilously close to the water's edge, ignoring its mother's warning call. For the brute lurking motionless among the mangroves, the opportunity is simply too good to pass up. It lunges landward, propelled by powerful hind limbs, and sinks its formidable teeth into the calf, dragging it back into the surf. The victim's frantic struggling subsides as it drowns, trapped in the viselike jaws of its captor. Victorious, the beast shambles out of the water to devour its kill on terra firma. At first glance, this fearsome predator resembles a crocodile, with its squat legs, stout tail, long snout and eyes that sit high on its skull. But on closer inspection, it has not armor but fur, not claws but hooves. And the cusps on its teeth clearly identify it not as a reptile but as a mammal. In fact, this improbable creature is *Ambulocetus*, an early whale, and one of a series of intermediates linking the land-dwelling ancestors of cetaceans to the 80 or so species of whales, dolphins and porpoises that rule the oceans today.

Until recently, the emergence of whales was one of the most intractable mysteries facing evolutionary biologists. Lacking fur and hind limbs and unable to go ashore for so much as a sip of freshwater, living cetaceans represent a dramatic departure from the mammalian norm. Indeed, their piscine form led Herman Melville in 1851 to describe Moby Dick and his fellow whales as fishes. But to 19th-century naturalists such as Charles

Darwin, these air-breathing, warm-blooded animals that nurse their young with milk distinctly grouped with mammals. And because ancestral mammals lived on land, it stood to reason that whales ultimately descended from a terrestrial ancestor. Exactly how that might have happened, however, eluded scholars. For his part, Darwin noted in *On the Origin of Species* that a bear swimming with its mouth agape to catch insects was a plausible evolutionary starting point for whales. But the proposition attracted so much ridicule that in later editions of the book he said just that such a bear was "almost like a whale."

The fossil record of cetaceans did little to advance the study of whale origins. Of the few remains known, none were sufficiently complete or primitive to throw much light on the matter. And further analyses of the bizarre anatomy of living whales led only to more scientific head scratching. Thus, even a century after Darwin, these aquatic mammals remained an evolutionary enigma. In fact, in his 1945 classification of mammals, famed paleontologist George Gaylord Simpson noted that whales had evolved in the oceans for so long that nothing informative about their ancestry remained. Calling them "on the whole, the most peculiar and aberrant of mammals," he inserted cetaceans arbitrarily among the other orders. Where whales belonged in the mammalian family tree and how they took to the seas defied explanation, it seemed.

Over the past two decades, however, many of the pieces of this once imponderable puzzle have fallen into place. Paleontologists have uncovered a wealth of whale fossils spanning the Eocene epoch, the time between 55 million and 34 million years ago when archaic whales, or archaeocetes, made their transition from land to sea. They have also unearthed some clues from the ensuing Oligocene, when the modern suborders of cetaceans—the mysticetes (baleen whales) and the odontocetes (toothed whales)—arose. That fossil material, along with analyses of DNA from living animals, has enabled scientists to paint a detailed picture of when, where and how whales evolved from their terrestrial forebears. Today their transformation—from landlubbers to Leviathans—stands as one of the most profound evolutionary metamorphoses on record.

Evolving Ideas

AT AROUND THE SAME TIME that Simpson declared the relationship of whales to other mammals undecipherable on the basis of anatomy, a new comparative approach emerged, one that looked at antibody-antigen reactions in living animals. In response to Simpson's assertion, Alan Boyden of Rutgers University and a colleague applied the technique to the whale question. Their results showed convincingly that among living animals, whales are most closely related to the even-toed hoofed

Guide to Terminology

CETACEA is the order of mammals that comprises living whales, dolphins and porpoises and their extinct ancestors, the archaeocetes. Living members fall into two suborders: the odontocetes, or toothed whales, including sperm whales, pilot whales, belugas, and all dolphins and porpoises; and the mysticetes, or baleen whales, including blue whales and fin whales. The term "whale" is often used to refer to all cetaceans.

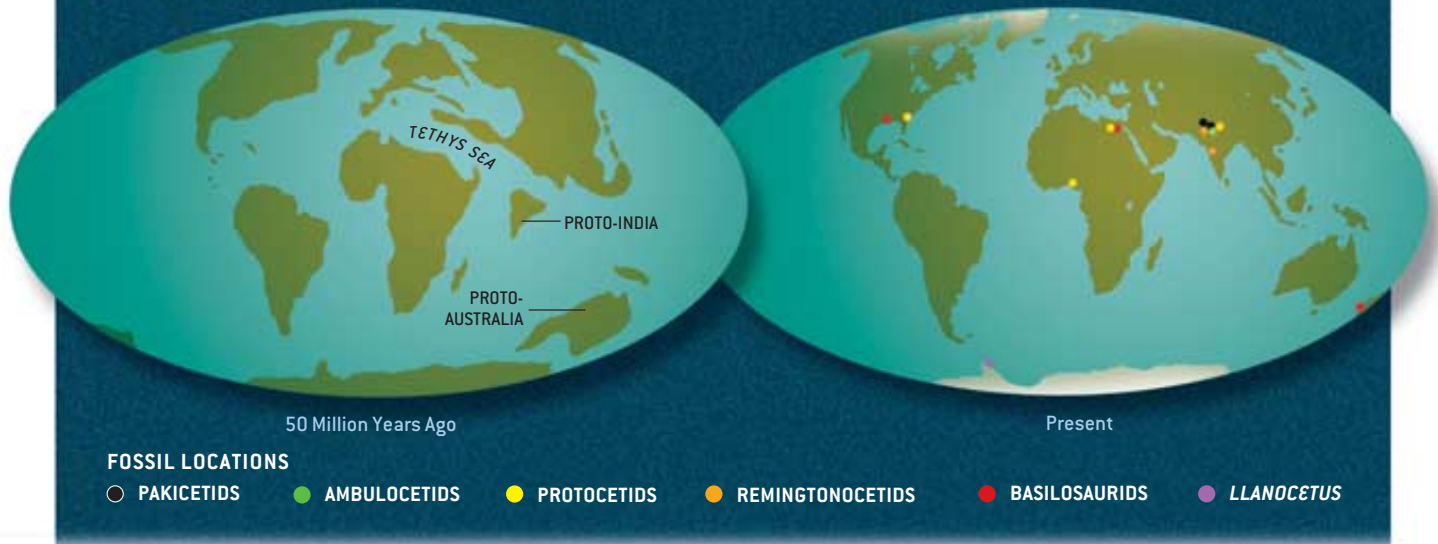
MESONYCHIDS are a group of primitive hoofed, wolflike mammals once widely thought to have given rise to whales.

ARTIODACTYLA is the order of even-toed, hoofed mammals that includes camels; ruminants such as cows; hippos; and, most researchers now agree, whales.

EOCENE is the epoch between 55 million and 34 million years ago, during which early whales made their transition from land to sea.

OLIGOCENE is the epoch between 34 million and 24 million years ago, during which odontocetes and mysticetes evolved from their archaeocete ancestors.

THE WHALE'S CHANGING WORLD



It might seem odd that 300 million years after vertebrates first established a toehold on land, some returned to the sea. But the setting in which early whales evolved offers hints as to what lured them back to the water. For much of the Eocene epoch (roughly between 55 million and 34 million years ago), a sea called Tethys, after a goddess of Greek mythology, stretched from Spain to Indonesia. Although the continents and ocean plates we know now had taken shape, India was still adrift, Australia hadn't yet fully separated from Antarctica, and great swaths of Africa and Eurasia lay submerged under Tethys. Those shallow, warm waters incubated abundant nutrients and teemed with fish. Furthermore, the space vacated by the plesiosaurs, mosasaurs and other large marine reptiles that perished along with the dinosaurs created room for new top predators (although sharks and crocodiles still provided a healthy dose of competition). It is difficult to imagine a more enticing invitation to aquatic life for a mammal.

During the Oligocene epoch that followed, sea levels sank and India docked with the rest of Asia, forming the crumpled interface we know as the Himalayas. More important, University of Michigan paleontologist Philip Gingerich notes, Australia and Antarctica divorced, opening up the Southern Ocean and creating a south circumpolar current that eventually transformed the balmy Eocene earth into the ice-capped planet we inhabit today. The modern current and

climate systems brought about radical changes in the quantity and distribution of nutrients in the sea, generating a whole new set of ecological opportunities for the cetaceans.

As posited by paleontologist Ewan Fordyce of the University of Otago in New Zealand, that set the stage for the replacement of the archaeocetes by the odontocetes and mysticetes (toothed and baleen whales, respectively). The earliest known link between archaeocetes and the modern cetacean orders, Fordyce says, is *Llanocetus*, a 34-million-year-old protobaleen whale from Antarctica that may well have trawled for krill in the chilly Antarctic waters, just as living baleen whales do. Odontocetes arose at around the same time, he adds, specializing to become echolocators that could hunt in the deep.

Unfortunately, fossils documenting the origins of mysticetes and odontocetes are vanishingly rare. Low sea levels during the middle Oligocene exposed most potential whale-bearing sediments from the early Oligocene to erosive winds and rains, making that period largely "a fossil wasteland," says paleontologist Mark Uhen of the Cranbrook Institute of Science in Bloomfield Hills, Mich. The later fossil record clearly shows, however, that shortly after, by about 30 million years ago, the baleen and toothed whales had diversified into many of the cetacean families that reign over the oceans today.

—K.W.

mammals, or artiodactyls, a group whose members include camels, hippopotamuses, pigs and ruminants such as cows. Still, the exact nature of that relationship remained unclear. Were whales themselves artiodactyls? Or did they occupy their own branch of the mammalian family tree, linked to the artiodactyl branch via an ancient common ancestor?

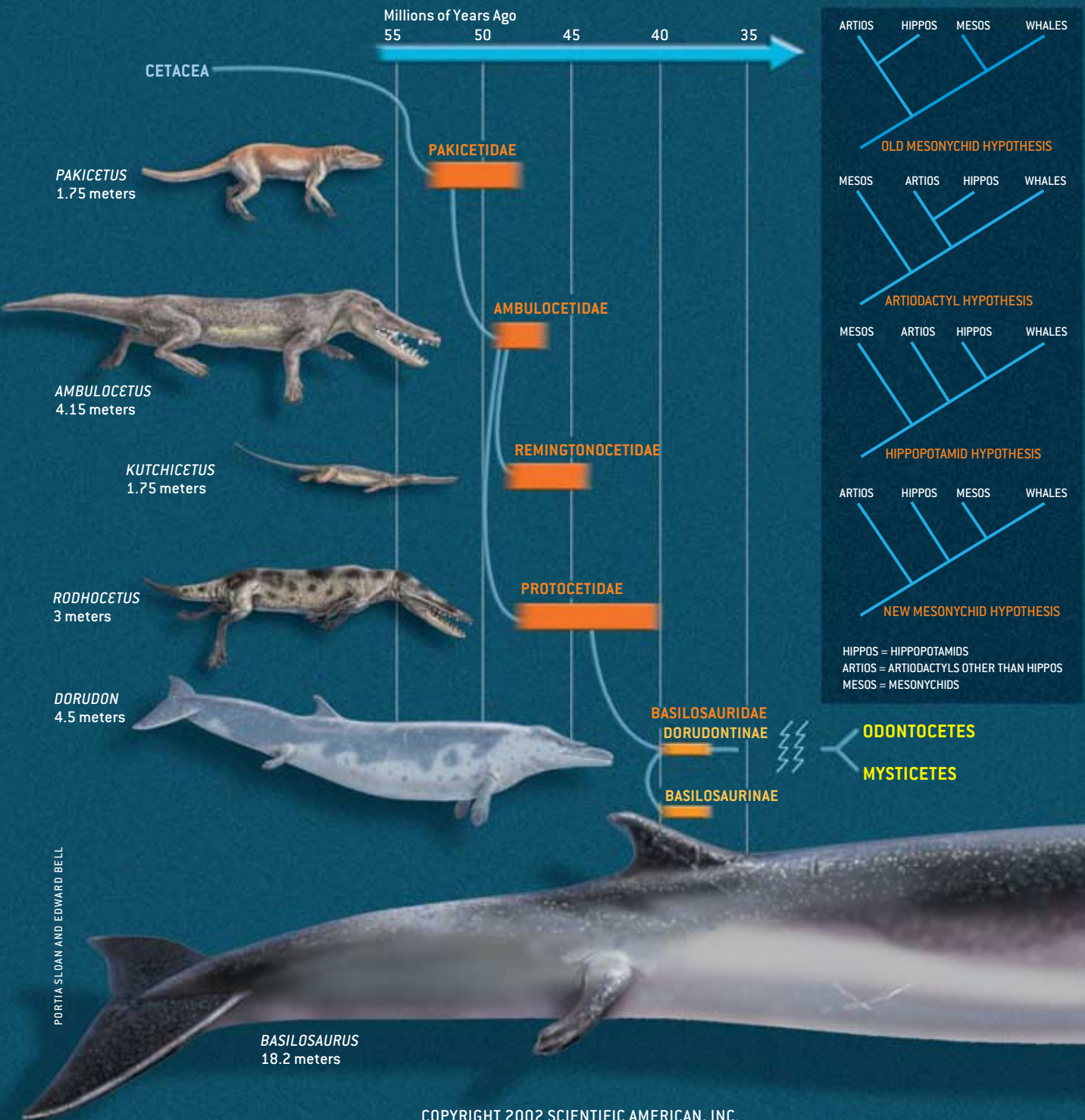
Support for the latter interpretation came in the 1960s, from studies of primitive hoofed mammals known as condylarths that had not yet evolved the specialized characteristics of artiodactyls or the other mammalian orders. Paleontologist

Leigh Van Valen, then at the American Museum of Natural History in New York City, discovered striking resemblances between the three-cusped teeth of the few known fossil whales and those of a group of meat-eating condylarths called mesonychids. Likewise, he found shared dental characteristics between artiodactyls and another group of condylarths, the arctocyonids, close relatives of the mesonychids. Van Valen concluded that whales descended from the carnivorous, wolflike mesonychids and thus were linked to artiodactyls through the condylarths.

CETACEAN RELATIONS

FAMILY TREE OF CETACEANS shows the descent of the two modern suborders of whales, the odontocetes and mysticetes, from the extinct archaeocetes. Representative members of each archaeocete family or subfamily are depicted (*left*). Branching diagrams illustrate various hypotheses of the relationship of whales to other mammals (*right*). The old mesonychid hypothesis, which posits that extinct wolflike beasts known as mesonychids are the closest relatives of whales, now seems unlikely in light of new fossil whale discoveries. The anklebones of those ancient whales bear the distinctive characteristics of artiodactyl ankles, suggesting that whales are

themselves artiodactyls, as envisioned by the artiodactyl hypothesis. Molecular studies indicate that whales are more closely related to hippopotamuses than to any other artiodactyl group. Whether the fossil record can support the hippopotamid hypothesis, however, remains to be seen. A fourth scenario, denoted here as the new mesonychid hypothesis, proposes that mesonychids could still be the whale's closest kin if they, too, were included in the artiodactyl order, instead of the extinct order Condylarthra, in which they currently reside. If so, they would have to have lost the ankle traits that characterize all known artiodactyls. —K.W.



PORTIA SLOAN AND EDWARD BELL

Walking Whales

A DECADE OR SO PASSED before paleontologists finally began unearthing fossils close enough to the evolutionary branching point of whales to address Van Valen's mesonychid hypothesis. Even then the significance of these finds took a while to sink in. It started when University of Michigan paleontologist Philip Gingerich went to Pakistan in 1977 in search of Eocene land mammals, visiting an area previously reported to shelter such remains. The expedition proved disappointing because the spot turned out to contain only marine fossils. Finding traces of ancient ocean life in Pakistan, far from the country's modern coast, is not surprising: during the Eocene, the vast Tethys Sea periodically covered great swaths of what is now the Indian subcontinent [see box on page 73]. Intriguingly, though, the team discovered among those ancient fish and snail remnants two pelvis fragments that appeared to have come from relatively large, walking beasts. "We joked about walking whales," Gingerich recalls with a chuckle. "It was unthinkable." Curious as the pelvis pieces were, the only fossil collected during that field season that seemed important at the time was a primitive artiodactyl jaw that had turned up in another part of the country.

Two years later, in the Himalayan foothills of northern Pakistan, Gingerich's team found another weird whale clue: a partial braincase from a wolf-size creature—found in the company of 50-million-year-old land mammal remains—that bore some distinctive cetacean characteristics. All modern whales have features in their ears that do not appear in any other vertebrates. Although the fossil skull lacked the anatomy necessary for hearing directionally in water (a critical skill for living whales), it clearly had the diagnostic cetacean ear traits. The team had discovered the oldest and most primitive whale then known—one that must have spent some, if not most, of its time on land. Gingerich christened the creature *Pakicetus* for its place of origin and, thus hooked, began hunting for ancient whales in earnest.

At around the same time, another group recovered additional remains of *Pakicetus*—a lower jaw fragment and some isolated teeth—that bolstered the link to mesonychids through strong dental similarities. With *Pakicetus* showing up around 50 million years ago and mesonychids known from around the same time in the same part of the world, it looked increasingly likely that cetaceans had indeed descended from the mesonychids or something closely related to them. Still, what the earliest whales looked like from the neck down was a mystery.

Further insights from Pakistan would have to wait,

however. By 1983 Gingerich was no longer able to work there because of the Soviet Union's invasion of Afghanistan. He decided to cast his net in Egypt instead, journeying some 95 miles southwest of Cairo to the Western Desert's Zeuglodon Valley, so named for early 20th-century reports of fossils of archaic whales—or zeuglodon, as they were then known—in the area. Like Pakistan, much of Egypt once lay submerged under Tethys. Today the skeletons of creatures that swam in that ancient sea lie entombed in sandstone. After several field seasons, Gingerich and his crew hit pay dirt: tiny hind limbs belonging to a 60-foot-long sea snake of a whale known as *Basilosaurus* and the first evidence of cetacean feet.

Earlier finds of *Basilosaurus*, a fully aquatic monster that slithered through the seas between some 40 million and 37 million years ago, preserved only a partial femur, which its discoverers interpreted as vestigial. But the well-formed legs and feet revealed by this discovery hinted at functionality. Although at less than half a meter in length the diminutive limbs probably would not have assisted *Basilosaurus* in swimming and certainly would not have enabled it to walk on land, they may well have helped guide the beast's serpentine body during the difficult activity of aquatic mating. Whatever their purpose, if any, the little legs had big implications. "I immediately thought, we're 10 million years after *Pakicetus*," Gingerich recounts excitedly. "If these things still have feet and toes, we've got 10 million years of history to look at." Suddenly, the walking whales they had scoffed at in Pakistan seemed entirely plausible.

Just such a remarkable creature came to light in 1992. A team led by J.G.M. (Hans) Thewissen of the Northeastern Ohio Universities College of Medicine recovered from 48-million-year-old marine rocks in northern Pakistan a nearly complete skeleton of a perfect intermediate between modern whales and their terrestrial ancestors. Its large feet and powerful tail bespoke strong swimming skills, while its sturdy leg bones and mobile elbow and wrist joints suggested an ability to locomote on land. He dubbed the animal *Ambulocetus natans*, the walking and swimming whale.

Shape Shifters

SINCE THEN, Thewissen, Gingerich and others have unearthed a plethora of fossils documenting subsequent stages of the whale's transition from land to sea. The picture emerging from those specimens is one in which *Ambulocetus* and its kin—themselves descended from the more terrestrial pakicetids—spawned needle-nosed beasts known as remingtonocetids and the intrepid protocetids—the first whales seaworthy enough to fan out from Indo-Pakistan across the globe. From the protocetids arose the dolphinlike dorudontines, the probable progenitors of the snakelike basilosaurines and modern whales [see box on opposite page].

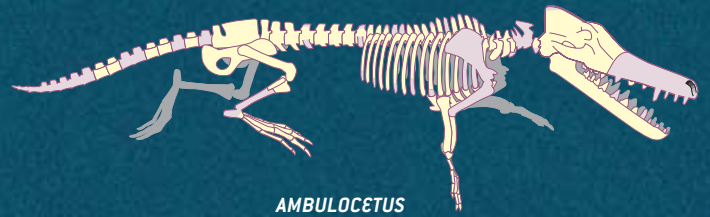
In addition to furnishing supporting branches for the whale family tree, these discoveries have enabled researchers to chart many of the spectacular anatomical and physiological changes that allowed cetaceans to establish permanent residency in the



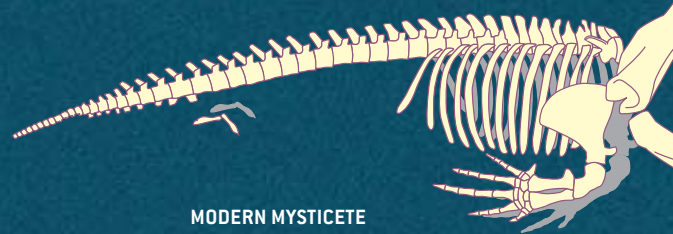
BECOMING LEVIATHAN



PAKICETUS



AMBULOCETUS



MODERN MYSTICETE

REPRESENTATIVE ARCHAEOCETES in the lineage leading to modern odontocetes and mysticetes trace some of the anatomical changes that enabled these animals to take to the seas (reconstructed bone appears in lavender). In just 15 million years, whales shed their terrestrial trappings and became fully adapted to aquatic life. Notably, the hind limbs diminished, the forelimbs transformed into flippers, and the vertebral column evolved to permit tail-powered swimming. Meanwhile the skull changed to enable underwater hearing, the nasal opening moved backward to the top of the skull, and the teeth simplified into pegs for grasping instead of grinding. Later in whale evolution, the mysticetes' teeth were replaced with baleen.

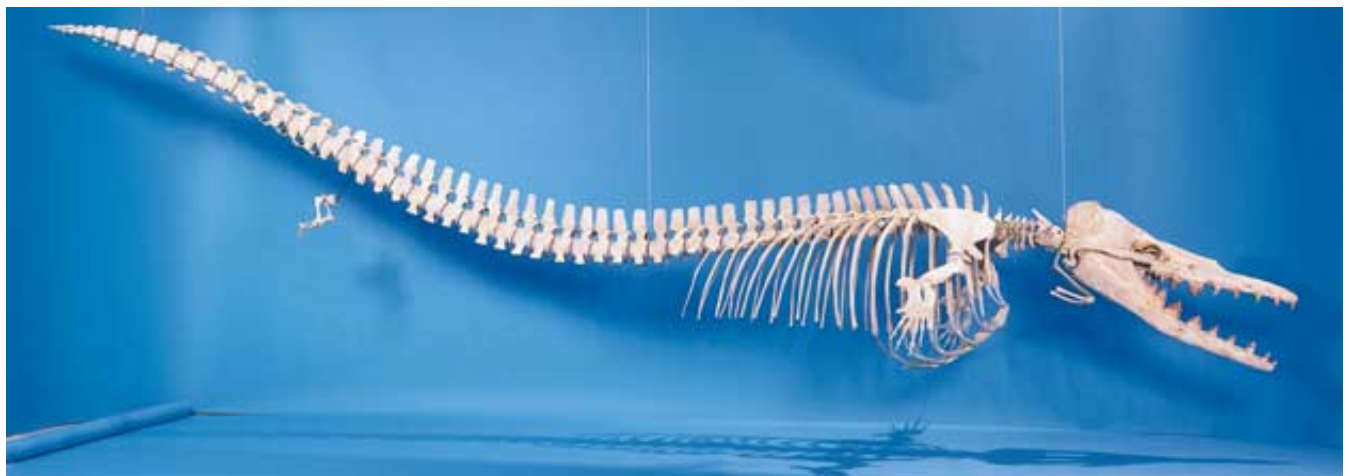
ocean realm. Some of the earliest of these adaptations to emerge, as *Pakicetus* shows, are those related to hearing. Sound travels differently in water than it does in air. Whereas the ears of humans and other land-dwelling animals have delicate, flat eardrums, or tympanic membranes, for receiving airborne sound, modern whales have thick, elongate tympanic ligaments that cannot receive sound. Instead a bone called the bulla, which in whales has become quite dense and is therefore capable of transmitting sound coming from a denser medium to deeper parts of the ear, takes on that function. The *Pakicetus* bulla shows some modification in that direction, but the animal retained a land mammal-like eardrum that could not work in water.

What, then, might *Pakicetus* have used its thickened bullae

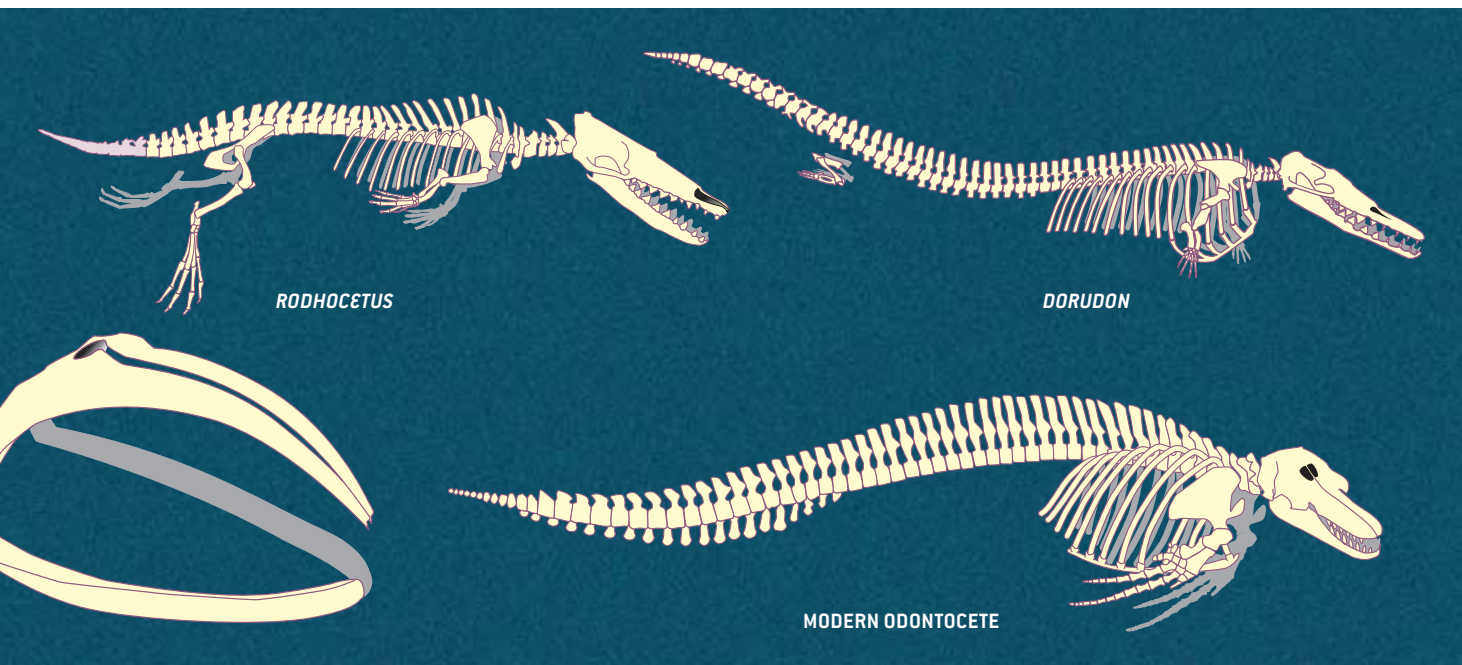
for? Thewissen suspects that much as turtles hear by picking up vibrations from the ground through their shields, *Pakicetus* may have employed its bullae to pick up ground-borne sounds. Taking new postcranial evidence into consideration along with the ear morphology, he envisions *Pakicetus* as an ambush predator that may have lurked around shallow rivers, head to the ground, preying on animals that came to drink. *Ambulocetus* is even more likely to have used such inertial hearing, Thewissen says, because it had the beginnings of a channel linking jaw and ear. By resting its jaw on the ground—a strategy seen in modern crocodiles—*Ambulocetus* could have listened for approaching prey. The same features that allowed early whales to receive sounds from soil, he surmises, preadapted them to hearing in the water.

Zhe-Xi Luo of the Carnegie Museum of Natural History in Pittsburgh has shown that by the time of the basilosaurines and dorudontines, the first fully aquatic whales, the ropelike tympanic ligament had probably already evolved. Additionally, air

DORUDON, a 4.5-meter-long, dolphinlike archaeocete that patrolled the seas between roughly 40 million and 37 million years ago, may be the ancestor of modern whales.



PORTIA SLOAN (illustrations); UNIVERSITY OF MICHIGAN EXHIBIT MUSEUM OF NATURAL HISTORY (photograph)



sinuses, presumably filled with spongelike tissues, had formed around the middle ear, offering better sound resolution and directional cues for underwater hearing. Meanwhile, with the external ear canal closed off (a prerequisite for deep-sea diving), he adds, the lower jaw was taking on an increasingly important auditory role, developing a fat-filled canal capable of conducting sound back to the middle ear.

Later in the evolution of whale hearing, the toothed and baleen whales parted ways. Whereas the toothed whales evolved the features necessary to produce and receive high-frequency sounds, enabling echolocation for hunting, the baleen whales developed the ability to produce and receive very low frequency sounds, allowing them to communicate with one another over vast distances. Fossil whale ear bones, Luo says, show that by around 28 million years ago early odontocetes already had some of the bony structures necessary for hearing high-pitched sound and were thus capable of at least modest echolocation. The origin of the mysticete's low-frequency hearing is far murkier, even though the fossil evidence of that group now dates back to as early as 34 million years ago.

Other notable skull changes include movement of the eye sockets from a crocodilelike placement atop the head in *Pakicetus* and *Ambulocetus* to a lateral position in the more aquatic protocetids and later whales. And the nasal opening migrated back from the tip of the snout in *Pakicetus* to the top of the head in modern cetaceans, forming the blowhole. Whale dentition morphed, too, turning the complexly cusped, grinding molars of primitive mammalian ancestors into the simple, pronglike teeth of modern odontocetes, which grasp and swallow their food without chewing. Mysticetes lost their teeth altogether and developed comblike plates of baleen that hang from their upper jaws and strain plankton from the seawater.

The most obvious adaptations making up the whale's pro-

tean shift are those that produced its streamlined shape and unmatched swimming abilities. Not surprisingly, some bizarre amphibious forms resulted along the way. *Ambulocetus*, for one, retained the flexible shoulder, elbow, wrist and finger joints of its terrestrial ancestors and had a pelvis capable of supporting its weight on land. Yet the creature's disproportionately large hind limbs and paddlelike feet would have made walking somewhat awkward. These same features were perfect for paddling around in the fish-filled shallows of Tethys, however.

Moving farther out to sea required additional modifications, many of which appear in the protocetid whales. Studies of one member of this group, *Rodhocetus*, indicate that the lower arm bones were compressed and already on their way to becoming hydrodynamically efficient, says University of Michigan paleontologist Bill Sanders. The animal's long, delicate feet were probably webbed, like the fins used by scuba divers. *Rodhocetus* also exhibits aquatic adaptations in its pelvis, where fusion between the vertebrae that form the sacrum is reduced, loosening up the lower spine to power tail movement. These features, says Gingerich, whose team discovered the creature, suggest that *Rodhocetus* performed a leisurely dog paddle at the sea surface and a swift combination of otterlike hind-limb paddling and tail propulsion underwater. When it went ashore to breed or perhaps to bask in the sun, he proposes, *Rodhocetus* probably hitched itself around somewhat like a modern eared seal or sea lion.

By the time of the basilosaurines and dorudontines, whales were fully aquatic. As in modern cetaceans, the shoulder remained mobile while the elbow and wrist stiffened, forming flippers for steering and balance. Farther back on the skeleton, only tiny legs remained, and the pelvis had dwindled accordingly. Analyses of the vertebrae of *Dorudon*, conducted by Mark D. Uhen of the Cranbrook Institute of Science in Bloomfield Hills, Mich., have revealed one tail vertebra with a rounded profile.

Modern whales have a similarly shaped bone, the ball vertebra, at the base of their fluke, the flat, horizontal structure capping the tail. Uhen thus suspects that basilosaurines and dorudontines had tail flukes and swam much as modern whales do, using so-called caudal oscillation. In this energetically efficient mode of locomotion, motion generated at a single point in the vertebral column powers the tail's vertical movement through the water, and the fluke generates lift.

Exactly when whales lost their legs altogether remains un-

WATER, WATER EVERYWHERE

MOST MAMMALS—big ones in particular—cannot live without freshwater. For marine mammals, however, freshwater is difficult to come by. Seals and sea lions obtain most of their water from the fish they eat (some will eat snow to get freshwater), and manatees routinely seek out freshwater from rivers. For their part, cetaceans obtain water both from their food and from sips of the briny deep.

When did whales, which evolved from a fairly large (and therefore freshwater-dependent) terrestrial mammal, develop a system capable of handling the excess salt load associated with ingesting seawater? Evidence from so-called stable oxygen isotopes has provided some clues. In nature, oxygen mainly occurs in two forms, or isotopes: ^{16}O and ^{18}O . The ratios of these isotopes in freshwater and seawater differ, with seawater containing more ^{18}O . Because mammals incorporate oxygen from drinking water into their developing teeth and bones, the remains of those that imbibe seawater can be distinguished from those that take in freshwater.

J.G.M. (Hans) Thewissen of the Northeastern Ohio Universities College of Medicine and his colleagues thus analyzed the oxygen isotope ratios in ancient whale teeth to gain insight into when these animals might have moved from a freshwater-based osmoregulatory system to a seawater-based one. Oxygen isotope values for pakicetids, the most primitive whales, indicate that they drank freshwater, as would be predicted from other indications that these animals spent much of their time on land. Isotope measurements from amphibious *Ambulocetus*, on the other hand, vary widely, and some specimens show no evidence of seawater intake. In explanation, the researchers note that although *Ambulocetus* is known to have spent time in the sea (based on the marine nature of the rocks in which its fossils occur), it may still have had to go ashore to drink. Alternatively, it may have spent the early part of its life (when its teeth mineralized) in freshwater and only later entered the sea.

The protocetids, however, which show more skeletal adaptations to aquatic life, exhibit exclusively marine isotope values, indicating that they drank only seawater. Thus, just a few million years after the first whales evolved, their descendants had adapted to increased salt loads. This physiological innovation no doubt played an important role in facilitating the protocetids' dispersal across the globe. —K.W.

known. In fact, a recent discovery made by Lawrence G. Barnes of the Natural History Museum of Los Angeles County hints at surprisingly well developed hind limbs in a 27-million-year-old baleen whale from Washington State, suggesting that whale legs persisted far longer than originally thought. Today, however, some 50 million years after their quadrupedal ancestors first waded into the warm waters of Tethys, whales are singularly sleek. Their hind limbs have shrunk to externally invisible vestiges, and the pelvis has diminished to the point of serving merely as an anchor for a few tiny muscles unrelated to locomotion.

Making Waves

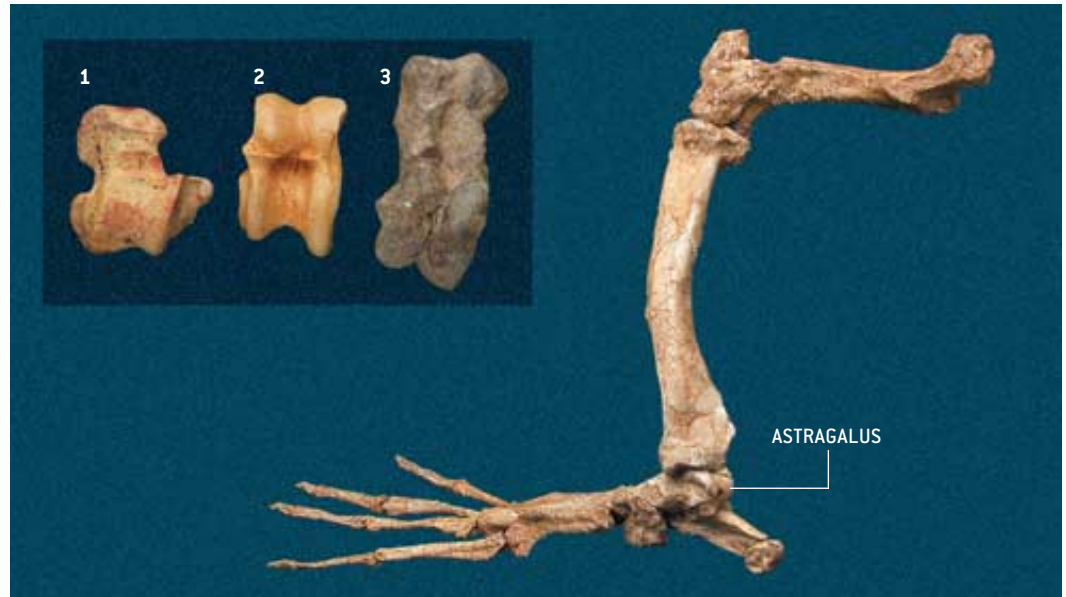
THE FOSSILS UNCOVERED during the 1980s and 1990s advanced researchers' understanding of whale evolution by leaps and bounds, but all morphological signs still pointed to a mesonychid origin. An alternative view of cetacean roots was taking wing in genetics laboratories in the U.S., Belgium and Japan, however. Molecular biologists, having developed sophisticated techniques for analyzing the DNA of living creatures, took Boyden's 1960s immunology-based conclusions a step further. Not only were whales more closely related to artiodactyls than to any other living mammals, they asserted, but in fact whales were themselves artiodactyls, one of many twigs on that branch of the mammalian family tree. Moreover, a number of these studies pointed to an especially close relationship between whales and hippopotamuses. Particularly strong evidence for this idea came in 1999 from analyses of snippets of noncoding DNA called SINES (short interspersed elements), conducted by Norihiro Okada and his colleagues at the Tokyo Institute of Technology.

The whale-hippo connection did not sit well with paleontologists. "I thought they were nuts," Gingerich recalls. "Everything we'd found was consistent with a mesonychid origin. I was happy with that and happy with a connection through mesonychids to artiodactyls." Whereas mesonychids appeared at the right time, in the right place and in the right form to be considered whale progenitors, the fossil record did not seem to contain a temporally, geographically and morphologically plausible artiodactyl ancestor for whales, never mind one linking whales and hippos specifically. Thewissen, too, had largely dismissed the DNA findings. But "I stopped rejecting it when Okada's SINE work came out," he says.

It seemed the only way to resolve the controversy was to find, of all things, an ancient whale anklebone. Morphologists have traditionally defined artiodactyls on the basis of certain features in one of their anklebones, the astragalus, that enhance mobility. Specifically, the unique artiodactyl astragalus has two grooved, pulleylike joint surfaces. One connects to the tibia, or shinbone; the other articulates with more distal anklebones. If whales descended from artiodactyls, researchers reasoned, those that had not yet fully adapted to life in the seas should exhibit this double-pulleyed astragalus.

That piece of the puzzle fell into place last fall, when Gingerich and Thewissen both announced discoveries of new primitive whale fossils. In the eastern part of Baluchistan Province,

HIND LIMB of an ancient whale, *Rodhocetus*, preserves a long-sought anklebone known as the astragalus (at right). Shown in the inset beside a mesonychid astragalus [1] and one from a modern artiodactyl [2], the *Rodhocetus* astragalus [3] exhibits the distinctive double-pulley shape that characterizes all artiodactyl astragali, suggesting that whales descended not from mesonychids as previously thought but from an ancient artiodactyl.



Gingerich's team had found partially articulated skeletons of *Rodhocetus balochistanensis* and a new protocetid genus, *Artiocetus*. Thewissen and his colleagues recovered from a bone bed in the Kala Chitta Hills of Punjab, Pakistan, much of the long-sought postcranial skeleton of *Pakicetus*, as well as that of a smaller member of the pakicetid family, *Ichthyolestes*. Each came with an astragalus bearing the distinctive artiodactyl characteristics.

The anklebones convinced both longtime proponents of the mesonychid hypothesis that whales instead evolved from artiodactyls. Gingerich has even embraced the hippo idea. Although hippos themselves arose long after whales, their purported ancestors—dog- to horse-size, swamp-dwelling beasts called anthracotheres—date back to at least the middle Eocene and may thus have a forebear in common with the cetaceans. In fact, Gingerich notes that *Rodhocetus* and anthracotheres share features in their hands and wrists not seen in any other later artiodactyls. Thewissen agrees that the hippo hypothesis holds much more appeal than it once did. But he cautions that the morphological data do not yet point to a particular artiodactyl, such as the hippo, being the whale's closest relative, or sister group. "We don't have the resolution yet to get them there," he remarks, "but I think that will come."

What of the evidence that seemed to tie early whales to mesonychids? In light of the new ankle data, most workers now suspect that those similarities probably reflect convergent evolution rather than shared ancestry and that mesonychids represent an evolutionary dead end. But not everyone is convinced. Maureen O'Leary of the State University of New York at Stony Brook argues that until all the available evidence—both morphological and molecular—is incorporated into a single phylogenetic analysis, the possibility remains that mesonychids belong at the base of the whale pedigree. It is conceivable, she says, that mesonychids are actually ancient artiodactyls but ones that reversed the ankle trend. If so, mesonychids could still be the

whales' closest relative, and hippos could be their closest living relative [see box on page 74]. Critics of that idea, however, point out that although folding the mesonychids into the artiodactyl order offers an escape hatch of sorts to supporters of the mesonychid hypothesis, it would upset the long-standing notion that the ankle makes the artiodactyl.

Investigators agree that figuring out the exact relationship between whales and artiodactyls will most likely require finding additional fossils—particularly those that can illuminate the beginnings of artiodactyls in general and hippos in particular. Yet even with those details still unresolved, "we're really getting a handle on whales from their origin to the end of archaeocetes," Uhen reflects. The next step, he says, will be to figure out how the mysticetes and odontocetes arose from the archaeocetes and when their modern features emerged. Researchers may never unravel all the mysteries of whale origins. But if the extraordinary advances made over the past two decades are any indication, with continued probing, answers to many of these lingering questions will surface from the sands of time. SA

Kate Wong is a writer and editor for *ScientificAmerican.com*

MORE TO EXPLORE

The Emergence of Whales: Evolutionary Patterns in the Origin of Cetacea. Edited by J.G.M. Thewissen. Plenum Publishing, 1998.

Skeletons of Terrestrial Cetaceans and the Relationship of Whales to Artiodactyls. J.G.M. Thewissen, E. M. Williams, L. J. Roe and S. T. Hussain in *Nature*, Vol. 413, pages 277–281; September 20, 2001.

Origin of Whales from Early Artiodactyls: Hands and Feet of Eocene Protocetidae from Pakistan. Philip D. Gingerich, Munir ul Haq, Iyad S. Zalmout, Intizar Hussain Khan and M. Sadiq Malkani in *Science*, Vol. 293, pages 2239–2242; September 21, 2001.

The Encyclopedia of Marine Mammals. Edited by W. F. Perrin, Bernd G. Würsig and J.G.M. Thewissen. Academic Press, 2002.

A broadcast version of this article will run on *National Geographic Today*, a show on the National Geographic Channel. Please check your local listings.

Genetic analysis of hair samples attributed to yeti, bigfoot and other anomalous primates

Bryan C. Sykes, Rhettman A. Mullis, Christophe Hagenmuller, Terry W. Melton and Michel Sartori

Proc. R. Soc. B 2014 **281**, 20140161, published 2 July 2014

References

[This article cites 11 articles, 3 of which can be accessed free](#)

<http://rspb.royalsocietypublishing.org/content/281/1789/20140161.full.html#ref-list-1>

[Article cited in:](#)

<http://rspb.royalsocietypublishing.org/content/281/1789/20140161.full.html#related-urls>

 **open access**

This article is free to access

Subject collections

Articles on similar topics can be found in the following collections

[evolution](#) (1869 articles)

[genetics](#) (145 articles)

[taxonomy and systematics](#) (198 articles)

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)



Cite this article: Sykes BC, Mullis RA, Hagenmuller C, Melton TW, Sartori M. 2014 Genetic analysis of hair samples attributed to yeti, bigfoot and other anomalous primates. *Proc. R. Soc. B* **281**: 20140161. <http://dx.doi.org/10.1098/rsob.2014.0161>

Received: 21 January 2014

Accepted: 27 March 2014

Subject Areas:

genetics, taxonomy and systematics, evolution

Keywords:

yeti, almasty, bigfoot, sasquatch, mitochondrial DNA

Author for correspondence:

Bryan C. Sykes

e-mail: bryan.sykes@wolfsonox.ac.uk

Genetic analysis of hair samples attributed to yeti, bigfoot and other anomalous primates

Bryan C. Sykes¹, Rhettman A. Mullis², Christophe Hagenmuller³, Terry W. Melton⁴ and Michel Sartori^{5,6}

¹Institute of Human Genetics, Wolfson College, University of Oxford, Oxford OX2 6UD, UK

²PO Box 40143, Bellevue, WA 98005, USA

³NaturAlpes, Annecy-Le-Vieux 74940, France

⁴Mitotyping Technologies, 2565 Park Center Boulevard, State College, PA 16801, USA

⁵Museum of Zoology, Palais de Rumine, Lausanne 1014, Switzerland

⁶Museum of Zoology and Grindel Biocentre, Hamburg 20146, Germany

In the first ever systematic genetic survey, we have used rigorous decontamination followed by mitochondrial 12S RNA sequencing to identify the species origin of 30 hair samples attributed to anomalous primates. Two Himalayan samples, one from Ladakh, India, the other from Bhutan, had their closest genetic affinity with a Palaeolithic polar bear, *Ursus maritimus*. Otherwise the hairs were from a range of known extant mammals.

1. Introduction

Despite several decades of research, mystery still surrounds the species identity of so-called anomalous primates such as the yeti in the Himalaya, almasty in central Asia and sasquatch/bigfoot in North America. On the one hand, numerous reports including eye-witness and footprint evidence, point to the existence of large unidentified primates in many regions of the world. On the other hand, no bodies or recent fossils of such creatures have ever been authenticated. There is no shortage of theories about what these animals may be, ranging from surviving populations of collateral hominids such as *Homo neanderthalensis*, *Homo floresiensis* [1] or Denisovans [2], extinct apes such as *Gigantopithecus* [3] or even unlikely hybrids between *Homo sapiens* and other mammals [4]. Modern science has largely avoided this field and advocates frequently complain that they have been 'rejected by science' [5]. This conflicts with the basic tenet that science neither rejects nor accepts anything without examining the evidence. To apply this philosophy to the study of anomalous primates and to introduce some clarity into this often murky field, we have carried out a systematic genetic survey of hair samples attributed to these creatures. Only two 'tongue-in-cheek' scientific publications report DNA sequence data from anomalous primates. Milinkovitch *et al.* [6], after analysis of a Nepalese sample, confirmed Captain Haddock's suspicions that the yeti was an ungulate [7]. The same conclusion was reached by Coltman *et al.* [8] after analysis of sasquatch hair from Alaska.

2. Material and methods

Hair samples submissions were solicited from museum and individual collections in a joint press release issued on 14 May 2012 by the Museum of Zoology, Lausanne and the University of Oxford. A total of 57 samples were received and subjected to macroscopic, microscopic and infrared fluorescence examination to eliminate obvious non-hairs. This excluded one sample of plant material and one of glass fibre. Of the screened samples, 37 were selected for genetic analysis based on their provenance or historic interest. Lengths (2–4 cm) of individual hair shaft

Table 1. Origin and GenBank sequence matches of hair samples attributed to anomalous primates. (All sequence matches were 100%.)

ref. no.	location	attribution	GenBank sequence match	common name
25025	Ladakh, India	yeti	<i>U. maritimus</i>	polar bear
25191	Bhutan	yeti/migyhur	<i>U. maritimus</i>	polar bear
25092	Nepal	yeti	<i>Capricornis sumatraensis</i>	serow
25027	Russia	almasty	<i>U. arctos</i>	brown bear
25039	Russia	almasty	<i>Equus caballus</i>	horse
25040	Russia	almasty	<i>Bos taurus</i>	cow
25041	Russia	almasty	<i>Equus caballus</i>	horse
25073	Russia	almasty	<i>Equus caballus</i>	horse
25074	Russia	almasty	<i>U. americanus</i>	American black bear
25075	Russia	almasty	<i>P. lotor</i>	raccoon
25194	Russia	almasty	<i>U. arctos</i>	brown bear
25044	Sumatra	orang pendek	<i>Tapirus indicus</i>	Malaysian tapir
25035	AZ, USA	bigfoot	<i>P. lotor</i>	raccoon
25167	AZ, USA	bigfoot	<i>Ovis aries</i>	sheep
25104	CA, USA	bigfoot	<i>U. americanus</i>	American black bear
25106	CA, USA	bigfoot	<i>U. americanus</i>	American black bear
25081	MN, USA	bigfoot	<i>Erethizon dorsatum</i>	N. American porcupine
25082	MN, USA	bigfoot	<i>U. americanus</i>	American black bear
25202	OR, USA	bigfoot	<i>U. americanus</i>	American black bear
25212	OR, USA	bigfoot	<i>C. lupus/lutrans/domesticus</i>	wolf/coyote/dog
25023	TX, USA	bigfoot	<i>Equus caballus</i>	horse
25072	TX, USA	bigfoot	<i>Homo sapiens</i>	human
25028	WA, USA	bigfoot	<i>U. americanus</i>	American black bear
25029	WA, USA	bigfoot	<i>C. lupus/lutrans/domesticus</i>	wolf/coyote/dog
25030	WA, USA	bigfoot	<i>Bos taurus</i>	cow
25069	WA, USA	bigfoot	<i>Odocoileus virginianus/hemionus</i>	white-tailed/mule deer
25086	WA, USA	bigfoot	<i>Bos taurus</i>	cow
25093	WA, USA	bigfoot	<i>C. lupus/lutrans/domesticus</i>	wolf/coyote/dog
25112	WA, USA	bigfoot	<i>Bos taurus</i>	cow
25113	WA, USA	bigfoot	<i>C. lupus/lutrans/domesticus</i>	wolf/coyote/dog

were thoroughly cleaned to remove surface contamination, ground into a buffer solution in a glass homogenizer then incubated for 2 h at 56°C in a solution containing proteinase K before extraction with phenol/chloroform/isoamyl alcohol. PCR amplification of the ribosomal mitochondrial DNA 12S fragment corresponding to bps 1093–1196 of the human mitochondrial genome was carried out [9,10]. Recovered sequences were compared to GenBank accessions for species identification.

3. Results and discussion

The table 1 shows the GenBank species identification of sequences matching the 30 samples from which DNA was recovered. Seven samples failed to yield any DNA sequences despite multiple attempts. As the sequence of mitochondrial 12S RNA segment is identical in *H. sapiens* and *H. neanderthalensis*, amplification and sequencing of mitochondrial DNA hypervariable region 1 (bps 16 000–16 400) of no. 25072 was carried out and identified the source as being

identical to the revised Cambridge Reference Sequence [11] and thus *H. sapiens* of likely European matrilineal descent. Other submitted samples were of known mammals that in most cases were living within their normal geographical range, the exceptions being sample nos. 25025 and 25191 (*Ursus maritimus*, polar bear) from the Himalayas, no. 25074 (*Ursus americanus*, American black bear) and no. 25075 (*Procyon lotor*, raccoon) that were submitted from Russia even though they are native to North America.

Despite the wide range of age and condition of the submitted hair shafts, which ranged from fresh to museum specimens more than 50 years old, the majority yielded mitochondrial 12S RNA sequences which allowed species identification with 100% sequence identity. Of the recovered sequences, only one (no. 25072) yielded a human sequence, indicating that the rigorous cleaning and extraction protocol had been effective in eliminating extraneous human contamination which often confounds the analysis of old material and may lead to misinterpretation of a sample as human or even as an unlikely and

unknown human x mammalian hybrid [4]. The deliberately permissive primer combination used here allowed a wide range of mammalian DNA to be amplified within a single reaction, although this meant that some identification did not go beyond the level of genus. For example, no. 25029 was identified as *Canis* but did not distinguish between *Canis lupus* (wolf), *Canis latrans* (coyote) and *Canis domesticus* (domestic dog).

Sequences derived from hair sample nos. 25025 and 25191 had a 100% match with DNA recovered from a Pleistocene fossil more than 40 000 BP of *U. maritimus* (polar bear) [12] but not to modern examples of the species. Hair sample no. 25025 came from an animal shot by an experienced hunter in Ladakh, India ca 40 years ago who reported that its behaviour was very different from a brown bear *Ursus arctos* with which he was very familiar. Hair sample no. 25191 was recovered from a high altitude (ca 3500 m) bamboo forest in Bhutan and was identified as a nest of a migyur, the Bhutanese equivalent of the yeti. The Ladakh hairs (no. 25025) were golden-brown, whereas the hair from Bhutan (no. 25191) was reddish-brown in appearance. As the match is to a segment only 104 bp long, albeit in the very conserved 12S RNA gene, this result should be regarded as preliminary. Other than these data, nothing is currently known about the genetic affinity of Himalayan bears and although there are anecdotal reports of white bears in Central Asia and the Himalayas [13,14], it seems more likely that the two hairs reported here are from either a previously unrecognized bear species, colour variants of *U. maritimus*, or *U. arctos/U. maritimus* hybrids. Viable *U. arctos/U. maritimus* hybrids are known from the Admiralty, Barayanov and Chicagov (ABC) islands off the coast of Alaska though in the ABC hybrids the mitochondrial sequence homology is with modern rather than ancient polar bears [15]. If they are hybrids, the Ladakh and Bhutan specimens are probably descended from a different hybridization event during the early stages of species divergence between *U. arctos* and *U. maritimus*. Genomic sequence data are needed to decide between these alternatives. If these bears are widely distributed in the Himalayas, they may well contribute to the biological foundation of the yeti legend,

especially if, as reported by the hunter who shot the Ladakh specimen, they behave more aggressively towards humans than known indigenous bear species.

With the exception of these two samples, none of the submitted and analysed hairs samples returned a sequence that could not be matched with an extant mammalian species, often a domesticate. While it is important to bear in mind that absence of evidence is not evidence of absence and this survey cannot refute the existence of anomalous primates, neither has it found any evidence in support. Rather than persisting in the view that they have been 'rejected by science', advocates in the cryptozoology community have more work to do in order to produce convincing evidence for anomalous primates and now have the means to do so. The techniques described here put an end to decades of ambiguity about species identification of anomalous primate samples and set a rigorous standard against which to judge any future claims.

Acknowledgements. We thank Reinhold Messner, Peter Byrne, Justin Smeja, Bart Cutino, Derek Randles, Dan Shirley, Garland Fields, Loren Coleman, Betty Klopp, Marcel Cagney, Sam Cagney, Lori Simmons, Adam Davies, Dr Mike Amaranthus, Mike Long, Patrick Spell, Maxwell David, Mark McClurkan, Rob Kryder, Jack Barnes, Jeff Anderson, David Ellis, Steve Mattice, Brenda Harris, Stuart Fleming, Igor Burtsev, Dmitri Pirkulov, Michael Trachtengerts and Dmitri Bayanov for submitting samples and for their progressive stance in doing so. Thanks also to, Ray Crowe, Ronnie Roseman, Greg Roberts and Tom Graham for discussing their experiences and to Jeff Meldrum and Anna Nekaris for advice and guidance. We are very grateful to Ken Goddard, Ed Espinoza, Mike Tucker, Barry Baker, Bonnie Yates, Cookie Smith and Dyan Straughan of the US Fish and Wildlife Service Forensic Laboratory, Ashland, OR, USA, for help with forensic methods of trace evidence analysis and to Charity Holland, Bonnie Higgins, Gloria Dimick and Michele Yon for technical assistance.

Data accessibility. DNA sequences: GenBank accession nos. KJ155696–KJ155724 and KJ607607. Voucher samples of the research materials have been deposited in the Heuvelmans Archive at the Museum of Zoology, Lausanne, Switzerland.

Funding statement. We also thank Harry Marshall and Icon Films for their contribution to the costs of analysis.

References

- Brown P, Sutikna T, Morwood M, Soejono R, Jatmiko E, Saptomo E, Awa Due R. 2004 A new small-bodied hominin from the Late Pleistocene of Flores, Indonesia. *Nature* **431**, 1055–1061. (doi:10.1038/nature02999)
- Reich D *et al.* 2010 Genetic history of an archaic hominin group in Denisova Cave in Siberia. *Nature* **468**, 1053–1060. (doi:10.1038/nature09710)
- Strauss W. 1957 Jaw of *Gigantopithecus*. *Science* **125**, 685. (doi:10.1126/science.125.3250.685)
- Ketchum M *et al.* 2013 North American hominins. *De Novo* **1**, 1–15.
- Regal B. 2011 Searching for Sasquatch, p. 5. New York, NY: Palgrave Macmillan.
- Milinkovitch MC *et al.* 2004 Molecular phylogenetic analyses indicate extensive morphological convergence between the 'yeti' and primates. *Mol. Phylogenet. Evol.* **31**, 1–3. (doi:10.1016/j.ympev.2004.01.009)
- Herge. 1960 *Tintin au Tibet*. Tournai, Belgium: Casterman.
- Coltman D, Davis C. 2005 Molecular cryptozoology meets the Sasquatch. *Trends Ecol. Evol.* **21**, 60–61. (doi:10.1016/j.tree.2005.11.010)
- Melton T, Dimick G, Higgins B, Lindstrom Nelson K. 2005 Forensic mitochondrial DNA analysis of 691 casework hairs. *J. Forensic Sci.* **50**, 73–80. (doi:10.1520/JFS2004230)
- Melton T, Holland C. 2007 Routine forensic use of the mitochondrial 12S ribosomal RNA genes for species identification. *J. Forensic Sci.* **52**, 1305–1307.
- Andrews R, Kubacka I, Chinnery C, Lightowlers RN, Turnbull DM, Howell N. 1999 Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* **23**, 147. (doi:10.1038/13779)
- Lingqvist C *et al.* 2010 Complete mitochondrial genome of a Pleistocene jawbone unveils the origin of polar bear. *Proc. Natl Acad. Sci. USA* **107**, 5053–5057. (doi:10.1073/pnas.0914266107)
- Brunner B. 2007 *Bears: a brief history*, p. 64. New Haven, CT: Yale University Press.
- Smythe F. 1936 *The valley of flowers*, p. 144. London, UK: Hodder and Stoughton.
- Hailer F, Kutschera V, Hallström B, Klassert D, Fain S, Leonard J, Arnason U, Janke A. 2012 Nuclear genomic sequences reveal that polar bears are an old and distinct bear lineage. *Science* **336**, 344–347. (doi:10.1126/science.1216424)

Molecular analysis of 'anomalous primate' hair samples: a commentary on Sykes *et al.*

Norman MacLeod

Proc. R. Soc. B 2014 **281**, 20140843, published 2 July 2014

References

This article cites 2 articles, 1 of which can be accessed free

<http://rsjb.royalsocietypublishing.org/content/281/1789/20140843.full.html#ref-list-1>

Subject collections

Articles on similar topics can be found in the following collections

[genetics](#) (145 articles)

[molecular biology](#) (146 articles)

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)



CrossMark
click for updates

Commentary

Cite this article: MacLeod N. 2014 Molecular analysis of 'anomalous primate' hair samples: a commentary on Sykes *et al.* *Proc. R. Soc. B* **281**: 20140843.

<http://dx.doi.org/10.1098/rsob.2014.0843>

Received: 8 April 2014

Accepted: 6 June 2014

Author for correspondence:

Norman MacLeod

e-mail: n.macleod@nhm.ac.uk

Molecular analysis of 'anomalous primate' hair samples: a commentary on Sykes *et al.*

Norman MacLeod

The Natural History Museum, London, UK

While the correct logical formulation of a scientific hypothesis test is taught to virtually every child in their secondary school curriculum, the manner in which scientific researchers approach the resolution of questions concerning the cause(s) of natural phenomena is often misunderstood and/or misconstrued by the general public, usually aided or abetted by media reports that 'scientists have proved' this or that. With the exception of discoveries of species, minerals, compounds, etc., new to science, or known to science but found to occur at a place or time previously unanticipated, very little is 'proved' by science. Indeed, even in these cases all that is proved is that the phenomenon exists or existed at the place and time where it was found. The day-to-day work of most scientists lies not only with the discovery of new phenomena and/or occurrences, but also with the seemingly more mundane, though infinitely more complex, task of interpreting nature: how did the processes and objects we see in nature come to be? How do they function? What influences them and what do they influence? In statistics (which is often used as a tool for testing scientific hypotheses), the hypothetico-deductive formalism scientists most often used to explore these issues is enshrined in the concept of the null hypothesis which states that there is no relationship between two observed or measured phenomena [1]. Thus, in R. A. Fisher's classic 'lady tasting tea' experiment, the ability of the lady in question (Dr Muriel Bristol-Roach) to determine whether the milk was placed in the cup before or after the tea was evaluated by performing a series of randomized blind tests, recording the number of correct identifications and determining whether this number was sufficient to preclude the null hypotheses that they were obtained through random guessing. If the null hypothesis cannot be refuted no alternative hypotheses need be sought.

In the case of cryptozoology, its proponents have, for many years, claimed that the scientific establishment has failed to live up to the tenets of its own philosophy by failing to acknowledge the evidence they have offered for the existence of large species presently unknown to science. In most cases, scientific researchers have regarded this evidence—typically anecdotal observations recounted by individuals backed up on occasion by photographic and/or sound recordings, usually of quite poor quality—as hopelessly ambiguous and so not suitable for rigorous hypothesis testing. In such cases, the 'evidence' that links the observation with an unknown species (the alternative hypothesis) can be attributed reasonably to lack of familiarity of the observer with the regional biota, uncharacteristic behaviour of a known species, unusual lighting or fraudulent staging. In cases of ambiguous evidence such as these the null hypotheses of no link between such evidence and any unknown species is accepted because it cannot be refuted specifically. This stricture also applies to certain types of ephemeral physical evidence (e.g. trackways) that have often been documented photographically.

Of a different character altogether, however, is direct physical evidence in the form of bodies or body parts. These could, in principle, be compared with the body parts of known species and identified unambiguously as either having a combination of characteristics known to occur in a species described previously (the null hypothesis) or a set of characteristics of sufficient novelty to warrant establishment of a new species (the alternative hypothesis). Scientists who have looked into the claims of cryptozoologists have often been struck by the lack of such physical evidence in the form of collected individuals, dead bodies, fossils and/or parts thereof. This lack of direct and unambiguously testable evidence supporting the recognition of animals such as the

yeti, Loch Ness Monster, and Morag, not to mention sewer alligators, and the various beasts of Bodmin, Dartmoor, Exmoor and Dean, is the primary reason why many regard cryptozoology as a pseudoscience that accepts the existence of species in the absence of unambiguous physical evidence relying instead on personal observation, anecdote, legend and myth.

Until recently, hair samples reported to have been collected from areas where mammalian cryptozoological species are suspected to have occurred fell into this category of ambiguous evidence owing to the lack of morphological characteristics sufficient to rule out the possibility that they could be derived from extant species. However, owing the recognition that naturally occurring hair samples often include bits of skin and parts of hair follicles, the cells of which contain DNA, along with current DNA sequencing technologies, this physical evidence has moved out of the category of ambiguous, untestable evidence and into the realm of scientifically acceptable physical evidence that can be used, at least in principle, to identify unknown species. The reason for this alteration in the status of hair samples is that DNA sequences recovered therefrom could, in principle, be compared with those of extant species and the null hypothesis that the hair sample was derived from a species already known to science tested empirically.

The results of such tests on a series of 37 hair samples reported anecdotally to come from cryptozoological species is the subject of the Sykes *et al.* [2] article in this issue of the *Proceedings of the Royal Society*. These 37 samples were a subset of 58 samples submitted to the Sykes team for analysis. Of these 58 samples, two were excluded as being non-hair and 37 of the remaining 56 samples were selected for DNA analysis. The 19 samples excluded from DNA analysis were so designated for a variety of reasons including budget constraints, prioritization of samples of particular historical interest and amount of material available. In this reduced sample, seven of the samples selected for sequencing yielded no DNA. However, all of the 30 samples that did yield DNA contained base-pair sequences that were 100% compatible with known mammal species, though in certain instances the hair sample was reported to have been obtained from a region well outside the species' known geographical range. In two instances (samples 25 025 and 25 191), the gene sequence matched not an extant species, but a fossil sequence obtained from a Pleistocene polar bear (*Ursus maritimus*). As polar bears are not known to occur on the Tibetan Plateau, the Sykes team speculate that these samples may have come either from a previously unknown bear species

or possibly from a hybrid between *U. maritimus* and the brown bear (*Ursus arctos*). Viable hybrids of these species are known to occur in North America. A hybrid between two known bear species does not conform to the model offered by cryptozoologists to account for these samples, though if a hybrid bear species does occur in this region it may explain some of the anecdotal observations reported by individuals.

Does this evidence disprove the legends of the Yeti, Migyhur, Almasty, Sasquatch/Bigfoot? It does not. Scientific hypothesis testing of this sort is not designed to, and cannot, prove hypotheses alternative to the null hypothesis. All that can be said with confidence is that the results obtained by the Sykes team for the 29 samples that yielded DNA sequences failed to reject the null hypothesis that these samples came from species already known to science. Interestingly, despite the fact that most cryptozoologists have suggested the cryptids in question are unknown primate species, not one of the Sykes team's sequences yielded DNA that could be shown to have come from any non-human primate. Nevertheless, 19 of the original 55 bona fide hair samples submitted originally to the Sykes team did not produce DNA sequences. The taxonomic affinity of these samples remains unknown and science has nothing further to say about them, at least for the moment. From a scientific point of view, these samples return to the category of ambiguous evidence insofar as they cannot offer any unambiguous information that can be used to refute the null hypothesis of no link to any presently unknown (primate) species.

On a more general note, and as the Sykes *et al.* [2] report mentions in its last paragraph, this type of analysis opens the way for cryptozoologists and mainstream biological zoologists to enter into a productive dialogue. Cryptozoologists must now either accept the findings of the Sykes team or show where they are in error. Mainstream zoologists must also now recognize that, in the case of hair samples, the claims of the cryptozoological community are now amenable to scientific testing and potential verification. In this area, these two communities can and should speak the same language, the language of hard scientific data and hypothesis testing. Will this ultimately lead to the recognition of new large mammalian species in out-of-the-way corners of the world? No one—certainly no scientist—can say for sure. What we do know is that scientific discoveries just as strange and unexpected as those advocated by cryptozoologists in these cases have happened before (e.g. the coelacanth [3] and the okapi [4]).

References

1. Fisher RAS. 1935 *The design of experiments*. Edinburgh, UK: Oliver & Boyd.
2. Sykes BC, Mullis RA, Hagenmuller C, Melton TW, Sartori M. 2014 Genetic analysis of hair samples attributed to yeti, bigfoot and other anomalous primates. *Proc. R. Soc. B* **281**, 20140161. (doi:10.1098/rspb.2014.0161)
3. Forey PL. 1998 *History of the coelacanth fishes*. London, UK: Chapman & Hall.
4. Simpson GG. 1984 Mammals and cryptozoology. *Proc. Am. Phil. Soc.* **128**, 1–19.

Mammals on the EDGE: Conservation Priorities Based on Threat and Phylogeny

Nick J. B. Isaac*, Samuel T. Turvey, Ben Collen, Carly Waterman, Jonathan E. M. Baillie

Institute of Zoology, Zoological Society of London, London, United Kingdom

Conservation priority setting based on phylogenetic diversity has frequently been proposed but rarely implemented. Here, we define a simple index that measures the contribution made by different species to phylogenetic diversity and show how the index might contribute towards species-based conservation priorities. We describe procedures to control for missing species, incomplete phylogenetic resolution and uncertainty in node ages that make it possible to apply the method in poorly known clades. We also show that the index is independent of clade size in phylogenies of more than 100 species, indicating that scores from unrelated taxonomic groups are likely to be comparable. Similar scores are returned under two different species concepts, suggesting that the index is robust to taxonomic changes. The approach is applied to a near-complete species-level phylogeny of the Mammalia to generate a global priority list incorporating both phylogenetic diversity and extinction risk. The 100 highest-ranking species represent a high proportion of total mammalian diversity and include many species not usually recognised as conservation priorities. Many species that are both evolutionarily distinct and globally endangered (EDGE species) do not benefit from existing conservation projects or protected areas. The results suggest that global conservation priorities may have to be reassessed in order to prevent a disproportionately large amount of mammalian evolutionary history becoming extinct in the near future.

Citation: Isaac NJB, Turvey ST, Collen B, Waterman C, Baillie JEM (2007) Mammals on the EDGE: Conservation Priorities Based on Threat and Phylogeny. *PLoS ONE* 2(3): e296. doi:10.1371/journal.pone.0000296

INTRODUCTION

Our planet is currently experiencing a severe anthropogenically driven extinction event, comparable in magnitude to prehistoric mass extinctions. Global extinction rates are now elevated up to a thousand times higher than the background extinction rates shown by the fossil record, and may climb another order of magnitude in the near future [1–3]. The resources currently available for conservation are, unfortunately, insufficient to prevent the loss of much of the world's threatened biodiversity during this crisis, and conservation planners have been forced into the unenviable situation of having to prioritise which species should receive the most protection—this is 'the agony of choice' [4] or the 'Noah's Ark problem' [5].

A range of methods for setting species-based conservation priorities have been advocated by different researchers or organisations, focusing variously on threatened species, restricted-range endemics, 'flagship', 'umbrella', 'keystone', 'landscape' or 'indicator' species, or species with significant economic, ecological, scientific or cultural value [6–8]. To date, global priority-setting exercises have tended to focus on endemic (or restricted range) species [6,9,10], presumably because endemism is easier to measure than competing methods. However, recent data show that endemism is a poor predictor of total species richness or the number of threatened species [11].

It has also been argued that maximising Phylogenetic Diversity (PD) should be a key component of conservation priority setting [4,12–14]. Species represent different amounts of evolutionary history, reflecting the tempo and mode of divergence across the Tree of Life. The extinction of a species in an old, monotypic or species-poor clade would therefore result in a greater loss of biodiversity than that of a young species with many close relatives [15,16]. However, conserving such lineages may be difficult, since there is some evidence that they are more likely to be threatened with extinction than expected by chance [17]. This clumping of extinction risk in species-poor clades greatly increases the loss of PD compared with a null model of random extinction [18] and suggests that entire vertebrate orders may be lost within centuries

[19]. Among mammals alone, at least 14 genera and three families have gone extinct since AD 1500 [20], and all members of a further 19 families and three orders are considered to be in imminent danger of extinction [2]. Many academic papers have suggested ways to maximise the conservation of PD [e.g. 12,13,21–23] and measure species' contributions to PD [e.g. 4,23–25], but these have rarely been incorporated into conservation strategies. Therefore, it is possible that evolutionary history is being rapidly lost, yet the most distinct species are not being identified as high priorities in existing conservation frameworks.

There are several reasons why PD has not gained wider acceptance in the conservation community. First, although evolutionary history consists of two distinct components (the branching pattern of a phylogenetic tree and the length of its branches), complete dated species-level phylogenies for large taxonomic groups have only recently become available [26]. Early implementations of PD-based approaches were therefore unable to incorporate branch length data, and focused solely on measurements of branching pattern [4]. Second, PD removes the focus from species and so may lack wider tangible appeal to the public; conserving PD may be seen as less important than the protection of endemic or

.....
Academic Editor: Walt Reid, The David and Lucile Packard Foundation, Conservation and Science Program, United States of America

Received January 16, 2007; **Accepted** February 19, 2007; **Published** March 14, 2007

Copyright: © 2007 Isaac et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors did not receive any funding to conduct the research described in this paper.

Competing Interests: The authors have declared that no competing interests exist.

* **To whom correspondence should be addressed.** E-mail: nick.isaac@ioz.ac.uk

threatened species [16]. However, the current instability in species taxonomy [27] means that decisions based on PD might be more objective than those based on different species concepts [13,16,27]. Combining species' conservation status with a measure of their contribution to PD is therefore desirable, because species can be retained as units but weighted appropriately [5,22]. This would generate a useful and transparent means for setting global priorities for species-based conservation [25].

This paper describes a new method for measuring species' relative contributions to phylogenetic diversity [the 'originality' of species: ref 24]. We explore the statistical properties of the resulting measure, which we call Evolutionary Distinctiveness (ED), and test its robustness to changing species concepts. ED scores are calculated for the Class Mammalia, and combined with values for species' extinction risk to generate a list of species that are both evolutionarily distinct and globally endangered ('EDGE species'). The resultant list provides a set of priorities for mammalian conservation based not only on the likelihood that a species will be lost, but also on its irreplaceability.

Evolutionary Distinctiveness and its use in priority-setting

In order to calculate ED scores for each species, we divide the total phylogenetic diversity of a clade amongst its members. This is achieved by applying a value to each branch equal to its length divided by the number of species subtending the branch. The ED of a species is simply the sum of these values for all branches from which the species is descended, to the root of the phylogeny. For the examples in this paper, we have measured ED in units of time, such that each million years of evolution receives equal weighting and the branches terminate at the same point (i.e. the phylogeny is ultrametric). The method could be applied to non-ultrametric phylogenies if the conservation of other units [e.g. character diversity 28,29] was prioritised [although see ref 30].

The basic procedure for calculating ED scores is illustrated in figure 1, which describes a clade of seven species (A–G). The ED score of species A is given by the sum of the ED scores for each of the four branches between A and the root. The terminal branch contains just one species (A) and is 1 million years (MY) long, so receives a score of 1 MY. The next two branches are both 1 MY long and contain two and three species, so each daughter species (A, B and C) receives 1/2 and 1/3 MY respectively. The deepest branch that is ancestral to species A is 2 MY long and is shared among five species (A to E), so the total ED score for species A is given by $(1/1+1/2+1/3+2/5) = 2.23$ MY. Species B is the sister taxon of A, so receives the same score. By the same arithmetic, C has a score of $(2/1+1/3+2/5) = 2.73$ MY, both D and E receive $(1/1+2/2+2/5) = 2.4$ MY, and both F and G receive $(0.5/1+4.5/2) = 2.75$ MY. The example illustrates that ED is not solely determined by a species' unique PD (i.e. the length of the terminal branch). Species F and G are the top-ranked species based on their ED scores, even though each represents just a small amount of unique evolutionary history (0.5 MY). This suggests that the conservation of both F and G should be prioritised, because the extinction of either would leave a single descendant of the oldest and most unusual lineage in the phylogeny [c.f. 15,24]. The ED calculation is similar to the Equal Splits measure [25], which apportions branch length equally among daughter clades, rather than among descendent species.

In order to represent a useful tool in priority setting, ED scores must be applicable in real phylogenies of large taxonomic groups. To do this, we modified the basic procedure described above to control for missing species, incomplete phylogenetic resolution and

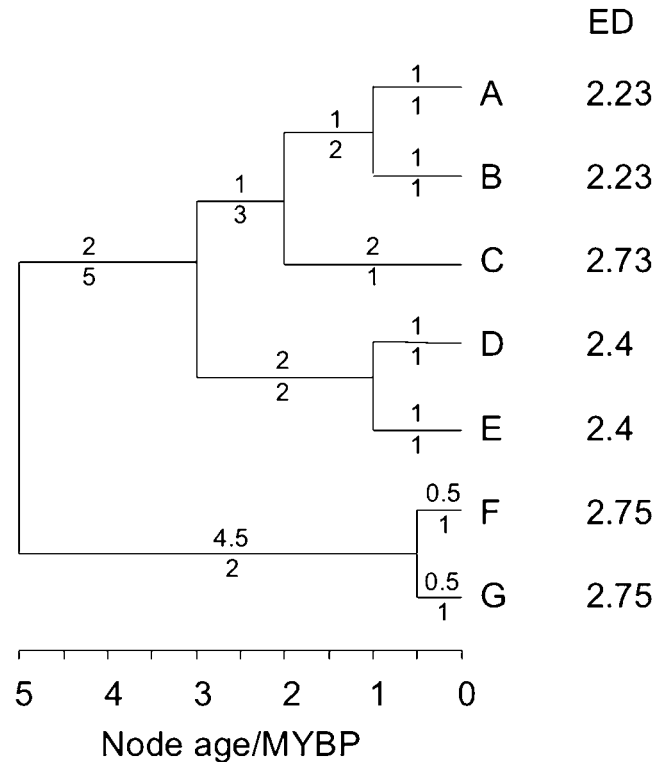


Figure 1. Hypothetical phylogeny of seven species (A–G) with Evolutionary Distinctiveness (ED) scores. Numbers above each branch indicate the length; numbers below show the number of descendent species. MYBP, millions of years before present. doi:10.1371/journal.pone.0000296.g001

uncertainty in node ages (see Materials and Methods). The approach is implemented using a dated phylogeny of the Class Mammalia that is nearly complete (>99%) at the species level [31]. We then combined ED and extinction risk to identify species that are both evolutionarily distinct and globally endangered ('EDGE species'). We measured extinction risk using the quantitative and objective framework provided by the World Conservation Union (IUCN) Red List Categories [2]. We follow previous researchers in treating the Red List categories as intervals of extinction risk and by assuming equivalence among criteria [32,33, but see 34]. The resulting list of conservation priorities ('EDGE scores') was calculated as follows:

$$EDGE = \ln(1 + ED) + GE * \ln(2) \quad (1)$$

where GE is the Red List category weight [Least Concern = 0, Near Threatened and Conservation Dependent = 1, Vulnerable = 2, Endangered = 3, Critically Endangered = 4, ref 32], here representing extinction risk on a log scale. EDGE scores are therefore equivalent to a \log_e -transformation of the species-specific expected loss of evolutionary history [5,25] in which each increment of Red List category represents a doubling ($e^{\ln(2)}$) of extinction risk. For the purposes of these analyses, we did not calculate EDGE scores for species listed as Extinct in the Wild ($n=4$), domesticated populations of threatened species and 34 species (mostly of dubious taxonomic status) for which an evaluation has not been made.

RESULTS

Statistical properties of ED

We measured ED in clades of different sizes to test whether ED scores from different taxonomic groups are likely to be comparable. We found that most ED is derived from a few branches near the tips (i.e. those shared with few other species) and that virtually no ED is gained in clades above ~ 180 species (figure 2). Median ED in clades of 60 species is 88% of the total accumulated using the whole tree ($n = 10$, figure 2). Moreover, the rank order of ED scores is unaffected by the size of the clade under consideration, except in very small clades and among species with low overall ED (i.e. few of the lines in figure 2 cross one another). These findings suggest that ED scores of different taxonomic groups measured on separate phylogenies (i.e. with no nodes in common) will be comparable, so long as each phylogeny is larger than a threshold size. Based on the scaling observed in figure 2, we suggest a minimum species richness of 100 as a useful rule of thumb to ensure comparability among taxa.

Although most species (90% in figure 2) derive at least two-thirds of their total ED from the terminal branch (which is not shared with others), this branch length is a poor predictor of total ED ($r^2 = 0.03$ on a log-log scale). For species on short branches, there is an order of magnitude difference between the length of the terminal branch and ED. For example, the pale-throated and brown-throated three-toed sloths (*Bradypus tridactylus* and *B. variegatus*) share a common ancestor thought to be just over a million years old, but the total ED of both species is 20.4 MY (Table S1) since they have few close living relatives.

ED scores are also robust to taxonomic changes. For example, ED scores in primates under the biological species concept [35] are tightly correlated with ED scores under the phylogenetic

species concept [36] ($r^2 = 0.65$ on a log-log scale), in spite of the fact that there are substantial differences between the two: the number of primate species differs by 50%. Furthermore, the highest-ranking species do not change their identity: 45 of 58 biological species in the upper quartile of ED scores are also in the upper quartile as phylogenetic species. However, species that have been split into three or more species do tend to lose a large portion of their ED. For example, the fork-marked lemur (*Phaner furcifer*) is the second most distinct biological species of primate, with an ED score of 38.33. It was split into four phylogenetic species [36] with an ED score of 10.45 (Table S2), which is just inside the upper quartile.

ED and EDGE scores in mammals

Mammal ED scores range from 0.0582 MY (19 murid rodents) to 97.6 MY (duck-billed platypus, *Ornithorhynchus anatinus*). Scores are approximately log-normally distributed, with a median of 7.86 MY and geometric mean of 6.28 MY.

Evolutionary Distinctiveness is not evenly distributed among the Red List categories. Least Concern species have significantly lower ED than the other categories ($F_{1,4180} = 26.3$, $p < 0.0001$, using \log_e transformed scores); there are no significant differences among the remaining categories. This suggests that species with low ED scores tend to suffer from low levels of extinction risk, although the explanatory power of this model is extremely low ($r^2 = 0.006$).

EDGE scores range from 0.0565 (10 murid rodents) to 6.48 (Yangtze River dolphin or baiji, *Lipotes vexillifer*) and are approximately normally distributed around a mean of 2.63 (± 0.017 ; figure 3). The 100 highest priority (EDGE) species includes several large-bodied and charismatic mammals, including the giant and lesser pandas, the orang-utan, African and Asian

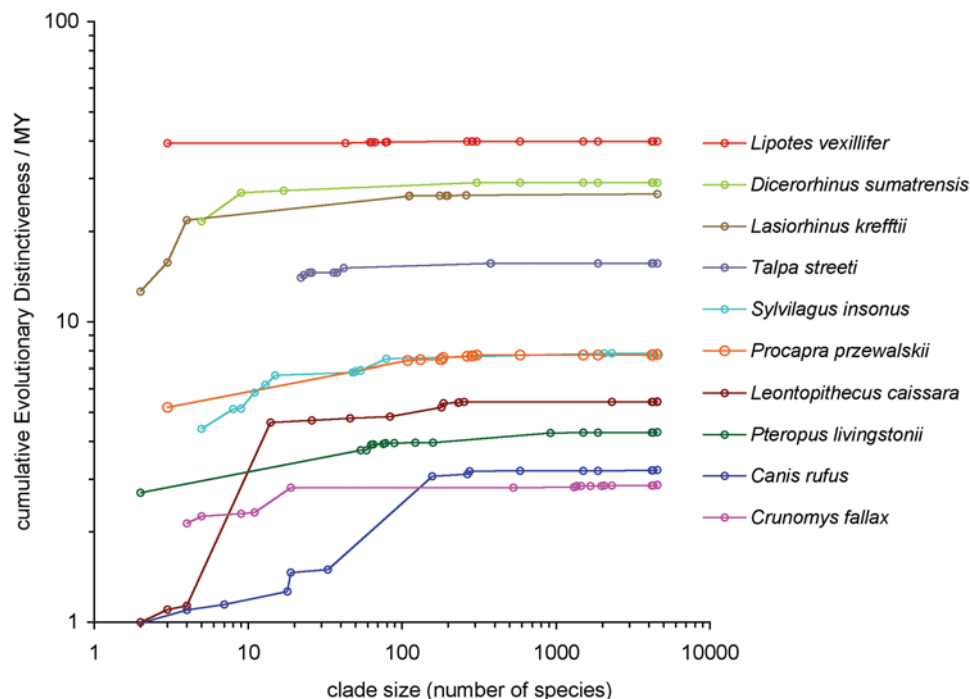


Figure 2. Scaling of ED scores with clade size for ten Critically Endangered mammal species. ED scores were calculated at each node between the tips and root for ten species in different orders. Species chosen are: the baiji (*Lipotes vexillifer*), sumatran rhino (*Dicerorhinus sumatrensis*), northern hairy-nosed wombat (*Lasiorhinus krefftii*), persian mole (*Talpa streeti*), Omiltemi rabbit (*Sylvilagus insonus*), Przewalski's gazelle (*Procapra przewalskii*), black-faced lion tamarin (*Leontopithecus caissara*), Livingstone's flying fox (*Pteropus livingstonii*), red wolf (*Canis rufus*) and northern Luzon shrew rat (*Crunomys fallax*). See Materials and Methods for further details. doi:10.1371/journal.pone.0000296.g002

elephants, four rhinoceroses, two tapirs, two baleen whales, a dugong and a manatee. However, many smaller and less appreciated species also receive high priority, including sixteen rodents, thirteen eulipotyphlans, twelve bats, four lagomorphs and an elephant shrew (Table S1). The top 100 also includes at least 37 species that would not qualify for most area-based definitions of endemism, since they are listed as threatened under Red List criterion A (reduction in population size) without qualifying for criteria B–D, which are based on population size or geographical range. Whilst the highest-ranked species, by definition, are all highly threatened (44 of the top 100 species are Critically Endangered, a further 47 are Endangered), threat status alone does not guarantee a high priority. For example, 10 Critically Endangered species (in the genera *Gerbillus*, *Peromyscus* and *Crocidura*), as well as 32 Endangered species, fail to make the top 1000, whilst 130 Near Threatened species do.

DISCUSSION

It is important that conservation priority-setting approaches are able to satisfy two conditions: they capture biodiversity and are robust to uncertainty. The method described herein satisfies the first condition because EDGE scores incorporate species value (in terms of originality, or irreplaceability) weighted by urgency of action (i.e. risk of extinction). Our approach satisfies the second condition because the scores are also robust to clade size, missing species and poor phylogenetic resolution. EDGE scores are also easy to calculate, as all that is required is a set of Red List assessments and a near-complete phylogeny containing at least 100 species.

In particular, EDGE priorities are much less sensitive to taxonomic uncertainty than alternate methods. The current trend towards the adoption of the phylogenetic species concept among biologists [27] is likely to produce a large number of ‘new’ threatened and endemic species [37], potentially altering the distribution of hotspots [38] and distorting other biodiversity patterns [27]. The EDGE approach is robust to such distortion because any increase in extinction risk due to splitting is balanced

by a decrease in ED. A good example is that of the ruffed lemurs (*Varecia* spp.), which consist of one Endangered biological species (ED = 19.8; EDGE = 5.11) or two phylogenetic species (Endangered and Critically Endangered; ED = 10.3; EDGE = 4.50 and 5.20). Using the same approach, we estimate that the long-beaked echidna (*Zaglossus bruijnii*) would fall from the second-ranked priority to the 20th after the addition of two new congeners [suggested by 39]. Thus, EDGE scores for existing species are robust to the ongoing discovery of new species.

EDGE priorities are also robust to several other forms of uncertainty. Like all phylogenetic methods, the precise EDGE scores are dependent on the topology and branch lengths of the phylogeny. However, errors in the phylogeny are unlikely to alter the identity of high-ranking species, particularly for clades of several hundred species. Topological uncertainty is usually expressed in supertrees as polytomies, which are accounted for using simple correction factors. Likewise, branch length uncertainty has been incorporated into the scoring system to down-weight the priority of species descended from nodes with imprecisely estimated ages (see Materials and Methods). These developments make it possible to estimate robustly the contribution to phylogenetic diversity of species in poorly known clades. The other major source of uncertainty is in estimating extinction risk: most recent changes in Red List category have come about through improvements in knowledge, rather than genuine changes in status [32]. EDGE scores will inevitably be affected by future changes in extinction risk, although no more so than other approaches using the Red List categories.

A minority of mammal species could not be assigned EDGE scores. Around 300 species are classified as Data Deficient and could not be meaningfully included, although in reality they may have a high risk of extinction [17]. By far the most likely candidate for high EDGE status following future Red List re-assessment is the franciscana or La Plata River dolphin *Pontoporia blainvillei* (ED = 36.3 MY). In addition, fifty extant species are missing from the phylogeny. The highest ranked of these are probably a pair of Critically Endangered shrews (*Sorex cansulus* and *S. kizilovi*); median and maximum ED scores for the genus are 4.55 and 14.6 MY, giving potential respective EDGE scores of 4.49 and 5.52 for these species (cf. figure 3). A further 260 species have been described since the chosen taxonomy was published [40]. Of these, the recently described Annamite striped rabbit *Nesolagus timminsi* [41] is the sister species to the tenth-ranked Sumatran rabbit *N. netscheri*, so would be a high priority if similarly threatened.

It has been suggested that species with few close relatives (i.e. high ED) are ‘relicts’ or ‘living fossils’ that have limited ability to generate novel diversity. This view implies that conservation efforts should instead be focused on recent radiations containing species with low ED scores (e.g. murid rodents), which represent ‘cradles’ rather than ‘museums’ of diversity [e.g. 16,42]. However, the assumption that we are able to predict future evolutionary potential is dubious and no general relationships between phylogeny and diversity over geological time have yet been established [43,44]. Furthermore, phylogenetic diversity is clearly related to character diversity [30], and so ED may be a useful predictor of divergent properties and hence potential utilitarian value [14]. Moreover, because species with low ED scores tend to suffer from low levels of extinction risk, phylogenetic cradles of mammalian diversity are likely to survive the current extinction crisis even without specific interventions. Focusing on lower risk species, at the expense of EDGE priorities, would therefore result in a severe pruning of major branches of the Tree of Life comparable to that seen in previous mass extinction events [45,46].

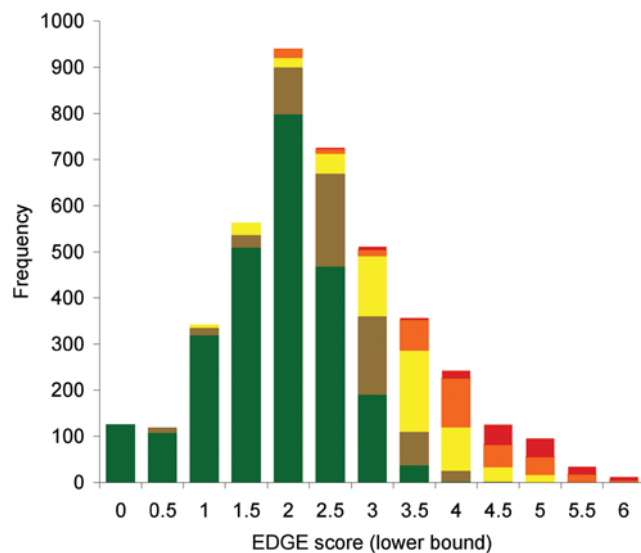


Figure 3. Histogram of EDGE scores for 4182 mammal species, by threat category. Colours indicate the Red List category: Least Concern (green), Near Threatened and Conservation Dependent (brown), Vulnerable (yellow), Endangered (orange) and Critically Endangered (red). doi:10.1371/journal.pone.0000296.g003

The top 100 EDGE species span all the major mammalian clades [being distributed among 18 orders and 52 families recognised by ref 35] and display a comparable range of morphological and ecological disparity, including the largest and smallest mammals, most of the world's freshwater cetaceans, an oviparous mammal and the only species capable of injecting venom using their teeth. However, around three-quarters of species-based mammal conservation projects are specifically aimed at charismatic megafauna [47], so conventional priority-setting tools may not be sufficient to protect high priority EDGE species. This concern is supported by two additional lines of evidence. First, we found that species not found in protected areas ['gap species' defined by ref 48] tended to have higher EDGE scores than those found inside protected areas (logistic regression: $\chi^2_{1,3994} = 69.46$, $p < 0.0001$). Second, an assessment of published conservation strategies and recommendations (including IUCN Specialist Group Conservation Action Plans, captive breeding protocols and the wider scientific literature listed in the 1978–2005 Zoological Record database) reveals that no species-specific conservation actions have even been suggested for 42 of the top 100 EDGE species. Most of these species are from poorly known regions or taxonomic groups and until now have rarely been highlighted as conservation priorities. Little conservation action is actually being implemented for many other top EDGE species, despite frequent recommendations in the conservation literature. Indeed, the top-scoring EDGE species, the Yangtze River dolphin (*Lipotes vexillifer*), is now possibly the world's most threatened mammal despite two decades of debate over a potential *ex situ* breeding programme, and may number fewer than 13 surviving individuals [49]. The lack of conservation attention for priority EDGE species is a serious problem for mammalian biodiversity and suggests that large amounts of evolutionary history are likely to be lost in the near future. This phenomenon of diversity slipping quietly towards extinction is likely to be much more severe in less charismatic groups than mammals.

The approach described in this paper can be used for conservation in a number of ways. First, conservation managers with limited resources at their disposal typically need to conserve populations of several threatened species. If all other factors were equal, the management of the most evolutionarily distinct species should be prioritized. Second, a list of high-priority species requiring urgent conservation action can be generated easily. In this paper, we have selected the 100 highest-ranking species, but one might equally choose all threatened (Vulnerable and above) species with above average ED. This would result in a list of 521 (using median) or 630 (using geometric mean) 'EDGE species' that are both evolutionarily distinct and globally endangered. Third, EDGE scores could also be used to weight species' importance in selecting reserve networks, building on previous studies that have used phylogenetic diversity [50–52] or threatened species [11] to identify priority areas for conservation. The statistical properties of EDGE scores (they are both normally-distributed and bounded at zero) make them especially suitable for these kinds of analysis. In this way, the EDGE approach is not an alternative to existing conservation frameworks [e.g. 6] but complements them.

The EDGE approach identifies the species representing most evolutionary history from among those in imminent danger of extinction. Our methods extend the application of PD-based conservation to a wider range of taxa and situations than previous approaches [4,5,13,22,24,25]. Future work might incorporate socioeconomic considerations [5,14] and the fact that a species' value depends also on the extinction risk of its close relatives [53]. However, our results suggest that large numbers of evolutionarily distinct species are inadequately served by existing conservation

measures, and that more work is carried out to prevent the imminent loss of large quantities of our evolutionary heritage. It is hoped that this approach will serve to highlight their importance to biodiversity and emphasize the need for urgent conservation action.

MATERIALS AND METHODS

Implementing ED scores for mammals

We used a composite 'supertree' phylogeny [31] to calculate ED scores for mammals. The supertree presents several challenges to the estimation of ED when compared with the (unknown) true phylogeny: poor resolution, missing species and uncertainty in node ages. Accordingly, we modified the basic procedure to control for these problems.

Phylogenetic information is poor in many mammalian clades (especially bats and rodents, which together make up >60% of species) and the whole supertree contains only 47% of all possible nodes, many of which are polytomies (nodes with more than two daughter branches). Across the whole phylogeny, ~40% of species are immediately descended from bifurcations, ~20% from small polytomies (3–5 daughters), ~15% from medium-sized polytomies (6–10 daughters) and the remainder from large polytomies with >10 daughters. Polytomies in supertrees result from poor or conflicting data rather than a true representation of the speciation process, so the distinctiveness of branches subtending them is overestimated [54], thus leading to biased ED scores. For example, the common ancestor of species X, Y and Z is believed to be 1 MY old, but the branching pattern within the clade is unknown. The polytomy appears to show that each species represents 1 MY of unique evolutionary history. In reality, the phylogeny is bifurcating, with one species aged 1 MY and the others sharing a more recent common ancestor. The bias induced by polytomies can be corrected by estimating the expected ED of descendant species under an appropriate null model of diversification. We achieved this by applying a scaling factor based on the empirical distribution of ED scores in a randomly generated phylogeny of 5000 species grown under constant rates of speciation (0.1) and extinction (0.08). The mean ED score of species in 819 clades of three species was 0.81 of the clade age; ED scores for nodes of 2–20 species scale according to $(\text{branch length}) * (1.081 - 0.267 * \ln\{d\})$, where d is the number of descendent branches ($n = 2873$ clades, $r^2 = 0.69$). Quantitatively similar values were obtained in bifurcating clades of primates [$1.117 - 0.246 * \ln\{d\}$, $n = 78$, ref 55] and carnivores [$1.139 - 0.269 * \ln\{d\}$, $n = 101$, ref 56].

The mammal supertree contains 4510 of the 4548 (>99%) extant species listed in Wilson & Reeder [35]. Although few in number, the missing species need to be taken into account because their absence will tend to inflate the ED scores of close relatives. For example, omitting species A from the phylogeny in figure 1 would elevate B from the joint lowest ranking species (with A) to the joint highest-ranking (with C), with an ED score of $(2/1 + 1/2 + 2/4) = 3.5$ MY. The problem is acute in real datasets since missing species tend not to be a random sample: 22 of the 38 missing mammals are from the genus *Sorex*. We account for this problem using a simple correction factor that allocates the missing species among their presumed closest relatives. For example, we correct for the omission of the bare-bellied hedgehog (*Hemiechinus nudiventris*) by treating the other five *Hemiechinus* spp. as $6/5 = 1.2$ species, and we correct for the omission of both *Cryptochloris* species by spreading the two missing species evenly between other Chrysochloridae.

Variation among morphological and molecular estimates of divergence times (node ages) can lead to considerable uncertainty

in ED scores. To reduce the effects of this uncertainty, we estimated ED using three sets of branch lengths. One set was based on the best (i.e. mean) estimates of node age; the others were derived from the upper and lower 95% confidence intervals around these dates. Species values of ED were calculated as the geometric mean of scores under the three sets of branch lengths. The geometric mean was preferred since it down-weights species whose scores are based on nodes with symmetrical but wide confidence intervals in estimate age, and is therefore more conservative than the arithmetic mean.

Tests of robustness

To test whether ED scores are comparable among taxonomic groups, we examined how species' ED accumulates as progressively larger clades are considered. If ED scores are truly comparable, their rank order will be independent of the size of the clade considered. We randomly selected one Critically Endangered species from each of ten mammal orders and measured the cumulative ED score at each node between the species and the root of the mammal supertree, thus redefining and enlarging the clade (and so increasing the number of species it contained) at each step.

Taxonomic changes have the potential to dramatically alter the ED scores of individual species. Splitting a species in two reduces the distinctiveness of all branches ancestral to the split, particularly those near the tips. If ED scores are highly sensitive to taxonomic changes then it may be meaningless to apply them in setting conservation priorities. The effects of taxonomic changes on ED scores were therefore investigated in the primates, which have recently experienced considerable taxonomic inflation [27]. We compared primate ED scores under a biological species concept [35: 233 species] and a phylogenetic species concept [36: 358 species]. We employed a single phylogeny [31], but changed the number of species represented by each tip. We calculated the expected ED for multi-species tips by treating them as if they were descended from a polytomy of $\{n+r+1\}$ descendent branches, where n is the actual number of descendent branches and r is the number of species represented by the tip.

SUPPORTING INFORMATION

Table S1 Evolutionary Distinctiveness and EDGE scores for mammals. This table shows Evolutionary Distinctiveness (ED) and

EDGE scores for all species included in the mammal supertree [31] ranked by their EDGE score. Species that could not be assigned EDGE scores are appended to the bottom of the list, sorted by status and ED score. Species taxonomy follows Wilson & Reeder [35]. Red List categories follow the 2006 IUCN Red List [2]: CR = Critically Endangered, EN = Endangered, VU = Vulnerable, NT = Near Threatened, LC = Least Concern, CD = Conservation Dependent, DD = Data Deficient, NE = Not Evaluated. The NE category includes species in Wilson & Reeder [35] that could not be matched with any species or subspecies names in the Red List.

Found at: doi:10.1371/journal.pone.0000296.s001 (0.42 MB PDF)

Table S2 Evolutionary Distinctiveness for primates under two species concepts. This table lists ED scores for primates under the biological species concept [i.e. the taxonomy of ref 35], the number of phylogenetic species into which the biological species was split [36] and the estimated ED score of each phylogenetic species. See Materials and Methods for further information. ED scores are lower for phylogenetic species than biological species, even for taxa whose taxonomic status is the same under both concepts (i.e. the number of phylogenetic species is one). This occurs because the total number of species in the phylogeny is greater, so each receives a smaller share of the distinctiveness of ancestral branches. ED scores were calculated using just one set of branch lengths (the 'best' set), so differ from those in table S1.

Found at: doi:10.1371/journal.pone.0000296.s002 (0.05 MB PDF)

ACKNOWLEDGMENTS

We thank the PanTHERIA consortium for use of the mammal supertree. In particular, Olaf Bininda-Emonds generously provided us with electronic copies of the topology and alternate branch lengths. We also thank Ana Rodrigues for providing a list of species found in protected areas. We are also grateful to Guy Cowlshaw, Sarah Durant, Dan Faith, John Gittleman, Rich Grenyer, Kate Jones, Georgina Mace, Arne Mooers, Andy Purvis and three anonymous referees for useful comments and discussion.

Author Contributions

Conceived and designed the experiments: JB BC NI ST. Performed the experiments: NI. Analyzed the data: NI. Contributed reagents/materials/analysis tools: NI ST CW. Wrote the paper: JB BC NI ST.

REFERENCES

- Pimm SL, Russell GJ, Gittleman JL, Brooks TM (1995) The future of biodiversity. *Science* 269: 347–350.
- IUCN (2006) 2006 IUCN Red List of Threatened Species. <http://www.iucnredlist.org>.
- Mace GM, Masundire H, Baillie JEM (2005) Biodiversity. Chapter 4 in: Millennium ecosystem assessment, 2005 Current state and trends: findings of the condition and trends working group Ecosystems and human well-being, vol1. Washington DC: Island Press.
- Vane-Wright RI, Humphries CJ, Williams PH (1991) What to protect—systematics and the agony of choice. *Biological Conservation* 55: 235–254.
- Weitzman ML (1998) The Noah's Ark Problem. *Econometrica* 66: 1279–1298.
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.
- Mace GM, Collar NJ (2002) Priority-setting in species conservation. In: Norris K, Pain DJ, eds (2002) Conserving bird biodiversity: general principles and their application. Cambridge: Cambridge University Press. pp. 61–73.
- Entwistle A, Dunstone N, eds (2000) Has the panda had its day? Priorities for the conservation of mammalian diversity. Cambridge: Cambridge University Press. 455 p.
- Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, et al. (2001) Terrestrial ecoregions of the worlds: A new map of life on Earth. *Bioscience* 51: 933–938.
- Stattersfield AJ, Crosby MJ, Long AJ, Wege DC (1998) Endemic bird areas of the world: priorities for biodiversity conservation. Cambridge: BirdLife International. 846 p.
- Orme CDL, Davies RG, Burgess M, Eigenbrod F, Pickup N, et al. (2005) Global hotspots of species richness are not congruent with endemism or threat. *Nature* 436: 1016–1019.
- Witting L, Loeschcke V (1995) The optimization of biodiversity conservation. *Biological Conservation* 71: 205–207.
- Faith DP (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation* 61: 1–10.
- Crozier RH (1997) Preserving the information content of species: Genetic diversity, phylogeny, and conservation worth. *Annual Review of Ecology and Systematics* 28: 243–268.
- May RM (1990) Taxonomy as destiny. *Nature* 347: 129–130.
- Mace GM, Gittleman JL, Purvis A (2003) Preserving the tree of life. *Science* 300: 1707–1709.
- Purvis A, Agapow PM, Gittleman JL, Mace GM (2000) Nonrandom extinction and the loss of evolutionary history. *Science* 288: 328–330.
- Heard SB, Mooers AO (2000) Measuring the loss of evolutionary history from extinction: phylogenetically patterned speciation rates and extinction risks alter the calculus of biodiversity. *Proceedings of the Royal Society of London Series B-Biological Sciences* 267: 613–620.

19. McKinney ML (1998) Branching models predict loss of many bird and mammal orders within centuries. *Animal Conservation* 1: 159–164.
20. MacPhee RDE, Flemming C (1999) Requiem æternam: the last five hundred years of mammalian species extinctions. In: MacPhee RDE, ed (1999) *Extinctions in near time: causes, contexts, and consequences*. New York: Kluwer Academic/Plenum. pp. 333–371.
21. Nee S, May RM (1997) Extinction and the loss of evolutionary history. *Science* 278: 692–694.
22. Barker GM (2002) Phylogenetic diversity: a quantitative framework for measurement of priority and achievement in biodiversity conservation. *Biological Journal of the Linnean Society* 76: 165–194.
23. Steel M (2005) Phylogenetic diversity and the greedy algorithm. *Systematic Biology* 54: 527–529.
24. Pavoiné S, Ollier S, Dufour AB (2005) Is the originality of a species measurable? *Ecology Letters* 8: 579–586.
25. Redding DW, Mooers AO (2006) Incorporating evolutionary measures into conservation prioritization. *Conservation Biology* 20: 1670–1678.
26. Bininda-Emonds ORP (2004) The evolution of supertrees. *Trends in Ecology and Evolution* 19: 315–322.
27. Isaac NJB, Mallet J, Mace GM (2004) Taxonomic inflation: its influence on macroecology and conservation. *Trends in Ecology and Evolution* 19: 464–469.
28. Owens IPF, Bennett PM (2000) Quantifying biodiversity: a phenotypic perspective. *Conservation Biology* 14: 1014–1022.
29. Diniz-Filho JAF (2004) Phylogenetic diversity and conservation priorities under distinct models of phenotypic evolution. *Conservation Biology* 18: 698–704.
30. Faith DP (2002) Quantifying biodiversity: a phylogenetic perspective. *Conservation Biology* 16: 248–252.
31. Bininda-Emonds ORP, Cardillo M, Jones KE, MacPhee RDE, Beck RMD, et al. (2007) The delayed rise of present-day mammals. *Nature*, in press. doi:10.1038/nature05634
32. Butchart SHM, Stattersfield AJ, Bennun LA, Shutes SM, Akçakaya HR, et al. (2004) Measuring global trends in the status of biodiversity: Red List Indices for birds. *PLoS Biology* 2: 2294–2304.
33. Purvis A, Cardillo M, Grenyer R, Collen B (2005) Correlates of extinction risk: phylogeny, biology, threat and scale. In: Purvis A, Gittleman JL, Brooks T, eds (2005) *Phylogeny and conservation*. Cambridge: Cambridge University Press. pp. 295–316.
34. IUCN Red List Standards and Petitions Subcommittee (2006) Guidelines for using the IUCN Red List categories and criteria. <http://www.iucn.org/webfiles/doc/SSC/RedList/RedListGuidelines.pdf>.
35. Wilson DE, Reeder DM, eds (1993) *Mammal species of the world: a taxonomic and geographic reference*. 2nd ed. Washington DC: Smithsonian Institution Press.
36. Groves CP (2001) *Primate Taxonomy*; Erwin DH, Funk VA, eds. Washington and London: Smithsonian Institution Press.
37. Agapow PM, Bininda-Emonds ORP, Crandall KA, Gittleman JL, Mace GM, et al. (2004) The impact of species concept on biodiversity. *Quarterly Review of Biology* 79: 161–179.
38. Peterson AT, Navarro-Siguenza AG (1999) Alternate species concepts as bases for determining priority conservation areas. *Conservation Biology* 13: 427–431.
39. Flannery TF, Groves CP (1998) A revision of the genus *Zaglossus* (Monotremata, Tachyglossidae), with description of new species and subspecies. *Mammalia* 62: 367–396.
40. Wilson DE, Reeder DM, eds (2005) *Mammal Species of the World*. 3rd ed. Baltimore: Johns Hopkins University Press. 2142 p.
41. Averianov AO, Abramov AV, Tikhonov AN (2000) A new species of *Nesolagus* (Lagomorpha, Leporidae) from Vietnam with osteological description. *Contributions from the Zoological Institute, St Petersburg* 3: 1–22.
42. Erwin TL (1991) An evolutionary basis for conservation strategies. *Science* 253: 750–752.
43. Krajewski C (1991) Phylogeny and diversity. *Science* 254: 918–919.
44. Kemp TS (1999) *Fossils and evolution*. Oxford: Oxford University Press.
45. Hallam A, Wignall PB (1997) *Mass extinctions and their aftermath*. Oxford: Oxford University Press.
46. Briggs DEG, Crowther PR (2001) *Palaeobiology II*. Oxford: Blackwell Science.
47. Leader-Williams N, Dublin HT (2000) Charismatic megafauna as ‘flagship species’. In: Entwistle A, Dunstone N, eds (2000) *Has the panda had its day? Priorities for the conservation of mammalian diversity*. Cambridge: Cambridge University Press. pp. 53–81.
48. Rodrigues ASL, Andelman SJ, Bakarr MI, Boitani L, Brooks TM, et al. (2004) Effectiveness of the global protected area network in representing species diversity. *Nature* 428: 640–643.
49. Dudgeon D (2005) Last chance to see...: ex situ conservation and the fate of the baiji. *Aquatic Conservation: Marine and Freshwater Ecosystems* 15: 105–108.
50. Sechrest W, Brooks TM, da Fonseca GAB, Konstant WR, Mittermeier RA, et al. (2002) Hotspots and the conservation of evolutionary history. *Proceedings of the National Academy of Sciences of the United States of America* 99: 2067–2071.
51. Rodrigues ASL, Gaston KJ (2002) Maximising phylogenetic diversity in the selection of networks of conservation areas. *Biological Conservation* 105: 103–111.
52. Brooks T, Pilgrim JD, Rodrigues ASL, da Fonseca GAB (2005) Conservation status and geographic distribution of avian evolutionary history. In: Purvis A, Gittleman JL, Brooks T, eds (2005) *Phylogeny and conservation*. Cambridge: Cambridge University Press. pp. 267–294.
53. Hartmann K, Steel M (2007) Phylogenetic diversity: From combinatorics to ecology. In: Gascuel O, Steel M, eds (2007) *Reconstructing evolution: new mathematical and computational advances*. Oxford: Oxford University Press, in press.
54. Soutullo A, Dodsworth S, Heard SB, Mooers AO (2005) Distribution and correlates of carnivore phylogenetic diversity across the Americas. *Animal Conservation* 8: 249–258.
55. Purvis A (1995) A composite estimate of primate phylogeny. *Philosophical Transactions of the Royal Society of London Series B* 348: 405–421.
56. Bininda-Emonds ORP, Gittleman JL, Purvis A (1999) Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). *Biological Reviews of the Cambridge Philosophical Society* 74: 143–175.



A Common Rule for the Scaling of Carnivore Density

Author(s): Chris Carbone and John L. Gittleman

Source: *Science*, New Series, Vol. 295, No. 5563 (Mar. 22, 2002), pp. 2273-2276

Published by: [American Association for the Advancement of Science](#)

Stable URL: <http://www.jstor.org/stable/3076347>

Accessed: 14/08/2011 20:09

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



American Association for the Advancement of Science is collaborating with JSTOR to digitize, preserve and extend access to *Science*.

<http://www.jstor.org>

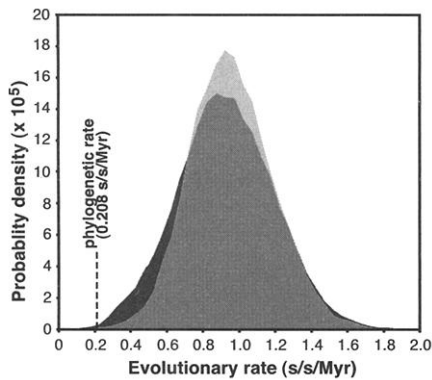


Fig. 3. Posterior probability densities of the evolutionary rate (s/s/Myr) under two models of population dynamics: constant population size (lighter distribution) and exponential growth (darker distribution).

of substitutions per nucleotide site against the time between serially preserved Adélie penguin samples. The regression estimated the rate of HVRI evolution to be 0.676 s/s/Myr; using a parametric bootstrap of 1000 replicates, the 95% confidence intervals were 0 to 2.04 s/s/Myr. The point estimate obtained from this analysis lies well within the two probability distributions obtained from the MCMC analyses. However, the wider confidence interval, which is expected because the method uses only summary distance information and ignores specific site patterns (18), does not exclude the phylogenetically derived estimate.

Mitochondrial HVRI sequences from Adélie penguins are evolving in a clock-like manner in that 89% of all samples belonging to the A and RS lineages passed a relative rate test (19) and a likelihood ratio test (20) ($P > 0.05$) [see the supplemental material (12)]. Estimates of the time of divergence of the A and RS lineages were produced by the MCMC analysis. The mean divergence times were 62,000 years (95% HPD interval 32,000 to 95,000) and 53,000 years (95% HPD interval 26,000 to 90,000) for constant and exponential growth, respectively. Both our point estimates and the 95% intervals indicate that the two lineages diverged during the last glacial cycle (21, 22). This is consistent with the fact that at the Last Glacial Maximum, there were few, if any, ice-free areas in the Ross Sea, and Adélie penguins are likely to have been restricted to refugia.

Although other studies have used ancient DNA to document changes in animal populations over time (23, 24), these data sets have not been used to estimate evolutionary rates. The fast evolutionary rate reported here of two to seven times that of the phylogenetic rate is concordant with the high rate of HVRI mutation found recently in humans (25). We suggest that an evolutionary rate of the mitochondrial HVRI of 0.4 to 1.4 s/s/Myr is more realistic than previous slower phylogenetic estimates, particularly for intraspecific studies and studies

of closely related species. The fact that we have been able to use ancient DNA to measure the tempo of evolution illustrates the importance of these unique Adélie penguin bone deposits.

References and Notes

1. R. L. Cann, M. Stoneking, A. C. Wilson, *Nature* **325**, 31 (1987).
2. G. F. Shields, A. C. Wilson, *J. Mol. Evol.* **24**, 212 (1987).
3. W. M. Brown, M. George, A. C. Wilson, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1967 (1979).
4. T. W. Quinn, *Mol. Ecol.* **1**, 105 (1992).
5. S. Pääbo, *Sci. Am.* **296**, 86 (1993).
6. D. G. Ainley, R. E. LeResche, W. J. L. Sladen, *Breeding Biology of the Adélie Penguin* (Univ. of California Press, Los Angeles, CA, 1983).
7. C. Baroni, G. Orombelli, *Geology* **22**, 23 (1994).
8. Appropriate ancient DNA procedures were employed in a dedicated facility. DNA sequences were deposited in GenBank, with accession numbers AF474792 through AF474887. See the supplemental material (12) and Table 2 for details.
9. DNA from blood samples of 380 Adélie penguins was isolated by means of standard procedures. PCR products were sequenced with the PRISM BigDye Terminator sequencing kit (Applied Biosystems) and analyzed on a 377A automated sequencer (Applied Biosystems) [see the supplemental material (12)]. DNA sequences were deposited in GenBank, with accession numbers AF474412 through AF474791.
10. H.-J. Bandelt, V. Macauley, M. Richards, *Mol. Phyl. Evol.* **16**, 8 (2000).
11. C. Baroni, G. Orombelli, *Quat. Res.* **36**, 157 (1991).
12. Supplemental material, including details of the ¹⁴C-dated material and latitude and longitudinal coordinates of each sampling location, is available on Science Online at www.sciencemag.org/cgi/content/full/295/5563/2270/DC1 or at www.massey.ac.nz/dmlamber.
13. N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
14. W. K. Hastings, *Biometrika* **57**, 97 (1970).
15. A. J. Drummond, G. K. Nicholls, A. G. Rodrigo, W. Solomon, *Genetics*, in press.
16. For details of the MCMC methods used, see the supplemental material (12). Software for implementing these methods is available at www.cebl.auckland.ac.nz/mepi/index.html.
17. A. Drummond, A. G. Rodrigo, *Mol. Biol. Evol.* **17**, 1807 (2000).
18. J. Felsenstein, *Genet. Res. Camb.* **59**, 139 (1992).
19. C.-I. Wu, W.-H. Li, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1741 (1985).
20. J. Felsenstein, *J. Mol. Evol.* **17**, 368 (1981).
21. J. Jouzel et al., *Nature* **364**, 403 (1987).
22. J. R. Petit et al., *Nature* **399**, 429 (1999).
23. E. A. Hadly, M. H. Kohn, J. A. Leonard, R. K. Wayne, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6893 (1998).
24. J. A. Leonard, R. K. Wayne, A. Cooper, *Proc. Natl. Acad. Sci. USA* **97**, 1651 (2000).
25. T. J. Parsons et al., *Nature Genet.* **15**, 363 (1997).
26. Supported by a Marsden grant to D.M.L. We gratefully acknowledge support from Massey University, Pisa University, the Italian Antarctic Research Programme, NSF, Antarctica New Zealand, and the U.S. Coast Guard. In addition, we thank P. Barrett, A. R. Bellamy, T. J. H. Chin, A. Cooper, L. Davis, S. Eyton, R. Forsberg, J. Heine, M. Hendy, K. Huber, K. Kerry, P. Lockhart, R. Marshall, L. Matisoo-Smith, V. Moulton, V. E. Neall, K. Newman, D. Penny, J. Robins, A. G. Rodrigo, L. Shepherd, P. Stapleton, C. Vleck, R. Ward, and E. C. Young. P.A.R. acknowledges the support of a Massey University doctoral scholarship. B.R.H. acknowledges A. Dress and the Program for Scientific and Technological Exchange between New Zealand and Germany. A.J.D. acknowledges a New Zealand Foundation for Research Science and Technology Bright Future Scholarship and research support from an NIH grant.

15 November 2001; accepted 14 February 2002

A Common Rule for the Scaling of Carnivore Density

Chris Carbone^{1*} and John L. Gittleman²

Population density in plants and animals is thought to scale with size as a result of mass-related energy requirements. Variation in resources, however, naturally limits population density and may alter expected scaling patterns. We develop and test a general model for variation within and between species in population density across the order Carnivora. We find that 10,000 kilograms of prey supports about 90 kilograms of a given species of carnivore, irrespective of body mass, and that the ratio of carnivore number to prey biomass scales to the reciprocal of carnivore mass. Using mass-specific equations of prey productivity, we show that carnivore number per unit prey productivity scales to carnivore mass near -0.75, and that the scaling rule can predict population density across more than three orders of magnitude. The relationship provides a basis for identifying declining carnivore species that require conservation measures.

Across communities in plants and animals, there is an inverse relationship between population density and body size, such that resource use and availability are driving con-

sistent statistical patterns (1–5). The critical factor is the individual species' rate of resource use. Typically, resource use is identified in general metabolic or physiological terms, as these represent the invariant properties of all biological systems at different levels. The precise measure and form of resource use have only been described indirectly (6–9).

We developed a general model (10) to predict carnivore density relative to resource-

¹Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, UK. ²Department of Biology, Gilmer Hall, University of Virginia, Charlottesville, VA 22904, USA.

*To whom correspondence should be addressed. E-mail: chris.carbone@ioz.ac.uk

REPORTS

es, expressed as prey biomass and prey productivity (11–14). We tested this model with data from the literature on density of 25 species of carnivores (15–20) and their most common prey (21) (Table 1). For each species, we calculated the average number of carnivores per unit prey biomass (i.e., carnivore number per 10,000 kg of prey). Controlling for prey biomass allows us to account for the wide variation in carnivore density resulting from variation in prey density within species, as well as to make comparisons between species.

Within carnivore species, population density is typically positively correlated with prey biomass (Fig. 1). In keeping with the assumption that a species' population density is influenced by individual rates of resource use (4), the number of carnivores supported on a given biomass of prey increases with decreasing body size. Comparing between species, we find a strong negative relationship between the number of carnivores per 10,000 kg of prey and carnivore body mass (Fig. 2A). The relationship takes the form of a power function [number per 10,000 kg of prey = $89.1 \times (\text{carnivore mass})^{-1.05}$; $N = 25$, $R^2 = 0.83$, $P < 0.0001$]. The exponent does not differ significantly from -1.0 [95% confidence limits, -0.845 (upper), -1.25 (lower); confidence limits for constant, 169 (upper), 47 (lower)] (22–24).

Our results depend on controlling for prey biomass. A plot of average carnivore population density (number per 100 km²) against carnivore body mass has considerably more variation than in the biomass-based analyses (Fig. 2B) [number per 100 km² = $197.6 \times (\text{carnivore mass})^{-0.88}$; $N = 25$, $R^2 = 0.63$, $P < 0.0001$; confidence limits for exponent, -0.59 (upper), -1.18 (lower); confidence limits for constant, 500 (upper), 78 (lower)].

An example of the importance of controlling for prey biomass can be seen by comparing the European badger (*Meles meles*) (15) and the coyote (*Canis latrans*) (19), both of which weigh about 13 kg. These species differ in average population density by a factor of almost 20, but this is due to a nearly 40-fold difference in the prey biomass density available to these species. Our biomass-based estimate of population density differs by a factor of only 1.6 (Table 1).

Previous studies have pointed out that density estimates of different-sized species may be confounded by sampling area (25, 26). Although the density values for carnivores and their prey may both be influenced by the sampling area, it is unlikely that this factor would bias our estimates of the predator-prey relationships in a way that would influence the overall allometric relationship shown in Fig. 2A. In addition, previous analyses of wolf population data (27) (Table 1) found that the inclusion of

sampling area in a multiple regression model did not substantially improve the model fit.

Ultimately, predator populations are sustained by population productivity rates of their prey rather than by standing biomass. Estimates of turnover on a population-by-population basis are not available, but biomass-based population productivity measures have been estimated in relation to body mass (11–13). We expected that the number of carnivores per unit prey biomass would vary with $(\text{carnivore mass})^{-1.0}$ and that the carnivore number per unit productivity would vary with $(\text{carnivore mass})^{-0.75}$ (10). We plotted the average ratio of carnivore number per unit productivity (number per 10,000 kg per year) against carnivore mass (Fig. 2C). This relationship has an exponent not significantly different from -0.75 [number per unit productivity = $56.2 \times (\text{carnivore mass})^{-0.66}$; $N = 24$, $R^2 = 0.70$, $P < 0.0001$; exponent confidence limits, -0.48 (upper), -0.85 (lower); confidence limits for constant, 101 (upper), 31 (lower)] (28). These findings support the notion that there is no systematic variation in prey productivity between carnivore species, and that carnivore density is constrained by metabolic rates and prey abundance.

We selected species that provide a range of body sizes, habitats, and feeding strategies; these include an invertebrate-feeder [the European badger (15)] and vertebrate hunter

Table 1. Summary of carnivore density and prey biomass. The number of carnivores per 10,000 kg of prey biomass was estimated from the ratio of carnivore population density (number per 100 km²) to biomass density (in units of 10,000 kg per 100 km²) of the main prey species averaged for each species. These values were used in Fig. 2A. Minimum and maximum estimates of the carnivore density and prey biomass density obtained for this study are provided (43).

Carnivore species	Average mass (kg)	Number of populations	Carnivore density		Prey biomass (10,000 kg per 100 km ²)
			Number per 10,000 kg of prey biomass (species average)	Number per 100 km ²	
Least weasel* (<i>Mustela nivalis</i>)	0.14	6	1656.49	183.67 to 8000.00	0.24 to 8.33
Ermine (<i>Mustela erminea</i>)	0.16	2	406.66	105.00 to 1333.33	0.38 to 2.47
Small Indian mongoose (<i>Herpestes javanicus</i>)	0.55	3	514.84	1300.00 to 2850.00	3.00 to 5.00
Pine marten (<i>Martes martes</i>)	1.3	1	31.84	56.80	1.78
Kit fox (<i>Atelocynus microtis</i>)	2.02	2	15.96	16.00 to 24.00	0.66 to 3.05
Channel Island fox (<i>Urocyon littoralis</i>)	2.16	1	145.94	957	6.56
Arctic fox (<i>Alopex lagopus</i>)	3.19	14	21.63	2.22 to 28.57	0.01 to 28.11
Red fox (<i>Vulpes vulpes</i>)	4.6	2	32.21	10.00 to 112.00	1.19 to 2.00
Bobcat (<i>Lynx rufus</i>)	10.0	1	9.75	3.40	0.35
Canadian lynx* (<i>Lynx canadensis</i>)	11.2	26	4.79	1.99 to 22.59	0.17 to 13.86
European badger (<i>Meles meles</i>)	13.0	8	7.35	110.00 to 840.00	6.20 to 714.00
Coyote* (<i>Canis latrans</i>)	13.0	16	11.65	2.29 to 44.44	0.35 to 14.85
Ethiopian wolf (<i>Canis simensis</i>)	14.5	2	2.70	19.00 to 120.00	25.00 to 31.34
Eurasian lynx (<i>Lynx lynx</i>)	20.0	2	0.46	0.98 to 2.90	2.31 to 5.90
Wild dog (<i>Lycaon pictus</i>)	25.0	10	1.61	0.07 to 15.00	0.16 to 110.00
Dhole (<i>Cuon alpinus</i>)	25.0	2	0.81	13.00 to 30.00	17.05 to 34.94
Snow leopard (<i>Uncia uncia</i>)	40.0	1	1.89	5.75	3.04
Wolf (<i>Canis lupus</i>)	46.0	23	0.62	0.50 to 4.20	0.89 to 8.10
Leopard (<i>Panthera pardus</i>)	46.5	19	6.17	0.50 to 37.04	0.03 to 41.63
Cheetah (<i>Acinonyx jubatus</i>)	50.0	7	2.29	0.61 to 7.79	0.16 to 6.69
Puma (<i>Puma concolor</i>)	51.9	3	0.94	0.37 to 7.00	1.00 to 11.00
Spotted hyena (<i>Crocuta crocuta</i>)	58.6	12	0.68	0.59 to 184.19	1.26 to 121.46
Lion (<i>Panthera leo</i>)	142.0	21	3.40	0.80 to 38.50	0.01 to 116.99
Tiger (<i>Panthera tigris</i>)	181.0	6	0.33	0.70 to 15.84	4.00 to 89.54
Polar bear* (<i>Ursus maritimus</i>)	310.0	8	0.60	0.28 to 2.11	0.42 to 3.37

*Includes more than one population estimate from the same area in relation to annual changes in prey density.

REPORTS

specialists [e.g., the African lion (*Panthera leo*) (16), leopard (*Panthera pardus*) (16, 17), and polar bear (*Ursus maritimus*) (18)]. Despite the wide variation in species' ecology, we find remarkable consistency in the average population density in relation to prey biomass and carnivore mass. However, some of the residual variation in population density can be explained in terms of species' biology. For example, interspecific predation and competition is a major factor influencing carnivore population density (29). African wild dogs (*Lycaon pictus*) and cheetahs (*Acinonyx jubatus*) can be found at lower densities in areas where prey are very abundant because of the abundance of competing lions and spotted hyenas in these areas (30, 31).

Clearly, all species are influenced to some degree by competition with other carnivores, and this must contribute to the variation found in density estimates across populations. Furthermore, the temporal responses of carnivore density to changes in prey may be somewhat related to turnover rates in different-sized prey (29). Lynx (*Lynx canadensis*) and coyotes (*Canis latrans*) feeding primarily on smaller prey such as rodents and hares show more rapid functional responses than do larger carnivores such as Isle Royale wolves (*Canis lupus*), which require 3 to 5 years to respond to population changes in moose numbers (29). As more data become available, our predictive model should be refined to quantitatively show the effect of these ecological differences in species abundance.

Allometric scaling, frequently used in biology to extrapolate trait values for species that are relatively unknown, is increasingly being applied to the prediction of population numbers

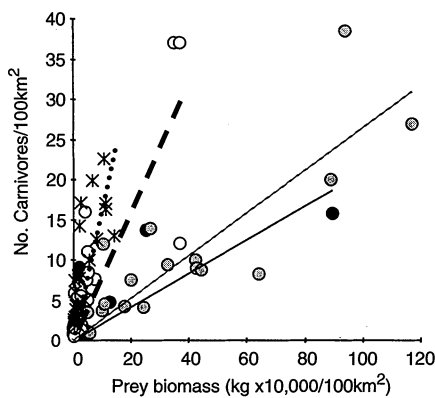


Fig. 1. Carnivore density (number per 100 km²) plotted against prey biomass density (in units of 10,000 kg per 100 km²) for different species of carnivores. For the purposes of illustration, we show the slopes of the regression (plotted through the origin) estimated for each species (see text for details): solid circles and solid line, tiger (*Panthera tigris*); shaded circles and gray line, lion (*Panthera leo*); open circles and dashed line, leopard (*Panthera pardus*); asterisks and dotted line, Canadian lynx (*Lynx canadensis*).

for rare, endangered, and threatened species (32–34). Scaling studies that control for key ecological variables (such as resource availability) may provide an important framework for identifying species that deviate from expected values because of other ecological processes. The data on the Eurasian lynx cited in this study provide an example (Fig. 2A). This species is rare relative to the estimated prey biomass availability (35–37). One population was recently reintroduced and both populations have been exposed to poaching, possibly contributing to the relatively low densities at these sites.

Mammalian carnivores are often uniquely characterized by fine-tuned relationships with

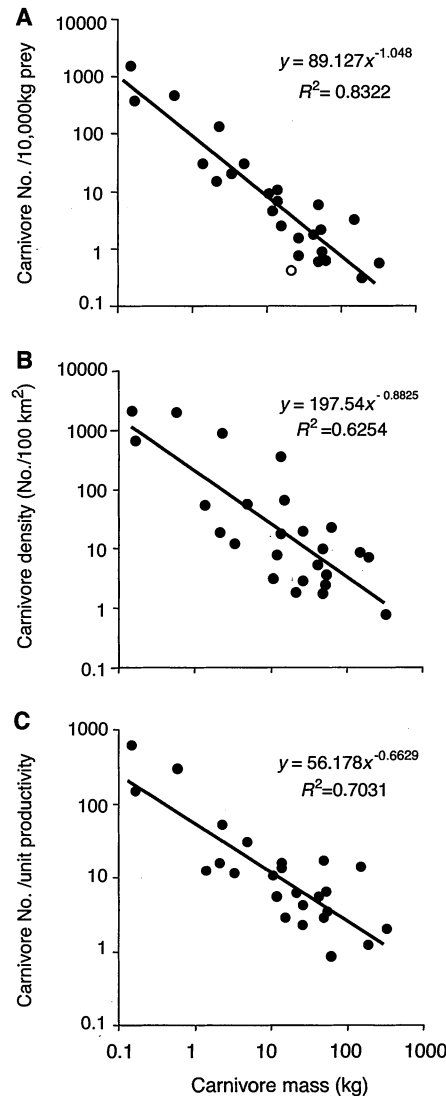


Fig. 2. Three measures of carnivore density plotted against carnivore body mass (plotted on a log-log scale): (A) number of carnivores per 10,000 kg of prey, (B) average carnivore density (number per 100 km²), and (C) number of carnivores per unit prey productivity (number per 10,000 kg of prey productivity per year) (see text for details). In (A), the Eurasian lynx is represented by the open circle; excluding this species, the regression is $y = 94.54x^{-1.03}$ ($R^2 = 0.86$).

their prey (38–40). It appears that carnivores are closely tied not only to prey size (14) but also to prey biomass. Carnivore populations and species are now rapidly dwindling in numbers. At least 90 carnivore species are currently listed as threatened or endangered (41). Our results show that prey density is a fundamental determinant of carnivore density both within and between species. Given that carnivore population density has been identified as a predictive factor influencing extinction risk (42), prey density is critical to the future of stable carnivore populations.

References and Notes

1. P. A. Marquet, *Science* **289**, 1487 (2000).
2. J. H. Brown, *Macroecology* (Univ. of Chicago Press, Chicago, 1995).
3. J. H. Brown, G. B. West, B. J. Enquist, in *Scaling in Biology: Patterns, Processes, Causes and Consequences*, J. H. Brown, G. B. West, Eds. (Oxford Univ. Press, Oxford, 2000), pp. 1–24.
4. B. J. Enquist, J. H. Brown, G. B. West, *Nature* **395**, 163 (1998).
5. J. Damuth, *Nature* **290**, 699 (1981).
6. ———, *Nature* **395**, 115 (1998).
7. F. Magnani, *Nature* **398**, 572 (1999).
8. B. J. Enquist, J. H. Brown, G. B. West, *Nature* **398**, 573 (1999).
9. R. C. Dewar, *Nature* **398**, 572 (1999).
10. Following (4–6), maximum population size (N_{max}) for species i can be expressed as $N_{max_{ij}} = a \times (R_j / M_i^b)$ where M_i^b represents the mass related metabolic rate (typically, $b \approx 0.75$) of species i in population j (assuming body mass is constant across populations), R_j represents the resources available to the population j , and a is a constant. We compiled data from studies that had estimates of both carnivore and prey density. We approximated R_j for each carnivore population j by (i) calculating the prey biomass per 100 km² [multiplying species weight (kg) by species density (number per 100 km²)], and (ii) calculating for each species the mass-specific productivity per unit biomass (11–13). The average ratios of carnivore density (number per 100 km²) to prey biomass (in units of 10,000 kg per 100 km²) and carnivore density to prey productivity (in units of 10,000 kg per 100 km² per year) were calculated across populations for each carnivore species. The productivity/biomass ratio typically varies with (mass)^{-0.25} (13), indicating that productivity is independent of body mass (5, 6). The ratio of carnivore number to prey productivity [(mass)^{-0.75}/(mass)⁰] should then simply scale as (mass)^{-0.75}, the exponent representing the inverse of the metabolic rate (b).
11. J. O. Farlow, *Ecology* **57**, 841 (1976).
12. K. Banse, S. Moshier, *Ecol. Monogr.* **50**, 355 (1980).
13. R. H. Peters, *The Ecological Implications of Body Size* (Cambridge Univ. Press, Cambridge, 1983).
14. C. Carbone, G. M. Mace, S. C. Roberts, D. W. Macdonald, *Nature* **402**, 286 (1999).
15. H. Kruuk, T. Parish, *J. Zool.* **196**, 31 (1982).
16. R. East, *Afr. J. Ecol.* **22**, 245 (1984).
17. T. N. Bailey, *The African Leopard—Ecology and Behavior of a Solitary Felid* (Columbia Univ. Press, New York, 1993).
18. I. Stirling, N. A. Oritsland, *Can. J. Fish Aquat. Sci.* **52**, 2594 (1995).
19. M. Odonoghue, S. Boutin, C. J. Krebs, E. J. Hofer, *Oikos* **80**, 150 (1997).
20. A. Angerbjorn, M. Tannerfeldt, S. Erlinge, *J. Anim. Ecol.* **68**, 34 (1999).
21. The most common prey species were usually identified in the literature (43) and constituted 70% or more of the diet. Where prey biomass was not estimated, species weights were multiplied by species density.
22. In a multivariate analysis, we find that carnivore density is significantly related to prey biomass but

- not to carnivore mass ($F = 26.94$; $df = 2, 22$; $P < 0.0001$; prey biomass: $F = 6.42$, $P < 0.016$; body mass: $F = 0.26$, not significant). Controlling for phylogeny (23, 24), we get a similar result: $\ln(\text{number per } 10,000 \text{ kg of prey}) + 1 = -1.12 \times \ln(\text{carnivore mass})$; $r = 0.753$, $P < 0.01$. One contrast in the body mass and prey biomass analysis was excluded, calculated between *Lynx* and *Panthera*, because it had a Studentized deleted residual greater than 3.
23. A. Purvis, A. Rambaut, *Comput. Appl. Biosci.* **11**, 247 (1995).
 24. O. R. P. Bininda-Emonds, J. L. Gittleman, A. Purvis, *Biol. Rev.* **74**, 143 (1999).
 25. K. S. Smallwood, C. Schonewald, *Oecologia* **105**, 329 (1996).
 26. T. M. Blackburn, K. J. Gaston, *Oikos* **75**, 303 (1996).
 27. T. K. Fuller, D. L. Murray, *Anim. Conserv.* **1**, 153 (1998).
 28. Controlling for phylogeny yields the following expression: $\ln(\text{number per productivity}) = -0.36 \times \ln(\text{carnivore mass})$; $r = 0.449$, $P < 0.05$. The contrast between canids and felids had a large leverage and was removed.
 29. T. K. Fuller, P. R. Sievert, in *Carnivore Conservation*, J. L. Gittleman, S. M. Funk, D. W. Macdonald, R. K. Wayne, Eds. (Cambridge Univ. Press, Cambridge, 2001), pp. 163–178.
 30. M. G. L. Mills, M. L. Gorman, *Conserv. Biol.* **11**, 1397 (1997).
 31. K. M. Laurenson, N. Wielebnowski, T. M. Caro, *Conserv. Biol.* **9**, 1329 (1995).
 32. W. A. Calder, in *Scaling in Biology: Patterns, Processes, Causes and Consequences*, J. H. Brown, G. B. West, Eds. (Oxford Univ. Press, Oxford, 2000), pp. 297–323.
 33. K. J. Gaston, T. M. Blackburn, *Patterns and Processes in Macroecology* (Blackwell Scientific, Oxford, 2000).
 34. S. P. Hubbell, *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, Princeton, NJ, 2001).
 35. A. Molinari-Jobin, P. Molinari, C. Breitenmoser-Würsten, U. Breitenmoser, *Wildlife Biol.*, in press.
 36. A. Jobin, P. Molinari, U. Breitenmoser, *Acta Theriol.* **45**, 243 (2000).
 37. B. Jedrzejewski, W. Jedrzejewski, *Predation in Vertebrate Communities: The Baialowieza Primeval Forest as a Case Study* (Springer-Verlag, Berlin, 1998).

38. A. F. Vezina, *Oecologia* **67**, 555 (1985).
39. J. L. Gittleman, M. E. Gompper, *Science* **291**, 997 (2001).
40. J. Berger, J. E. Swanson, I. L. Persson, *Science* **291**, 1036 (2001).
41. C. Hilton-Taylor (Compiler), *2000 IUCN Red List of Threatened Species* (IUCN–The World Conservation Union, Gland, Switzerland, 2000).
42. A. Purvis, J. L. Gittleman, G. Cowlshaw, G. M. Mace, *Proc. R. Soc. London Ser. B* **267**, 1947 (2000).
43. A version of the table with a full list of references can be found at *Science Online* at www.sciencemag.org/cgi/content/full/295/5563/2273/DC1.
44. We thank K. Gaston, J. Brown, K. Jones, P. Bennett, S. Funk, M. Rowcliffe, C. Mueller, A. Bourke, G. Mace, T. Coulson, S. Semple, J. Du Toit, J. Fulford, I. J. Gordon, and three anonymous referees for helpful discussions and comments on earlier drafts of the manuscript. We are grateful to the following for access to data: A. Jobin, Z. T. Ashenafi, K. Murphy, S. Roy, A. Venkataraman, A. T. Johnsingh, A. Angebjorn, and T. Coonan.

13 November 2001; accepted 21 January 2002

Neuronal Calcium Sensor 1 and Activity-Dependent Facilitation of P/Q-Type Calcium Currents at Presynaptic Nerve Terminals

Tetsuhiro Tsujimoto,^{1*†} Andreas Jeromin,^{2*} Naoto Saitoh,¹ John C. Roder,² Tomoyuki Takahashi¹

P/Q-type presynaptic calcium currents (I_{pCa}) undergo activity-dependent facilitation during repetitive activation at the calyx of the Held synapse. We investigated whether neuronal calcium sensor 1 (NCS-1) may underlie this phenomenon. Direct loading of NCS-1 into the nerve terminal mimicked activity-dependent I_{pCa} facilitation by accelerating the activation time of I_{pCa} in a Ca^{2+} -dependent manner. A presynaptically loaded carboxyl-terminal peptide of NCS-1 abolished I_{pCa} facilitation. These results suggest that residual Ca^{2+} activates endogenous NCS-1, thereby facilitating I_{pCa} . Because both P/Q-type Ca^{2+} channels and NCS-1 are widely expressed in mammalian nerve terminals, NCS-1 may contribute to the activity-dependent synaptic facilitation at many synapses.

Neurotransmitter release is triggered by Ca^{2+} influx through presynaptic voltage-dependent Ca^{2+} channels (I). Modulation in the presynaptic calcium current (I_{pCa}) results in robust alteration of synaptic efficacy because of their nonlinear relationship (2). At the calyx of Held nerve terminal, repetitive activation of Ca^{2+} channels increases the amplitudes of I_{pCa} (3–5). The magnitude of I_{pCa} facilitation is dependent on the extracellular Ca^{2+} concentration and is attenuated by intraterminal loading of Ca^{2+} chelating agents (4, 5). This

I_{pCa} facilitation is distinct from the voltage-dependent relief of Ca^{2+} channels from tonic inhibition by heterotrimeric guanine nucleotide binding (G) proteins (6, 7), because presynaptic loadings of guanine nucleotide analogs have no effect (4). A Ca^{2+} -binding protein may thus be involved in the activity-dependent I_{pCa} facilitation.

Among neuron-specific Ca^{2+} -binding proteins, frequenin was first cloned from *Drosophila T(X;Y) V7* mutants (8). Later, the frequenin homolog NCS-1 was cloned from a variety of species (9–14). NCS-1 (frequenin) is widely expressed in neuronal somata, dendrites, and nerve terminals (14–18) throughout embryonic and postnatal development (14, 17). Overexpression (19) or intracellular loading of NCS-1 in motoneurons (10) enhances neuromuscular transmission. We investigated whether NCS-1 is involved in the

activity-dependent I_{pCa} facilitation at the calyx of Held synapse.

Whole-cell voltage-clamp recordings were made from a calyceal nerve terminal (20), and I_{pCa} was elicited by an action potential waveform command pulse at 0.1 Hz. The half-width and the peak amplitude of a prerecorded action potential were similar to those reported for afferent fiber-stimulated action potentials in 14-day-old rats (21). After a stable epoch of I_{pCa} , NCS-1 was infused into a nerve terminal through a perfusion tube (Fig. 1A). After infusion, amplitudes of I_{pCa} gradually increased, reached a maximum in 5 min, and then gradually declined. This decline may be caused by “adaptation” in the mechanism of facilitation by NCS-1, because I_{pCa} elicited at 0.1 Hz does not undergo run-down for more than 20 min (22). The mean magnitude of I_{pCa} facilitation 5 min after the onset of NCS-1 infusion was $113 \pm 37\%$ (mean \pm SEM, $n = 3$).

We next examined the effect of NCS-1 on I_{pCa} elicited by a 5-ms depolarizing pulse. When NCS-1 was included in the presynaptic pipette solution, the rise time of I_{pCa} was significantly faster than rise times in the presence of heat-inactivated (H.I.) NCS-1 or in the absence of NCS-1 [Fig. 1, B (inset) and C]. The current-voltage (I - V) relationship of I_{pCa} measured at 1 ms after the onset of the command pulse had a peak at -10 mV in the presence of NCS-1, whereas the peaks were at 0 mV in the presence of H.I. NCS-1 or in the absence of NCS-1 (Fig. 1B). Similarly, in the presence of NCS-1, the half-activation voltage ($V_{1/2}$) calculated from the modified Boltzmann equation (20) was significantly more negative than those in the presence of H.I. NCS-1 or in the absence of NCS-1 (Fig. 1D). However, NCS-1 had no effect on the magnitude of plateau Ca^{2+} currents (Fig. 1E).

NCS-1 has four helix-to-helix Ca^{2+} -binding architectures (EF-hands) and binds three

¹Department of Neurophysiology, University of Tokyo Faculty of Medicine, Tokyo 113-0033, Japan.

²Mount Sinai Hospital Research Institute, Toronto, Ontario M5G 1X5, Canada.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: tujimoto-ty@umin.ac.jp

Ecology, Sexual Selection, and the Evolution of Mating Systems

Stephen T. Emlen; Lewis W. Oring

Science, New Series, Vol. 197, No. 4300 (Jul. 15, 1977), 215-223.

Stable URL:

<http://links.jstor.org/sici?sici=0036-8075%2819770715%293%3A197%3A4300%3C215%3AESSATE%3E2.0.CO%3B2-E>

Science is currently published by American Association for the Advancement of Science.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/aaas.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.



Ecology, Sexual Selection, and the Evolution of Mating Systems

Stephen T. Emlen and Lewis W. Oring

Mating systems (1) were first discussed in evolutionary terms by Darwin (2). Since then, major developments in genetic theory have allowed a better understanding of sex ratios, sexual dimorphism, and differential patterns of parental care (3-7). Important milestones toward an ecological understanding of mating systems have also been reached (8-16). Nevertheless, attempts at synthesizing natural history data into a unified theory of mating system evolution have lagged behind the development of population genetics theory.

One factor hindering development of a sociobiological framework of mating system theory has been a recurring tendency for field workers to search for and to discuss "adaptiveness" in the context of the survival or well-being of the population or species. To understand mating systems, we must abandon species- or group-selection viewpoints and return to the evolutionary tenet of natural selection operating at the level of the individual genome (17).

Fitness, in genetic theory, measures the reproductive success of an individual (or a genotype) measured relative to the reproductive success of other individuals (or genotypes) in the same or in other populations. Thus, we should expect a strong competitive element in many aspects of reproductive behavior. Darwin was fully aware of this intraspecific competition when he introduced the theory of sexual selection (2). Stated simply, when one sex becomes a limiting factor for the other, the result is an increase in intrasexual competition among members of the available sex for access to mates

of the limiting sex. The greater the "shortage" of one sex, the more intense the sexual selection. Sexual selection is relatively slight in monogamous groups while it is intense in highly polygamous societies. A better understanding of the causes of sexual selection is thus pivotal to the development of any mating system theory.

What accounts for the differing intensities of sexual selection found in different species and, frequently, between populations of the same species? We hypothesize that one important cause is the *ability of a portion of the population to control the access of others to potential mates*. This control can be direct, as in the physical herding of potential mates, and the physical exclusion of other members of the same sex from these mates, or indirect, by controlling resources that are critical either for mate attraction or for successful reproduction. The greater the degree of control or monopolization, the greater the resulting variance in mating success. As the difference in reproduction success of "haves" and "have nots" increases, so too does intrasexual competition to be in the former category. It becomes crucial to understand the manner in which access to mates can be controlled.

One of the emerging findings of sociobiology is that many aspects of an animal's social organization can be predicted on the basis of an understanding of a limited set of environmental variables. We propose that certain environmental factors determine the degree to which mates can be defended or monopolized. In effect, *ecological constraints impose*

limits on the degree to which sexual selection can operate. The greater the potential for multiple mate monopolization, the greater should be the potential intensity of sexual selection and the tendency for polygamy.

In this article, we discuss various ecological factors that influence the degree and form of polygamous mating systems. Our intention is to develop a theory adequate for predicting environmental influences on the evolution and expression of avian mating systems. We believe these predictions are also broadly applicable to many mammalian groups, as well as to certain insects and lower vertebrates. They are less applicable to strongly "r"-selected species (18), or for species that normally live in permanent groups with restricted membership where the potential for reciprocal interactions is high (19).

Economic Monopolizability and the Environmental Potential for Polygamy

There are two preconditions for the evolution of polygamy (see Fig. 1). First, multiple mates, or resources sufficient to attract multiple mates, must be energetically defendable by individuals. Many environmental factors affect the potential for such control. But primary among these are the spatial and temporal patterns of resource dispersion. The potential for controlling access to multiple mates should be viewed in a cost-benefit context. Stated simply, the prerequisite for a monogamous mating system is the economic defendability of a mate. The prerequisite for a polygamous mating system is the economic monopolizability of several mates. The *environmental potential for polygamy* depends on the degree to which multiple mates, or resources critical to gaining multiple mates, are economically defendable.

A second prerequisite is the ability of the animals to utilize this potential. Critical resources might be distributed such

Dr. Emlen is professor of animal behavior in the Division of Biological Sciences, Cornell University, Ithaca, New York 14853. Dr. Oring is associate professor in the Department of Biology, University of North Dakota, Grand Forks 58201.

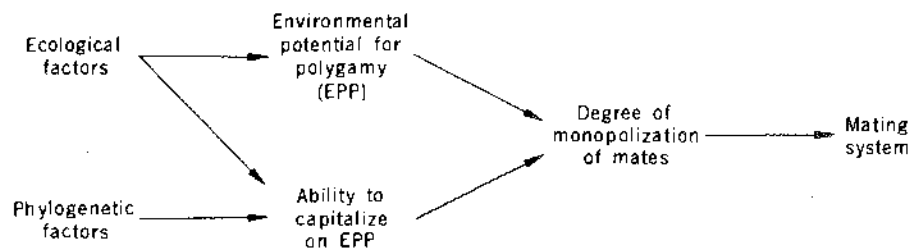


Fig. 1. A general schema of the determinants of a mating system.

that they are economically defensible for an individual that expends most of its time budget on such defense, but not for one that devotes considerable time to alternative activities. The degree to which an animal can take advantage of the "polygamy potential" of the environment depends in large part on the degree of parental care required for successful rearing of young.

Recent attempts at synthesizing the literature on mating systems have focused on this latter point. Polygamy is more prevalent in species where one sex is freed from parental care duties. Members of this sex can expend increased time and energy on intrasexual competition for resources and mates. Several predictions logically follow and have been stressed by others: polygamy should be more common among animals where (i) one sex is predisposed to assume most of the parental care (for example, mammals); (ii) parental care requirements are minimal (for example, birds with precocial as compared with altricial young); and (iii) a superabundant food resource enables a single parent to provide full parental care (7, 11-13). Emancipation by itself, however, need not lead to the evolution of polygamy. It merely assures that one sex can fully exploit the potential for mate monopolization inherent in the environment (see discussion on resource defense polygyny, below).

The spatial distribution of resources. Brown introduced the concept of economic defendability to the study of animal spacing patterns (20). We extend his approach to encompass mating systems.

When important resources are distributed uniformly in space, there is little opportunity for resource monopolization. If the resources are sufficiently abundant and stable through time, territoriality typically occurs. Members of the breeding population would tend toward even dispersion and the potential for multiple matings would be low. Sexual selection would be minimal, and the fitness of individuals might be maximized by sharing equally in parental care duties (see monogamy).

As critical resources become unevenly distributed in nature, the potential for obtaining additional mates increases. This is because some individuals may be able to control a larger quantity or better quality of resource than other individuals. To the degree that this influences mate choice, polygamous matings may be expected (Fig. 2, horizontal axis).

When important resources are highly clumped, the possibility arises for a small percentage of the population to monopolize a large proportion of the available resources. Sexual selection and variance in reproductive success should be high, and we speak of the environment as having a high polygamy potential (EPP).

The temporal distribution of mates. In our model, the benefits derived from resource defense lie in an increased probability of mate attraction and acquisition. The magnitude of this benefit is strongly influenced by the temporal pattern of availability of sexually receptive partners.

If females in a population become sexually receptive in unison, there is little potential for individual males to monopolize multiple females. This will be especially true if each female is sexually active for only a brief period. By the time

the sexes have located one another, and normal courtship and mating have taken place, most of the remaining available females have already been inseminated and few new partners are available. As long as the time involved in servicing a single sexual partner constitutes a significant portion of the total time that potential mates are available, trends toward polygamy will be minimal.

With increasing degrees of asynchrony among members of one sex, the potential for individuals of the other sex to accumulate multiple mates increases. Among species in which both sexes contribute to parental care, a moderate degree of asynchrony is essential for the expression of polygamy. In species where one sex is largely freed from parental duties, individuals of this sex should remain sexually active for the duration of the period during which members of the other sex become sexually receptive. The intensity of sexual selection then will be determined by the degree to which critical resources are differentially controlled by members of the limited sex, by the availability of members of the limiting sex (21), and by the process of mate selection.

As the degree of asynchrony becomes extreme, the rate of appearance of new potential mates reaches a point at which the cost of continued resource defense necessary for attracting an additional mate exceeds the additional benefits gained. Continued mate accumulation would no longer be energetically practical and tendencies toward polygamy will decrease (see Fig. 2, vertical axis).

The concept of operational sex ratios. To understand the intensity of sexual selection it is not the overall population ratio of males to females that is of importance but rather what we term the operational sex ratio (OSR)—defined as the average ratio of fertilizable females to sexually active males at any given time (22). This ratio is strongly affected by the degree of spatial and temporal clumping of the limiting sex. For example, continuous long periods of sexual activity by males, coupled with brief and asynchronous periods of receptivity by females, will produce a strong skew in the OSR.

The OSR provides an empirical measure of the degree of monopolizability of mates. The greater the degree of imbalance in the OSR, the greater the expected variance in reproductive success among members of the limited sex and the greater the degree of polygamy. Where the OSR is skewed toward males, polygyny is expected; when the skew is toward females, polyandry should occur (23).

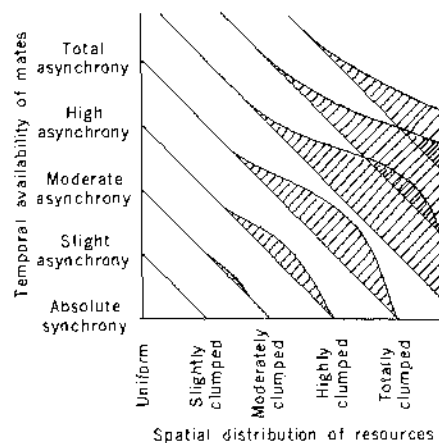


Fig. 2. Graphic representation of the environmental potential for polygamy (indicated by the perpendicular height of the shaded area) and its relation to the spatial distribution of resources and temporal availability of receptive mates.

Types of Mating Systems

Attempts to categorize mating systems have been hampered by a lack of generally accepted terminology. Classically, mating systems have been defined according to the number of mates that one sex can accumulate. Sometimes this is modified to separate simultaneous from sequential multiple mate acquisition. Selander (14) has proposed a new classification based, in part, on the duration of the pair bond or mate association. The difficulty with these classification schemes is that they leave functionally or causatively unrelated situations (or both) in the same category. They ignore the importance of environmental pressures on parental care and sexual selection, and the influence of these factors on mating system evolution. In this article we have characterized mating systems on the basis of the ecological and behavioral potential to monopolize mates, and by the means through which such monopolization takes place. Where male and female strategies conflict, we use terms appropriate for the sex that controls the resource base or monopolizes multiple mates (or both) (see Table 1).

Monogamy

Neither sex has the opportunity to monopolize additional members of the opposite sex, directly or through resource control. Multiple breedings per season may occur in sequence.

Monogamy is expected to occur when (i) there is no environmental "polygamy potential," or (ii) there is no opportunity to take advantage of what "polygamy potential" the environment affords. Monogamy is the dominant avian mating system, occurring in more than 90 percent of the species studied (24), but it is believed rare among mammals (25).

The prevalence of monogamy among birds is due primarily to the inability of most species to take advantage of any environmental "polygamy potential." Considerable parental care by both parents often is required for successful rearing of young. Thus, losses to an individual parent accrued by withholding care from one set of offspring while courting and mating with additional mates may be greater than the gains resulting from such behavior.

In mammals, the preponderance of female parental care allows males of most species to exploit whatever "polygamy potential" exists. Nevertheless, monogamy may be more common than is usually supposed, especially in populations

Table 1. An ecological classification of mating systems.

<i>Monogamy</i> : Neither sex has the opportunity of monopolizing additional members of the opposite sex. Fitness often maximized through shared parental care.
<i>Polygyny</i> : Individual males frequently control or gain access to multiple females.
<i>Resource defense polygyny</i> : Males control access to females <i>indirectly</i> , by monopolizing critical resources.
<i>Female (or harem) defense polygyny</i> : Males control access to females <i>directly</i> , usually by virtue of female gregariousness.
<i>Male dominance polygyny</i> : Males or critical resources are <i>not economically monopolizable</i> . Males aggregate during the breeding season and <i>females select mates</i> from these aggregations.
<i>Explosive breeding assemblages</i> : Both sexes converge for a short-lived, highly synchronized mating period. The operational sex ratio is close to unity and sexual selection is minimal.
<i>Leks</i> : Females are less synchronized and males remain sexually active for the duration of the females' breeding period. Males compete directly for dominant status or position within stable assemblages. Variance in reproductive success and skew in operational sex ratio reach extremes.
<i>Rapid multiple clutch polygamy</i> : Both sexes have substantial but relatively <i>equal</i> opportunity for increasing fitness through multiple breedings in rapid succession. Males and females each incubate separate clutches of eggs.
<i>Polyandry</i> : Individual females frequently control or gain access to multiple males.
<i>Resource defense polyandry</i> : Females control access to males <i>indirectly</i> , by monopolizing critical resources.
<i>Female access polyandry</i> : Females do not defend resources essential to males but, through interactions among themselves, may limit access to males. Among phalaropes, both sexes converge repeatedly at ephemeral feeding areas where courtship and mating occur. The mating system most closely resembles an explosive breeding assemblage in which the OSR may become skewed with an excess of females.

where individuals are widely dispersed over relatively uniform environments (26). As was mentioned previously, male emancipation by itself should only lead to polygyny under permissive environmental conditions.

If the potential of, or the gain from, mate monopolization is nonexistent, an individual should benefit by remaining with its initial mate and acting in such a manner as to maximize the survival chances of its offspring. Recent studies of long-lived birds show a clear advantage to long-term mate fidelity (27). Birds breeding with former mates show low levels of aggression and a high degree of within-pair synchronization, allowing them to breed more rapidly and efficiently and leading to a demonstrable increase in reproductive success. The longer the period of mate fidelity, the more the future physical condition of a mate becomes of importance to its partner. It then becomes adaptive to equalize the energetic burden of reproduction and to share in parental care.

Forms of Polygyny

Polygyny occurs if environmental or behavioral conditions bring about the clumping of females, and males have the capacity to monopolize them. Types of polygyny are classified according to the means that males use to control females.

1) *Resource defense polygyny*. Males defend resources essential to females. To

the degree that males can monopolize these resources, they can monopolize females.

When important resources are unevenly distributed or spatially clumped, certain males can defend areas containing a larger quantity or better quality of resource than others. If these resources are critical for female reproduction, then competition among males should revolve around subdividing and defending these resources. Female choice of mate should be influenced both by the quality of the defending male and the resources under his control (territory quality). The extent of polygyny will tend to increase with increasing variance in territory quality among the males of the population.

Among bird species where both sexes provide at least some parental care for offspring, females stand to lose if their mates take on additional females. Polygynous matings will be advantageous to the female only if the benefits gained by genetic access to a high-quality male or to the resources controlled by him more than offset her compensatory costs for the decreased contribution of the male in parental care. As stressed by Verner, Willson, and Orians, polygyny is expected when the distribution of resources is sufficiently irregular that a female mating with an already paired male on a superior-quality territory will have equal or better reproductive success than if she mated with an unpaired male occupying a poorer quality territory (8, 12, 13).

Studies on mating systems of North American passerines support these views. Resource defense polygyny is most prevalent in habitats with uneven resource distribution that results in a mosaic of male territories of different qualities (12, 28). Demographic studies further show that females forming polygynous pair bonds realize a reproductive success as great or greater than monogamous females (29-31). The specific resource attributes that comprise "territory quality" are expected to vary between species. Similarly, the extent of difference between territories that is sufficient to favor polygyny [the "polygyny threshold" of Orians (13)] will vary with such factors as the distribution of male parental care among his various mates, the total amount of parental care required by the young (for example, precocial versus altricial young), and the degree of dependency on the territory itself (for example, whether just for nesting or also for feeding).

Under certain environmental situations, the cost to the male of resource control or mate accumulation might be relatively low, while the potential benefit is extremely high. The cumulative advantage of multiple matings to the male could far outweigh the decreased reproductive success of individual females. In such conflict situations, females could be forced into assuming a larger fraction of the total parental care, even if it necessitated such long-term adaptations as reduced brood size or decreased growth rates of young. We suggest that some cases of male emancipation might best be viewed as an evolutionary result of a high environmental potential for polygamy rather than as an independently derived precursor to the evolution of polygamy.

Many passerine species believed to be monogamous show disparate parental investment by the male and female. (The female alone incubates the eggs while both sexes bring food for the nestlings.) Many of these species are sexually dimorphic, the male assuming a conspicuous breeding plumage while the female remains cryptically colored. The decreased male involvement in parental care preadapts such species to respond to slight changes in the environmental potential for polygamy by enabling them to become opportunistically or facultatively polygynous (31). We expect that many additional cases of such facultative polygyny will be discovered when individually marked populations are studied in areas where the feeding or nesting resources are limited or highly localized in space.

When male parental investment is minimal or nonexistent, a limited and clumped resource distribution can lead to extreme development of polygyny. In the orange-rumped honeyguide (*Indicator xanthonotus*), beeswax forms an essential part of the diet. Males do not provide defense of the nest site or food for the young, but they do maintain year-around territories at the locations of bee nests. These bee nests are found only on exposed cliffs; they are in short supply and a small proportion of the male population is able to control access to this resource. When females become sexually active, they enter the male territories and feed on the wax of the bee comb. Courtship is centered at the locations of bees' nests and copulation success is high for territory owners. One male was observed to copulate 46 times with at least 18 different females while nonterritorial males had minimal, if any, copulatory success. This species exemplifies an extreme degree of polygyny, based primarily on the ability of a small number of males to monopolize access to a critical resource (32).

Analogous, but less extreme, examples of resource defense polygyny have been reported for fiery-throated (*Pantherpe insignis*) and Anna hummingbirds (*Calypte anna*) (33, 34). Many aspects of hummingbird biology are closely linked to their habit of nectar feeding. When suitable flowers are sufficiently clumped and nectar production is high, territorial defense becomes economically feasible (34-36). Females incubate and rear the young alone, but they require a reliable nectar source to do so. In several species males allow females nesting within their territory to feed therein, but aggressively exclude all other hummingbirds (34, 35). The extent of polygyny presumably is determined by the distribution pattern of nectar-producing flowers which leads to differences in territory quality among males.

The yellow-bellied marmot (*Marmota flaviventris*) provides a mammalian example of resource defense polygyny. Overwintering sites constitute a monopolizable resource. Studies by Armitage and Downhower (37) suggest that areas of rock outcroppings provide the most suitable underground retreats for predator escape and for hibernation. Such sites are limited in number and are vigorously defended by males. High-quality locations are occupied by a colony including one or (rarely) a few dominant males, a variable number of mature females, plus young and juveniles. These units have been called "harems," but we feel they are better understood in the

context of resource defense polygyny. Male marmots achieve high reproductive success not by forcibly accumulating a harem of females but rather by economically monopolizing a resource that leads to female clumping.

2) *Female (or harem) defense polygyny.* Females are gregarious for reasons unrelated to reproduction. Their self-clumping tendencies facilitate direct monopolization by males.

If females themselves are defendable, we might expect males to forcibly accumulate females and to herd or maintain them under their jurisdiction by aggressively excluding all other males from the area. Such harem formation or "female defense polygyny" does occur, but usually in conjunction with male resource defense.

In many ungulates, females and young aggregate into small herds for part or all of the year. During times of parturition and sexual receptivity, these groups move into areas of preferred habitat (generally affording increased cover or abundant food, or both). Among the African impala (*Aepyceros melampus*) and waterbuck (*Kobus defassa*), for example, males at this time of year divide the habitat into defended territories. "The frequency with which a male has females in his territory is related to the amount of preferred habitat for that season that his territory contains. This may mean that a restricted number of territories may attract all the females for the duration of a limited mating season, and holders of those territories alone will contribute to breeding. In such circumstances it may be found that the marginal territories are held by young or old, rather than prime males" (38).

The reasons for female clumping may be totally unrelated to reproduction. Ungulates gain through increased predator detection and avoidance; other groups might gain through information exchange about locations of unpredictable food resources, increased foraging efficiency by observational learning, or cultural transmission of learned habits. But one result of the clumping is to increase the male potential for differential access and control of multiple mates.

Among many pinnipeds, females are sexually active shortly after giving birth (39). Females "haul out" onto land or ice to give birth, and the combination of female gregariousness, a shortage of suitable parturition locations, and a tendency to return annually to traditional areas gives rise to dense aggregations of sexually active females. By monopolizing access to these sites, dominant males are able to realize phenomenal

numbers of copulations. Competition among males is extremely intense and the resulting sexual selection presumably has led to the marked sexual dimorphism and intensity of aggressive behavior found in these species (10, 40).

The result of female clumping in pinnipeds may result in mates being directly defended as a resource (for example, in elephant seals) or may greatly increase the benefits derived from resource (site) defense (in many otariids such as the fur seals). In either case it leads to intensified sexual selection and increased imbalances in the reproductive success of different males in the population.

Female defense polygyny is rare among birds, being known for certain only in the greater rhea, a few tinamous, and several pheasants. The rhea provides one of the best known examples, but since males incubate and females sometimes mate with several males in succession, this system is described in the section on female selection and the evolution of polyandry.

In a number of pheasants and possibly in peafowl and chachalacas, female self-clumping occurs in combination with male resource defense. Female ring-necked pheasants (*Phasianus colchicus*) form groups that are attracted to certain male territories during the breeding season. Males directly defend female groups, and limited evidence suggests that larger harems and harems comprised of experienced females associate preferentially with males on prime territories (41). Membership in female groups is variable, and new individuals join throughout the summer. The result is a staggering of female breeding readiness (that is, asynchrony) enabling a single male to inseminate all sexually receptive members of the female group.

3) *Male dominance polygyny.* Males do not directly defend females or resources essential to females, but rather sort out among themselves their relative positions of dominance. Females choose males primarily on the basis of male status.

Communal displaying is frequent among species in which the male is totally emancipated from parental care and the environment provides little potential for resource or mate control. This is expected (i) when critical resources are superabundant but widely dispersed, or (ii) when they are sufficiently unpredictable in space and time as to be economically undefendable. It is also expected (iii) when resources or mates are clumped in a defensible pattern but the cost of successful defense is too high. High population density (increasing the number of

competitors) and increased intensity of competition (resulting from extreme limitation of resources or from strong skew in the OSR) can lead to situations where successful defense is energetically unfeasible. Cases where territorial defense is abandoned as the cost of defense increases have been documented in a variety of taxa (34, 36, 42). When access to females cannot be controlled through resource defense, male competition may take the form of direct male-male encounters, resulting in differential dominance relationships between the male members of the population.

Under what conditions should individual males aggregate and display communally? Such aggregations provide a forum for male-male competition and should increase the variance in reproductive performance among males. The copulation success of a high-ranking male would be increased by joining such a group; that of a low-ranking individual might decrease. Why then should subordinate males enter communally displaying groups?

If female movements or concentration areas are predictable, encounter rates would be high for males that position themselves in these areas, leading to localized concentrations of males. As males start to aggregate, their advertising and courtship signals become pooled, creating an enhanced stimulus situation that attracts females preferentially to larger aggregations (22, 43, 44). If isolated males have minimal chances of successfully attracting males, low-ranking males will be expected to adopt alternative, cryptic, or satellite strategies for obtaining females within the communal display areas rather than avoiding male aggregations altogether (45).

Females also benefit by male clumping since they can better compare a large number of potential mates in a minimum period of time. By relying on the male-male interactions to have done a preliminary sorting for them, females can preferentially select from a predetermined group of "tested" and "proven" males.

Other advantages have been suggested for male display aggregations. These include increased alertness and defense against potential predators as well as pooling of information about locations of patchy or ephemeral food resources (24, 46, 47). We believe that these are secondary advantages, accruing after the development of group displaying and being of greatest importance in species that remain sexually active and aggregated for long periods of time.

The form of male dominance polygyny will depend on the degree of synchrony

of sexual activity among females of the population. Where females are highly synchronized and converge at the male aggregation within a short period of time, a highly promiscuous, "explosive" breeding situation will occur. This is typical of many singing insects and chorusing amphibians (16, 48). As was discussed previously, the synchrony of the females restricts the potential for individual males to monopolize matings. Breeding activity is frenzied, but the operational sex ratio should not be highly skewed and sexual selection should not be intense (49).

If the females of a population are relatively asynchronous in their periods of sexual receptivity (50), the operational sex ratio becomes increasingly skewed and sexual selection intensifies. Males generally remain active for the duration of the population's breeding season. The intensity of male-male competition, together with the longer duration of the mating period, result in the establishment of stable dominance or position effects among the advertising males. The result is an organized aggregation usually referred to as a *lek*.

A lek is defined as a communal display area where males congregate for the sole purpose of attracting and courting females and to which females come for mating (51). Males jockey for a status or position that conveys maximal attractiveness to females. Frequently, central positions are occupied by older, more dominant, males that achieve a disproportionate share of female copulations (22, 44, 46, 47, 52, 53). Females visit the lek when sexually receptive and "sample" numerous males before selecting a mate. After copulation, the female leaves the display area and proceeds, on her own, to rear the young.

Lek mating systems have been described for several species of insects, mouth breeding fish, bullfrogs, a scattering of mammals and approximately 30 species of birds (54-56). We hypothesize that all cases will be typified by a fairly long breeding season, a heavily skewed operational sex ratio, and by the inability of individual males to economically control or monopolize the resources essential for female acquisition.

Male Incubation, Female Emancipation

Among most animals, female parental investment greatly exceeds that of males (3, 5-7). Consequently, most cases of polygamy involve male emancipation and the development of polygyny. However, a significant portion of avian spe-

cies share rather equally in parental care duties, including incubation of the eggs. And, in a very small percentage of species, males have assumed the full burden of incubation and brood rearing, emancipating the female and increasing the possibility of her mating repeatedly.

Males should assume the bulk of parental care only when their individual fitness is increased through such an action. This might occur in two general situations: (i) when the future physical condition of the female is of direct importance to the male (discussed under Monogamy), and (ii) when the lack of dependability of breeding conditions places a premium on female ability to produce additional or replacement clutches for the male. Lack of dependability can result either (i) from great fluctuations in environmental suitability for breeding or (ii) from very low success rates of reproductive attempts (caused, for example, by high predation rates).

From a female's viewpoint, the principal advantage of male incubation is a lessening of her metabolic burden, freeing her to devote increased time and energy to replenish nutrient and mineral reserves drained through egg production (57, 58).

From the male's point of view, female emancipation can be viewed as an insurance strategy—when the failure rate is sufficiently high, the cost involved in incubation may be less, on average, than the gain accrued by having a mate physiologically able to rapidly produce new clutches of eggs.

Complete male parental care is most likely to develop in groups with slight to moderate parental care needs (precocial young) and a phylogenetic history of shared incubation. It should also be more prevalent among determinate than indeterminate egg layers (59).

Rapid multiple clutch polygamy. Both sexes have substantial but relatively equal opportunity for increasing fitness through multiple breedings in rapid succession. Simultaneous brooding of two clutches usually occurs.

In some shorebirds and galliforms, the female may lay a first clutch that is incubated solely by the male, and a second clutch that she herself incubates. When conditions are good, the result is a doubling of reproductive potential with only a minimal increase in breeding time; when conditions are poor, the result is an increased ability to produce replacement clutches.

Species exemplifying such rapid multiple clutch polygamy include red-legged partridge (*Alectoris rufa*), sand-

erling (*Calidris alba*), mountain plover (*Charadrius montanus*), and Temminck's stint (*Calidris temminckii*) (58, 60, 61). All are ground-nesting birds whose precocial young suffer moderate to extremely high predation losses. The mountain plover inhabits marginal short-grass prairies and is further subjected to severe fluctuations in environmental suitability due to great variations in rainfall. The short arctic breeding season of the sanderling is characterized by great fluctuations in suitability for breeding. The California quail (*Lophortyx californicum*), which, at least occasionally, practices this mating system (62) also is subject both to severe fluctuations in environmental suitability and to extreme predation (63).

Rapid multiple clutch polygamy is probably more common than is currently known, especially among shorebirds. It may also prove to be common among small phasianids, where males of a number of species are known to sometimes incubate or brood young (or both) without female help (64).

Advantages of male incubation to males are maximized (i) when females preferentially return to their original mates to lay a subsequent clutch of eggs, and (ii) when males remain sexually active for as long as possible during or before the initiation of incubation.

Preferential treatment by females for their original mates has been investigated in red-legged partridge, mountain plover, and Temminck's stint [see also (77)]. In the former two species, females lay second clutches in the territories of their original mate and copulate with him providing he is sexually active. In Temminck's stint, on the other hand, mate fidelity between first and second clutch is reported from one locality (U.S.S.R.) but not from another (Finland) (58, 65).

Males exhibiting rapid multiple clutch polygamy continue to actively court additional females for a period following the completion of the first clutch. Such behavior is rarely reported among species that share incubation. In the mountain plover, this is accomplished by maintaining sexual activity during part of incubation, while in the Temminck's stint, sanderling, and red-legged partridge, males frequently delay incubation, leaving the first clutch untended for as long as 6 to 12 days while continuing to court additional females.

As a result of this behavior, rapid multiple clutch polygamy can lead to a doubling of reproductive potential without producing a strong skew in the operational sex ratio.

Female Sexual Selection and the Evolution of Polyandry

Male incubation preadapts a species for possible evolution toward polyandry. Emancipation allows the female opportunity to increase her fitness through continued production of multiple clutches. Female fitness can be increased only to the degree that males are sexually receptive and available to assume incubation of these additional clutches.

Since an individual male is not able to maintain sexual activity indefinitely during incubation, males sitting on eggs normally must be considered sexually "unavailable." To the degree that females can produce more clutches than can be serviced by males (caused either by high production rates of females or low failure rates of existent clutches), the operational sex ratio will become skewed with a shortage of males. These conditions lead to increased intrasexual competition among females for access to available males. The degree to which polyandry will develop then depends on the intensity of female sexual selection and the environmental potential for monopolization of mates (in this case female monopolization of males).

True polyandry is extremely rare among birds, being found in less than 1 percent of the species studied to date (24, 66). Most documented cases are restricted to the avian orders Gruiformes and Charadriiformes. Comprehensive studies involving individually marked, wild birds have been conducted on only three species, the spotted sandpiper (*Actitis macularia*), the American jacana (*Jacana spinosa*), and the northern phalarope (*Phalaropus lobatus*). The mating systems of the first two are considered resource defense polyandry, while the latter represents an explosive type of female access polyandry. There are no documented examples of female dominance polyandry with stable organization analogous to that typical of leks.

1) **Resource defense polyandry.** Females compete for and defend resources essential to males. To the degree that these resources are clumped and monopolizable, females can monopolize multiple males.

In the spotted sandpiper, males often perform all incubation and brood care. Incubation begins during egg laying, and by the time of clutch completion males are sexually disinterested. Predation losses are extremely high (67, 68), and the ability of females to rapidly lay replacement clutches is impressive (69).

Spotted sandpipers breed throughout

the mid- and lower latitudes of North America. Under favorable conditions, the long breeding season, coupled with a high egg-producing ability, results in the female reproductive output exceeding the availability of males to assume incubation responsibilities. Because of male incubation, the operational sex ratio is skewed in favor of females, competition exists for available mates, and female sexual selection is moderately strong.

The extent of polyandry seems to be determined, in part, by the "availability" of males and by the quality and the spatial dispersion pattern of essential resources. In one population in Minnesota, male density was low and breeding females on average mated with 1.17 males. In contrast, in dense populations in Minnesota and New York, where birds were concentrated on patches of high-quality habitat, successful females averaged 2.4 mates (67, 70). High nest predation and asynchronous patterns of arrival of males both influenced the incidence of polyandry through their effect on increasing the "availability" of males to accept additional clutches of eggs (71).

The American jacana exhibits the most extreme polyandry known. In parts of Costa Rica their breeding habitat is severely limited, and only a small fraction of either the male or female population is believed to breed in any one year. The few suitable ponds or lagoons are subdivided into small territories by males. Females control "super-territories," often encompassing the nesting areas of several males. Females frequently have multiple mates incubating clutches simultaneously. Predation of nests is high, and females readily provide replacement clutches for their males, who perform most parental care duties (66, 72).

As the degree of female sexual selection increases, so too does behavioral and size dimorphism. In spotted sandpipers the female is 25 percent larger than the male, both sexes defend territories although the female is dominant, and the female frequently shares incubation with her last mate of the year. In the jacana, breeding females weigh 50 to 75 percent more than males, are totally dominant over males in aggressive interactions, and provide minimal parental care for eggs or young. In essence, female jacanas are specialists in egg production.

2) *Male defense or "harem" polyandry?* Several species of tinamous as well as the greater rhea exhibit unique mating systems that have been termed harem polyandry. On the basis of under-

lying resource control, we interpret these cases not as polyandry but rather as specialized cases of polygyny coupled with male incubation.

Tinamous are primitive, ground-nesting birds that inhabit the neotropics. Most defend all-purpose territories from which they advertise vocally. Females of some species are loosely gregarious and travel between male territories. Males may have pair bonds simultaneously with several females who lay a communal clutch of eggs. Males then perform all incubation and care for the precocial young without female help. Competition for females appears to exist and indirect evidence suggests a considerable differential in reproductive success of neighboring males (73-76).

Male tinamous recoup the cost of incubation in three ways. (i) Females freed from parental duties are better able to provide replacement clutches when predation does occur (73, 77). (ii) If several females lay communally, the male obtains a compound clutch rapidly, thereby minimizing the times during which eggs are unattended and, hence, exposed to extreme predation pressure. (iii) A male can inseminate many females and increase his reproductive success if he is able to attract groups of synchronized females. Variance in clutch size is large among tinamous, and individual male *Tinamus* and *Rhynchotus* have been seen incubating as many as 12 eggs (76, 78).

Once incubation has begun, a male ceases calling and becomes sexually unavailable. Females maximize their fitness by continuing to produce additional eggs, but they now must seek out additional males that are sexually active and able to accept a new clutch. Because of the long, tropical breeding season, females are frequently able to breed with a number of males in sequence. The wandering of females between neighboring males can be viewed as a strategy that is maximally advantageous to both sexes.

Female gregariousness is highly pronounced in the greater rhea (79). At the onset of the breeding season, older males compete for possession of a self-clumping group of females. Males defend no resources other than females and nest sites. Subordinate males are physically driven from the vicinity of females. In this way, a dominant male gains access to or monopolizes a group of females that then communally lay a clutch of eggs. The male assumes full incubation. The emancipated females continue to produce eggs and move on, laying a communal clutch for a second, third, or even

fourth male. Many aspects of this breeding situation, including the reasons underlying male incubation, are presumed to be basically similar to those in the tinamous.

There is little evidence for sexual selection among females in the tinamous and rhea. Females are subordinate to males and sexual dimorphism is slight. All of this suggests a specialized form of resource defense polygyny (tinamous) or mate-defense polygyny (rheas) rather than an ecologically based polyandry.

3) *Female access polyandry.* Females do not directly defend resources essential to males but, through interactions among themselves, may limit access to males. Females sometimes defend males. This defense is longer than the reciprocal defense noted in male dominance polygyny since a female must remain with a male until a clutch is completed and incubation begins.

Phalaropes are highly specialized shorebirds that breed at mid- and high latitudes. They feed on a relatively small variety of insects and aquatic invertebrates in or adjacent to wetlands. They utilize a narrow range of nest sites located near these feeding areas. Because of the extremely unpredictable and ephemeral nature of their food resource, courting areas shift from year to year and even from week to week. Males perform all parental care, but there is no opportunity for a stable resource defense.

In all three phalarope species, males and females congregate at bodies of water where they feed, display, and copulate. In our opinion, the mating system is most analogous to an explosive breeding assemblage with the difference that the females can be the limited, and males the limiting, sex. We predict that the operational sex ratio will become increasingly skewed and polyandry will occur when (i) males arrive asynchronously on the breeding grounds, (ii) the breeding season is sufficiently long to allow re-nesting attempts, and (iii) nesting failure is frequent, increasing the importance of such replacement nestings. Female competition then occurs, and polyandry has been reported (80). Pair bonds are brief, and females attempt to maximize fitness through repeated matings (81). At high population levels, female-female interactions become severe and some members of these ephemeral aggregations are prevented from breeding (82). To the degree that certain females can influence the access of others to males during the period of copulation and nest initiation, the mating system can be called female access polyandry (83).

Plasticity of Mating Systems

Throughout this article, we have stressed that ecological variables influence or constrain the intensity of sexual selection. This, in turn, profoundly influences the form of the resulting mating system. Certain individual strategies for resource or mate control are adaptive to one set of ecological conditions, but not necessarily to another. As stated by Emlen "... ecological parameters impose limits on the range of types of social organization that will be adaptive. With differences in the dispersion of a critical resource, the availability of mates, or other factors, optimal social strategies shift, resulting in a fine tuning of social organization to ecological constraints" (84).

If the productivity or the spatial or temporal distribution of a critical resource changes from year to year or from area to area, we should expect corresponding changes in the environmental potential for polygamy. Similarly, if the energetic cost of resource or mate monopolization changes as a result of changes in population density, length of breeding season, and the like, we should expect corresponding changes in the ability to take advantage of the environmental potential for polygamy.

Considerable lability in mating systems is thus expected between different populations of a given species in different environmental or density situations. The form of the plasticity, as well as the conditions under which it should occur, should in themselves be predictable on the basis of the ecological framework presented here.

Shifts from monogamy to facultative polygyny (resource defense polygyny) and shifts in degree of polygyny have been documented in a wide variety of avian species—invariably in accordance with ecological predictions (8, 30, 31, 85). Similar variability in the amount of polyandry and its dependence on male availability and degree of resource monopolization has already been mentioned among spotted sandpipers. Predictable shifts along the continuum of monogamy (with shared incubation)—rapid multiple clutch polygamy—resource defense polyandry have also been suggested in both sanderlings and spotted sandpipers (61, 67).

Several organisms, including dragonflies, bullfrogs, turkeys, puku, topi, and Uganda kob, are known to exhibit lekking behavior at high population densities, but shift to a resource defense polygyny or mate defense polygyny at low densities (53, 54, 86, 87). We hypoth-

esize that these shifts are due to the changing energetic costs of mate or resource defense associated with density-related changes in the intensity of intrasexual competition (87).

Until recently, many field biologists have worked under a preconception that species specificity was a characteristic not only of courtship behavior but of mating systems as well. We are now coming to realize that variability in social organization, including mating systems, is widespread. The ecological model presented in this article should provide a basis for generating testable predictions concerning the expected form of such mating system variability.

Summary

We have attempted to provide an ecological framework for understanding and predicting the forms of animal mating systems. The underlying assumption is that intrasexual interactions associated with mating are basically competitive. An individual member of the limited sex is expected to maximize its inclusive fitness by attempting to control access to mates of the limiting sex. The degree to which this is possible depends on the costs and benefits associated with such control. Certain environmental factors, particularly the spatial dispersion pattern of key resources and the temporal availability of receptive mates, are important determinants of these costs and benefits. The greater the potential for individuals to monopolize resources or mates, the greater the intensity of sexual selection and the greater the environmental potential for polygamy.

The precise form of the mating system will depend on which sex is limiting and on the manner and the degree to which the limited sex controls the resource base or monopolizes mates (or both).

An ecological categorization of mating systems is presented (Table 1) that allows a better understanding of the selective forces shaping one mating system over another. Within this ecological framework, specific examples are discussed ranging from the occurrence of leks to the evolution of polyandry.

References and Notes

1. The term "mating system" of a population refers to the general behavioral strategy employed in obtaining mates. It encompasses such features as: (i) the number of mates acquired, (ii) the manner of mate acquisition, (iii) the presence and characteristics of any pair bonds, and (iv) the patterns of parental care provided by each sex.
2. C. Darwin, *The Descent of Man, and Selection in Relation to Sex* (Appleton, New York, 1871).
3. R. A. Fisher, *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930).
4. J. S. Huxley, in *Evolution: Essays on Aspects of Evolutionary Biology Presented to Professor E. S. Goodrich on His 70th Birthday*, G. R. DeBeer, Ed. (Clarendon, Oxford, 1938), pp. 11-42; *Am. Nat.* 72, 416 (1938); J. M. Smith, in *A Century of Darwin*, S. A. Barnett, Ed. (Heinemann, London, 1958), pp. 231-244; W. D. Hamilton, *Science* 156, 477 (1967); R. L. Trivers, *Am. Zool.* 14, 249 (1974).
5. G. C. Williams, *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought* (Princeton Univ. Press, Princeton, N.J., 1966).
6. ———, *Sex and Evolution* (Princeton Univ. Press, Princeton, N.J., 1975).
7. R. L. Trivers, in *Sexual Selection and the Descent of Man, 1871-1971*, B. Campbell, Ed. (Aldine, Chicago, 1972), pp. 136-179.
8. J. Verner, *Evolution* 18, 252 (1964).
9. J. H. Crook, *Behaviour* (Suppl. 10), (1964), p. 1; R. K. Selander, *Am. Nat.* 99, 129 (1965).
10. G. A. Bartholomew, *Evolution* 24, 546 (1970).
11. J. H. Crook, *Symp. Zool. Soc. London* 14, 181 (1965).
12. J. Verner and M. Willson, *Ecology* 47, 143 (1966).
13. G. H. Orians, *Am. Nat.* 103, 589 (1969).
14. R. K. Selander, in *Sexual Selection and the Descent of Man, 1871-1971*, B. Campbell, Ed. (Aldine, Chicago, 1972), p. 180.
15. F. A. Pitelka, R. T. Holmes, S. F. MacLean, Jr., *Am. Zool.* 14, 185 (1974).
16. R. D. Alexander, in *Insects, Science, and Society*, D. Pimental, Ed. (Academic Press, New York, 1975), p. 35.
17. We intend this to encompass the concept of inclusive fitness, as developed by W. D. Hamilton [*J. Theor. Biol.* 7, 1 (1964); *ibid.*, p. 17] and M. J. West-Eberhard [*Q. Rev. Biol.* 50, 1 (1975)], among others.
18. "r"-selected species generally tend to maximize quantity rather than quality of propagules. Parental care is relatively unimportant, if present at all. Hence the form of the mating system is less influenced by minor differences in resource control or mate assistance in the rearing of young.
19. The effects of competition or selfish behavior may not be fully expressed among long-lived animals that spend much of their lives in small, closed groups. This may be due to (i) the effects of kin selection, whereby selfish behavior might reduce the actor's inclusive fitness as a result of its negative effects on close relatives within the group, (ii) reciprocal interactions including reciprocal altruism, sensu R. L. Trivers [*Q. Rev. Biol.* 46, 35 (1971)], or (iii) the general advantages of group living being sufficiently great that selfish behavior is not carried to the point where it results in subordinate individuals leaving the group, sensu R. D. Alexander [*Ann. Rev. Ecol. Syst.* 4, 325 (1974)].
20. J. L. Brown, *Wilson Bull.* 76, 160 (1964).
21. In species where one sex provides the bulk of parental investment, this sex becomes temporarily "unavailable" for further reproductive attempts (for example, during incubation of eggs or lactation for young). This can cause a skew in the operational sex ratio that will lead to intensified competition for access to the few remaining partners of the limiting sex.
22. S. T. Emlen, *Behav. Ecol. Sociobiol.* 1, 283 (1976).
23. The operational sex ratio is measured by censusing the number of sexually active males and fertilizable females on different dates throughout the breeding season and combining these values to arrive at an average ratio. Direct measurement of OSR in the field can be difficult, since intrasexual competition often takes the form of exclusion of a portion of the population from the breeding area. Censuses of the breeding animals then would underrepresent the number of the limited sex. The problems of empirically measuring the OSR, however, are separate from, and should not detract from, its heuristic value in understanding and predicting the degree of polygamy to be expected under different environmental situations.
24. D. Lack, *Ecological Adaptations for Breeding in Birds* (Methuen, London, 1968).
25. J. F. Eisenberg, *Handb. Zool.* 10, 1 (1966).
26. Certain mammalian carnivores that hunt dispersed prey provide examples of species with low EPP. Such carnivores are rarely polygamous. Where males play a role in parental care, monogamy frequently results (among many canids and several felids and mustelids).
27. J. C. Coulson, *J. Anim. Ecol.* 35, 269 (1966); J. A. Mills, *ibid.* 42, 147 (1973); J. W. F. Davis, *ibid.* 45, 531 (1976).
28. J. Verner and M. F. Willson, *Ornithol. Monogr.* 9, 1 (1969).

29. J. L. Zimmerman, *Auk* 83, 534 (1966); S. V. Goddard and V. V. Board, *Wilson Bull.* 79, 283 (1967); C. H. Holm, *Ecology* 54, 356 (1973); J. G. Wittenberger, thesis, University of California, Davis (1976). For contradictory data, see S. T. Martin, thesis, Oregon State University (1971).
30. R. A. Howard, Jr., thesis, Cornell University (1977).
31. M. Carey and V. Nolan, Jr., *Science* 190, 1296 (1975).
32. E. W. Cronin, Jr., and P. W. Sherman, *Living Bird* 15, 5 (1976).
33. L. L. Wolf and F. G. Stiles, *Evolution* 24, 759 (1970).
34. F. G. Stiles, *Univ. Calif. Berkeley Publ. Zool.* 97, 1 (1973).
35. L. L. Wolf, *Auk* 86, 490 (1969); F. G. Stiles and L. L. Wolf, *ibid.* 87, 467 (1970).
36. See also F. B. Gill and L. L. Wolf, *Ecology* 56, 333 (1975); F. L. Carpenter and R. E. MacMillen, *Science* 194, 639 (1976).
37. K. B. Armitage, *Anim. Behav.* 10, 319 (1962); *ibid.* 13, 59 (1965); *J. Zool.* 172, 233 (1974); J. F. Downhower and K. B. Armitage, *Am. Nat.* 105, 355 (1971); D. C. Andersen, K. B. Armitage, R. S. Hoffman, *Ecology*, 57, 552 (1976).
38. The quotation is from P. J. Jarman [*Behaviour* 48, 215 (1974)], p. 248; the same point is made also by C. A. Spinage [*J. Zool.* 159, 329 (1969)].
39. In phocid seals, estrus occurs near the end of the lactation period and lasts for several days or longer. Most otariid seals exhibit a very brief estrus, early in the lactation period.
40. G. A. Bartholomew, *Univ. Calif. Berkeley Publ. Zool.* 47, 369 (1952); R. S. Peterson and G. A. Bartholomew, *Am. Soc. Mammal.* (Special Publ.) 1, 1 (1967); I. A. McLaren, *Ecology* 48, 104 (1967); R. S. Peterson, in *The Behavior and Physiology of Pinnipeds*, R. J. Harrison et al., Eds. (Appleton-Century-Crofts, New York, 1968), pp. 3-53; B. J. LeBoeuf and R. S. Peterson, *Science* 163, 91 (1969); B. J. LeBoeuf, *Am. Zool.* 14, 163 (1974); E. H. Miller, *Rapp. P. V. Réun. Cons. Int. Explor. Mer.* 169, 170 (1975).
41. R. D. Taber, *Condor* 51, 153 (1949).
42. D. E. Davis, *Anim. Behav.* 6, 207 (1959); R. D. Alexander, *Behaviour* 17, 130 (1961); J. I. Magnuson, *Can. J. Zool.* 40, 313 (1962); V. I. Paganin, *Ann. Zool. Fenn.* 3, 40 (1966); A. Zahavi, *Ibis* 113, 203 (1971).
43. D. Lack, *Br. Birds* 32, 290 (1939); F. Hamerstrom and F. Hamerstrom, *Proc. Int. Ornithol. Congr.* 11, 459 (1955); D. W. Snow, *ibid.* 13, 553 (1963); G. Morris, abstract, meeting of the Animal Behavior Society, Boulder, 1976; personal communication.
44. A. Lill, *Z. Tierpsychol.* 36, 1 (1974); *Adv. Ethol.* 10, 1 (1976).
45. M. Gadgil, *Am. Nat.* 106, 576 (1972).
46. I. Kivisto, *Finn. Game Res.* 26, 1 (1965).
47. J. P. Kruijt, G. J. DeVos, I. Bossema, *Proc. Int. Ornith. Congr.* 15, 339 (1972).
48. P. G. Pearson, *Ecol. Monogr.* 25, 233 (1955); W. F. Blair, *Tex. J. Sci.* 8, 87 (1956); C. M. Bogert, Commentary for recording of "Sound of North American Frogs" (Folkways Records, FX-6166), p. 1 (1958); in *Animal Sounds and Communication*, W. E. Lanyon and W. N. Tavolga, Eds. (Publ. 7, American Institute of Biological Sciences, Washington, D.C., 1960), p. 137; R. D. Alexander, *ibid.*, p. 38; and T. E. Moore, *Misc. Publ. Mus. Zool. Univ. Mich.* 121, 1 (1962); K. Wells, *Anim. Behav.*, in press.
49. To the extent that certain males limit the access of other males to females, some polygynous matings would occur. This situation could be called "male access polygyny" and would be analogous to the mating situation described for phalaropes under the section on female access polyandry.
50. Degree of asynchrony in a female population should be characterized relative to the time necessary for a male to attract, court, and service an individual female of the species in question—not in absolute units such as hours, days, or weeks. Asynchrony generally results from differences in the times at which individual females first become sexually active; but it may be augmented by high failure rates of early nests, causing females to return to the "sexually active" pool at later dates.
51. E. O. Wilson, *Sociobiology, The New Synthesis* (Harvard Univ. Press, Cambridge, Mass., 1975).
52. J. W. Scott, *Auk* 59, 477 (1942); H. G. Lumsden, *Ont. Dep. Lands For. Res. Rep. No. 66* (1965); *ibid.* No. 83 (1968); J. P. Kruijt and J. A. Hogan, *Ardea* 55, 203 (1967); R. J. Robel, *Proc. Am. Philos. Soc.* 111, 109 (1967); *J. Wildl. Manage.* 34, 306 (1970); F. Hamerstrom and F. Hamerstrom, *Wisc. Dept. Nat. Resour. Tech. Bull.* 64, 1 (1973); R. H. Wiley, *Anim. Behav. Monogr.* 6, 85 (1973); J. M. Shepard, *Living Bird* 14, 87 (1975); O. R. Floody and A. P. Arnold, *Z. Tierpsychol.* 37, 192 (1975).
53. H. D. Buechner and R. Schloeth, *Z. Tierpsychol.* 22, 209 (1965).
54. For discussions regarding insects, see H. T. Spieth, *Evol. Biol.* 2, 157 (1968); *Annu. Rev. Entomol.* 19, 385 (1974); R. Lederhouse, thesis, Cornell University (1977); and (55). For discussion regarding fish, see P. V. Loiselle and G. W. Barlow, in *Contributions in Behavior*, E. Reese, Ed. (Wiley-Interscience, New York, in press). For discussion regarding amphibians, see (22). For discussion regarding birds, see (24).
55. P. J. Campanella and L. L. Wolf, *Behaviour* 51, 49 (1974).
56. For discussion regarding mammals, see N. Monfort-Braham, *Z. Tierpsychol.* 39, 332 (1975); R. H. Schuster, *Science* 192, 1240 (1976); J. Bradbury, in *Biology of Bats*, W. Wimsatt, Ed. (Academic Press, New York, in press), vol. 3; and (53).
57. A clutch of eggs in many shorebirds may weigh 80 to 90 percent of the female's body weight and contain up to 200 percent of the female's total calcium [S. F. MacLean, Jr., *Ibis* 116, 552 (1974); L. W. Oring, personal communication].
58. O. Hilden, *Ornis Fenn.* 52, 117 (1975).
59. One means for varying reproductive output is to vary clutch size in accordance with resources. For those birds that are determinate layers (species in which the number of offspring are fixed by ovulation of a predetermined number of ova), this strategy is unattainable. Females of such species (including shorebirds) can adjust reproductive effort to environmental conditions only by altering the number of total reproductive attempts—through the laying of complete additional clutches.
60. D. Goodwin, *Ibis* 95, 581 (1953); D. Jenkins, *Bird Study* 4, 97 (1957); D. F. Parmelee, *Living Bird* 9, 97 (1970); W. Graul, *ibid.* 12, 69 (1973); but see also M. W. Pienkowski and G. H. Green, *Br. Birds* 69, 165 (1976).
61. D. F. Parmelee and R. B. Payne, *Ibis* 115, 218 (1973).
62. D. McLean, *Calif. Div. Fish Game Fish. Bull. No. 2*, (1930); I. I. McMillan, *J. Wildl. Manage.* 28, 702 (1964); W. Francis, *Condor* 67, 541 (1965); R. Anthony, *ibid.* 72, 276 (1970).
63. E. L. Sumner, Jr., *Calif. Fish Game* 21, 167 (1935); *ibid.*, p. 277; B. Glading, *ibid.* 24, 318 (1938).
64. H. L. Stoddard, *The Bobwhite Quail: Its Habits, Preservation, and Increase* (Scribner, New York, 1931); G. Gullion, *Condor* 58, 232 (1956); O. Wallmo, *Texas Game and Fish Commission*, Austin (1956); S. Schemnitz, *Wildl. Monogr.* 8, 1 (1961); A. H. Miller and R. C. Stebbins, *The Lives of Desert Animals in Joshua Tree National Monument* (Univ. of California Press, Berkeley, 1964).
65. V. D. Kokhanov, in *Fauna and Ecology of Waders*, V. E. Flint, Ed. (Moscow Univ. Press, Moscow, 1973), vol. 1, p. 66.
66. D. A. Jenni, *Am. Zool.* 14, 129 (1974).
67. L. W. Oring and M. Knudsen, *Living Bird* 11, 59 (1972).
68. L. W. Oring, unpublished observations.
69. Interclutch intervals can be as short as 3 days. As many as 20 eggs (totaling more than 400 percent of the female's body weight) have been laid by one female in a period of 42 days (68).
70. H. Hays, *Living Bird* 11, 43 (1972).
71. On an island in Minnesota, all nests were lost to predators in 1975. Males kept recycling into sexual activity and successful females laid for an average of 2.0 males, while males bred with an average of 1.5 females. In 1976, predators were artificially removed (making established males less available), but a number of late-arriving new males entered the breeding population asynchronously (increasing availability of new males). As a result, all males bred with only one female, but each female laid for an average of 2.0 males and copulated with an average of 3.7 males. In each year, the high incidence of polyandry resulted from the high and prolonged availability of males to accept new clutches—but the reasons underlying this availability differed in the two seasons (68).
72. D. A. Jenni and G. Collier, *Auk* 89, 743 (1972).
73. E. Schafer, *J. Ornithol.* 95, 219 (1954).
74. A. K. Pearson and O. P. Pearson, *Auk* 72, 113 (1955); D. A. Lancaster, *Condor* 66, 165 (1964A); *Bull. Am. Mus. Nat. Hist.* 127, 273 (1964); personal communication.
75. D. A. Lancaster, *Condor* 66, 253 (1964).
76. S. E. Weeks, *Zoologica (N.Y.)* 44, 13 (1973).
77. Lancaster (75) describes a situation in Bocard's tinamou in which two females demonstrated a preference for a particular male. The females left their first mate incubating, after laying a communal clutch of eggs, and paired with a new male. But when the first mate's nest was depredated and he began vocal advertising, they broke their new bond and rejoined the original male.
78. W. Beebe, *Zoologica (N.Y.)* 6, 195 (1925).
79. D. F. Bruning, *Living Bird* 13, 251 (1974).
80. O. Hilden and S. Vuolanto, *Ornis Fenn.* 49, 57 (1972); L. Ränner, *Fauna Flora* 67, 135 (1972); H. Mayfield, unpublished manuscript.
81. In both red and northern phalaropes the females that are successful in obtaining second mates (becoming polyandrous) are the first females to complete their initial clutches of eggs.
82. A. A. Kistchinski, *Ibis* 117, 285 (1975).
83. The nesting habitat of phalaropes comprises prairie wetlands and arctic tundra. Environmental and ecological conditions vary greatly from year to year in these habitats, and we should expect a corresponding amount of variability in the degree of polyandry shown by the phalaropes. (See section on plasticity of mating systems.)
84. S. T. Emlen, *Science* 192, 736 (1976).
85. E. A. Armstrong, *The Wren* (Collins, London, 1955), pp. 102-109; N. A. Case and O. H. Hewitt, *Living Bird* 2, 7 (1963); H. W. Kale II, *Publ. Nuttall Ornithol. Club* 5, 1 (1965).
86. W. Leuthold, *Behaviour* 27, 215 (1966); D. M. Smith, personal communication.
87. S. T. Emlen, in preparation.
88. We thank members of Cornell's Graduate Seminar in Social Behavior who offered both insights and criticisms, and D. A. Jenni who provided intellectual stimulation throughout his semester as a visiting scientist in Ithaca. D. Lancaster and the Cornell Laboratory of Ornithology aided this work through discussions and provision of space to L. W. O. H. F. Mayfield, E. H. Miller, G. H. Orrians, and F. A. Pitelka kindly offered helpful comments and criticisms on the manuscript. Supported, in part, by NSF grants BMS-75-18905 and BMS-76-81921 to S. T. E. and GB-42255 to L. W. O.

A framework for understanding semi-permeable barrier effects on migratory ungulates

Hall Sawyer^{1*}, Matthew J. Kauffman², Arthur D. Middleton^{3,4}, Thomas A. Morrison³, Ryan M. Nielson¹ and Teal B. Wyckoff^{3,5}

¹Western Ecosystems Technology, Inc., 200 South 2nd St., Suite B, Laramie, WY, 82070, USA; ²US Geological Survey, Wyoming Cooperative Fish and Wildlife Research Unit, Department of Zoology and Physiology, University of Wyoming, Laramie, WY, 82071, USA; ³Wyoming Cooperative Fish and Wildlife Research Unit, Department of Zoology and Physiology, University of Wyoming, Laramie, WY, 82071, USA; ⁴Program in Ecology, University of Wyoming, Laramie, WY, 82071, USA; and ⁵Wyoming Geographic Information Science Center, University of Wyoming, Laramie, WY, 82071, USA

Summary

1. Impermeable barriers to migration can greatly constrain the set of possible routes and ranges used by migrating animals. For ungulates, however, many forms of development are semi-permeable, and making informed management decisions about their potential impacts to the persistence of migration routes is difficult because our knowledge of how semi-permeable barriers affect migratory behaviour and function is limited.

2. Here, we propose a general framework to advance the understanding of barrier effects on ungulate migration by emphasizing the need to (i) quantify potential barriers in terms that allow behavioural thresholds to be considered, (ii) identify and measure behavioural responses to semi-permeable barriers and (iii) consider the functional attributes of the migratory landscape (e.g. stopovers) and how the benefits of migration might be reduced by behavioural changes.

3. We used global position system (GPS) data collected from two subpopulations of mule deer *Odocoileus hemionus* to evaluate how different levels of gas development influenced migratory behaviour, including movement rates and stopover use at the individual level, and intensity of use and width of migration route at the population level. We then characterized the functional landscape of migration routes as either stopover habitat or movement corridors and examined how the observed behavioural changes affected the functionality of the migration route in terms of stopover use.

4. We found migratory behaviour to vary with development intensity. Our results suggest that mule deer can migrate through moderate levels of development without any noticeable effects on migratory behaviour. However, in areas with more intensive development, animals often detoured from established routes, increased their rate of movement and reduced stopover use, while the overall use and width of migration routes decreased.

5. *Synthesis and applications.* In contrast to impermeable barriers that impede animal movement, semi-permeable barriers allow animals to maintain connectivity between their seasonal ranges. Our results identify the mechanisms (e.g. detouring, increased movement rates, reduced stopover use) by which semi-permeable barriers affect the functionality of ungulate migration routes and emphasize that the management of semi-permeable barriers may play a key role in the conservation of migratory ungulate populations.

Key-words: Brownian bridge movement model, connectivity, migration routes, mule deer, stopovers

*Correspondence author. E-mail: hsawyer@west-inc.com

Introduction

Migration is unique among animal movement strategies because of the impressive distances that animals travel, the predictability of their return and, for many species, the sheer number of individuals involved (Dingle 1996; Milner-Gulland, Fryxell & Sinclair 2011). Migratory ungulates have received much attention because of their role as drivers of ecosystem processes (McNaughton 1985; Hobbs 1996), their value to humans as harvestable resources (Vors & Boyce 2009) and their potential as flagship species for landscape-level conservation (Thirgood *et al.* 2004). Recent global declines in the abundance and distribution of migratory ungulates (Berger 2004; Bolger *et al.* 2008; Harris *et al.* 2009) underscore the need to better understand the consequences of disruptions to migratory behaviour. Declines in migratory ungulates have been clearly linked to excessive harvest and land-use changes (e.g. agricultural development) on seasonal ranges (Bolger *et al.* 2008), but neither overharvest nor fragmentation of seasonal ranges actually affect the migration route itself. In contrast, anthropogenic features, such as roads, fences, power lines and pipelines, often overlap or bisect migration routes and are commonly cited as sources of habitat fragmentation or barriers with the potential to impede animal movement (Bolger *et al.* 2008; Harris *et al.* 2009; Dobson *et al.* 2010). Despite this recognition, our knowledge of how such barriers affect migration when they overlap with a migration route is limited.

It is clear that impermeable barriers, such as game-proof fences, inhibit the connectivity of migration routes, such that entire seasonal ranges become inaccessible. A total loss of connectivity presumably eliminates the ecological benefits of migration, which can include tracking gradients in high-quality forage (McNaughton 1985; Wilmshurst *et al.* 1999), accessing water holes (Williamson

& Williamson 1984; Bolger *et al.* 2008) and reducing predation (Fryxell & Sinclair 1988; Hebblewhite & Merrill 2007). In some cases, impermeable barriers have caused population declines that resulted in the loss of thousands of migratory ungulates (Williamson & Williamson 1984; Whyte & Joubert 1988; Spinage 1992; Ben-Shahar 1993).

Most anthropogenic features, however, are at least semi-permeable to ungulates, and the assumption that semi-permeable barriers elicit similar effects (i.e., loss of migration function, population declines) is not yet supported by empirical evidence, nor have the potential mechanisms for such effects been explored. While the emergence of corridor ecology research (e.g. Hilty, Lidicker & Merenlender 2006) has improved the awareness of barrier effects, most conservation attention has focused on impermeable barriers (e.g. Dobson *et al.* 2010; Holdo *et al.* 2011). This is due in part to the difficulties associated with studying subtle and potentially long-term behavioural changes in migratory animals. However, recent improvements in GPS technology have advanced the study of migratory animals, and rapid increases in energy and urban development have prompted new interest in understanding how migratory ungulates might be influenced when semi-permeable barriers are constructed within their routes.

To facilitate a mechanistic understanding of semi-permeable barrier effects, we distinguish here between 'connectivity' and the 'functional attributes' of a migration route. For our purposes, connectivity simply describes whether or not animals are able to move from one seasonal range to another, whereas the functional attributes of a route include access of locally important resources such as stopover sites, movement corridors and escape terrain, which allow animals to track vegetation phenology and balance predation risk (Fig. 1). Thus, when connectivity is lost due to construction of an impermeable

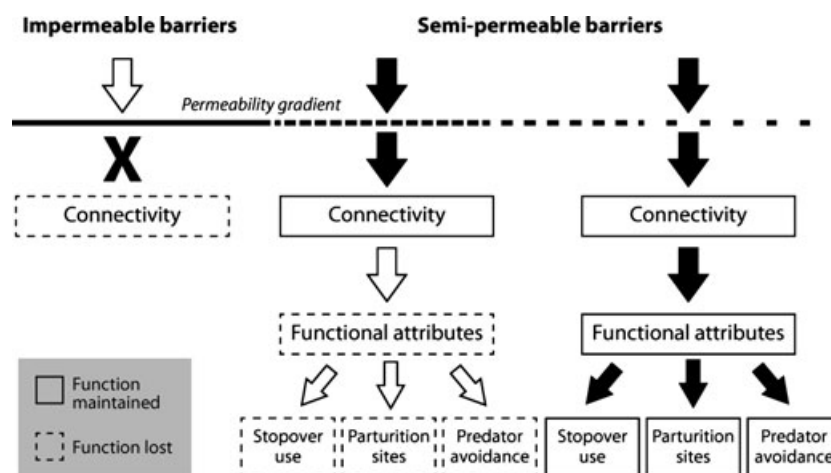


Fig. 1. Conceptual model that distinguishes between 'connectivity' and 'functional attributes' of a migration route and illustrates how each are affected by barriers (white arrows). Impermeable barriers impede connectivity such that animals can no longer migrate between seasonal ranges. In contrast, semi-permeable barriers often allow connectivity to be maintained, but the functional attributes of the migration route can be compromised, especially as permeability decreases.

barrier, the functional attributes of the migration route are also lost, along with the benefits of the seasonal range. Importantly, however, when connectivity remains intact because barriers are semi-permeable, the functional attributes of the migration routes may or may not be compromised. Thus, distinguishing between connectivity and functional attributes clarifies that impermeable and semi-permeable barriers may affect ungulate migration through different mechanisms.

Here, we propose a general framework to evaluate semi-permeable barrier effects on migratory ungulates, with the goal of expanding the discussion of barrier effects beyond the broad assumption that anthropogenic features will unconditionally impede migration. Our framework consists of three steps. First, the potential barrier is identified and measured in a way that facilitates the detection of development thresholds that alter behaviour. Roads, for example, are commonly viewed as potential barriers to migration (Dobson *et al.* 2010). However, a road or network of roads may not elicit a behavioural response until some threshold (e.g. road density, traffic levels, road width, etc.) is exceeded (Dyer *et al.* 2002; Frair *et al.* 2008). Thus, whether the potential barrier is a road, fence or other development, it should be measured in a way that considers likely thresholds. Second, the behavioural responses to a given anthropogenic feature are measured. We note that simply determining whether animals continue to migrate after construction of a potential barrier (e.g. Carruthers & Jakimchuk 1987; Ito *et al.* 2005) only provides information on connectivity and may overlook important behavioural changes. To examine whether semi-permeable barriers reduce the benefits of migration, specific migration behaviours (e.g. rate of movement, fidelity) must be quantified before and after the construction of the potential barrier (or in areas with and without barriers). These may include traditional metrics such as net-squared displacement and rate of movement, or the more advanced utilization distribution (UD) metrics now possible with movement-based kernel density estimation (MKDE; Benhamou 2011) and Brownian bridge movement models (BBMM; Horne *et al.* 2007; Kranstauber *et al.* 2012). Next, to predict how the observed behavioural changes may influence the functionality of the migration route, it is necessary to characterize functional attributes (e.g. stopover sites, escape terrain, parturition) of the migratory landscape. This third step highlights the importance of linking observed behavioural changes to functional attributes of the migratory landscape, thereby providing a means to evaluate how the benefits of migration may be altered by behavioural changes caused by barriers.

We illustrate our framework using empirical data from migratory mule deer *Odocoileus hemionus* in Wyoming, USA. Like many areas of western North America, ungulate ranges in Wyoming are experiencing unprecedented levels of energy development (Sawyer, Kauffman & Nielson 2009; Sawyer *et al.* 2009). Although the scale and intensity

of development are rapidly increasing (Copeland *et al.* 2009), we know little about whether energy infrastructure alters migratory behaviour, the functionality of migration routes or the ecological benefits of migration. Here, we use GPS movement data to examine the behavioural response of two migratory mule deer populations to varying levels of energy development. Using migration routes identified prior to large-scale natural gas development as the baseline, our goal was to determine how mule deer migration was influenced by increased levels of gas development. We examined several complementary metrics of behavioural change and evaluated how they affected the functional attributes of the migratory landscape, with an emphasis on understanding how semi-permeable barriers alter the benefits of migration. By revealing differential responses of mule deer to varying levels of development, our findings highlight the importance of considering semi-permeable barriers in land-use planning – an urgent goal amid ongoing global declines in ungulate migration.

Materials and methods

STUDY AREA

Our study was conducted in the 1100-km² Atlantic Rim Project Area (ARPA), located in south-central Wyoming. The ARPA is generally characterized by rolling topography, prominent ridges and dry canyons dominated by sagebrush *Artemisia* sp., black greasewood *Sarcobatus vermiculatus* and other mixed shrubs *Purshia tridentata*, *Chrysothamnus* sp., *Cercocarpus* sp. Elevations range from 1920 to 2530 m. The ARPA contains two distinct mule deer winter ranges known as the Dad and Wild Horse winter ranges. The Dad winter range supports 500–1000 mule deer, whereas the Wild Horse range supports 1500–2000. Population-level migration routes for both winter ranges were identified in 2005 and 2006 (Sawyer *et al.* 2009), during a period of exploratory energy development that we refer to as Phase 1 (Figs 2 and 3). Shortly thereafter, the Bureau of Land Management (BLM) approved development of *c.* 2000 wells to extract coal-bed methane from the ARPA (BLM 2007). Although most of the development was planned for areas outside of mule deer migration routes, there were two areas where development overlapped with migration routes, including the 33.6-km² Dry Cow Creek located northeast of the Dad winter range (Fig. 2) and the 15.5-km² Wild Horse Basin located east of the Wild Horse winter range (Fig. 3).

ANIMAL CAPTURE AND DATA COLLECTION

We captured 47 mule deer during Phase 1 and equipped animals with store-on-board GPS collars that collected locations every 2.5 h (Sawyer *et al.* 2009). Between February 2005 and November 2006, we collected 116 494 locations from the 47 deer to document spring and autumn migrations. We refer readers to Sawyer *et al.* (2009) for further details on Phase 1. During Phase 2, we captured 56 mule deer and equipped them with GPS collars programmed to collect locations every 2 hours during migration. Collars collected data for spring and autumn migrations of 2008,

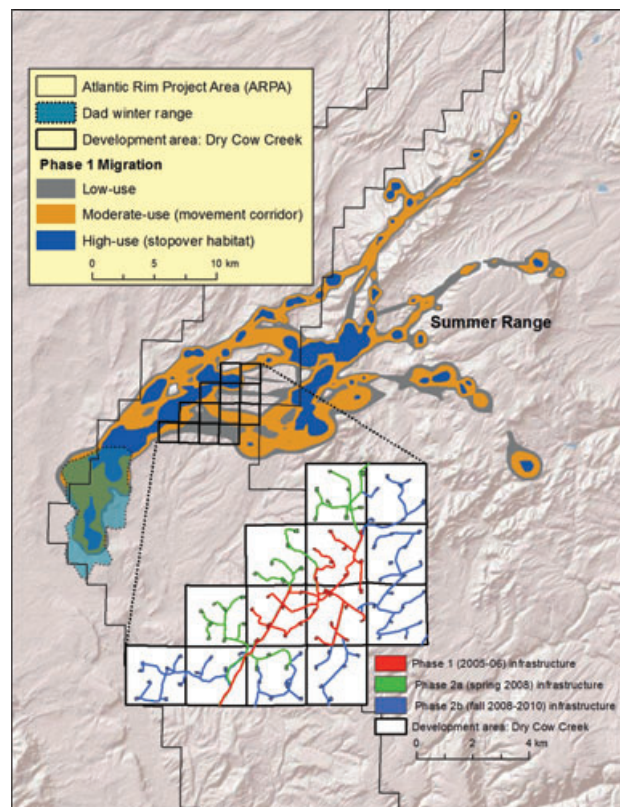


Fig. 2. Location of 33.6-km² Dry Cow Creek development area within the population-level migration route estimated for mule deer from the Dad winter range during Phase 1. Map insert shows the level of gas development in Dry Cow Creek during Phase 1 (2005–06), Phase 2a (spring 2008) and Phase 2b (autumn 2008–2010). Infrastructure includes roads (linear features) and gas pads (small squares).

2009 and 2010. During Phase 2, we recovered 191 302 GPS locations from 50 of the 56 marked animals. Of those 50 animals, 39 (26 in Wild Horse winter range and 13 in Dad winter range) lived long enough to complete at least one migration. Fix success of GPS collars was high (99%), so our analysis was not affected by missing locations.

IDENTIFYING THE POTENTIAL BARRIER

A critical component of studying barrier effects is to quantify potential barriers in terms that allow thresholds to be considered (Dyer *et al.* 2002; Frair *et al.* 2008). The potential barriers in our study included road networks and well pads associated with gas development. We used 10-m resolution satellite imagery acquired from Spot Image Corporation (Chantilly, VA, USA) to quantify road and well pad densities during each phase of development. We recognize that roads and well pads can have varying levels of human disturbance (e.g. traffic), depending on the type of wells (e.g. drilling vs. producing) and associated production facilities (Sawyer, Kauffman & Nielson 2009). However, we did not distinguish between road and well pad types because all roads in our development areas were improved gravel and *c.* 10 m wide, and well pads were similar in size and type.

DETECTING CHANGES IN MIGRATORY BEHAVIOUR

We sought to identify potential individual and population-level behavioural responses during migration. We calculated movement rates of mule deer ($n = 43$) through the development areas and

used a standard two-sample *t*-test ($\alpha = 0.10$) to determine whether movement rates varied between Phases 1 and 2. Movement rates were only calculated for animals that moved through development areas and were based on the movement sequence that included one location either side of the development area. To evaluate movement in the context of the larger migration route, we also calculated movement rates in undeveloped habitat, between the development areas and summer ranges. For a small sample of animals that collected data in both study phases ($n = 4$), we compared migration routes between years to assess whether animals detoured around the development area.

We used the Brownian bridge movement model (BBMM) to estimate population-level migration routes for GPS-collared deer from both the Dad and Wild Horse winter ranges. The BBMM uses time-specific location data to estimate a UD along a movement route, where the probability of being in an area is conditioned on the start and end locations, the elapsed time between locations and the speed of movement (Horne *et al.* 2007). We used the 'BBMM' package in R (R Foundation for Statistical Computing, Vienna, Austria) to estimate UDs for individual migration routes. Population-level migration routes were then estimated by averaging the individual UDs within each winter range and study phase. These population-level UDs provide a probabilistic measure of the migration route, where the height of UD reflects intensity of use and the contours of the UD delineate the surface area, or width of the route. Overall, the Phase 1 period (spring 2005–spring 2006) included 55 migrations (42 spring, 13 autumn) collected from 35 deer, whereas Phase 2 (spring 2008–autumn 2010) included 86 migration routes (56 spring, 30 autumn) from 39 deer. The Phase 1 population-level migration

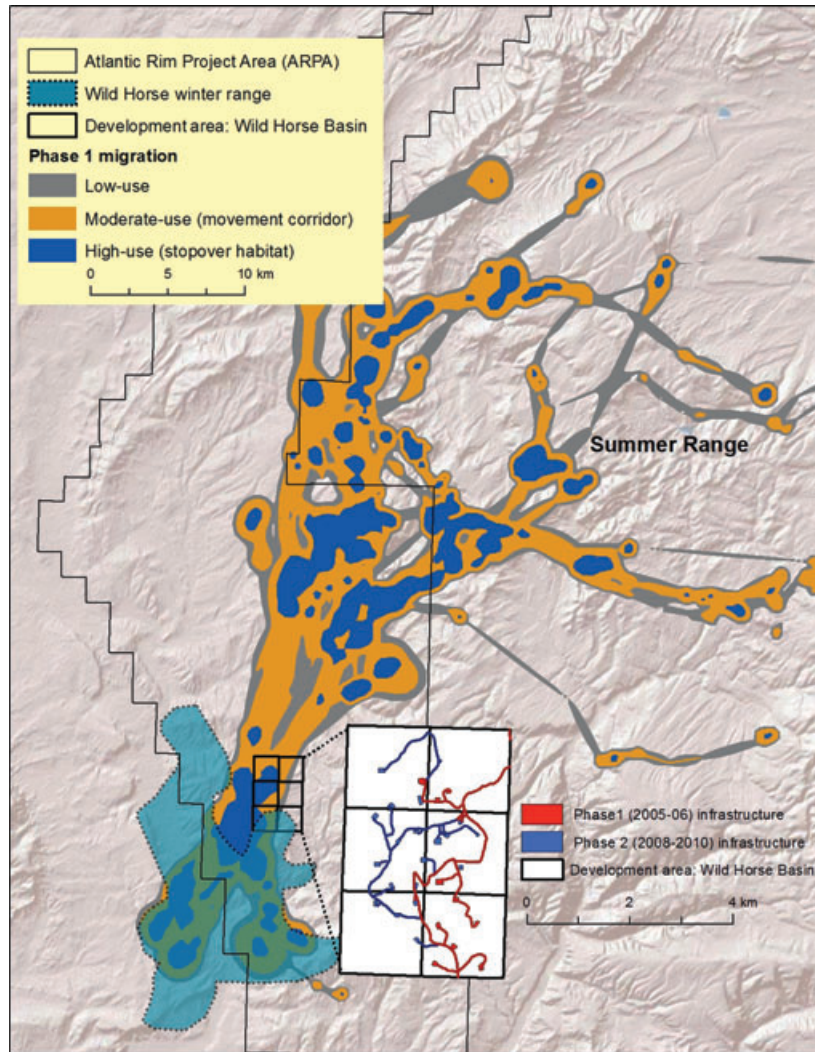


Fig. 3. Location of 15.5-km² Wild Horse Basin development area within the population-level migration route estimated for mule deer from the Dad winter range during Phase 1. Map insert shows the level of gas development in Wild Horse Basin during Phase 1 (2005–06) and Phase 2 (2008–2010). Infrastructure includes roads (linear features) and gas pads (small squares).

route for the Wild Horse winter range included 37 migrations by 23 deer, while the Dad winter range included 18 migrations by 12 deer (Figs 2 and 3). The Phase 2 population-level migration route for the Wild Horse winter range included 61 migrations by 23 deer. Phase 2 development in Dry Cow Creek was split into Phase 2a (spring 2008) and 2b (autumn 2008–autumn 2010), to account for the development activity during the summer of 2008. The population-level route for the Dad winter range included 12 migrations by 12 deer in Phase 2a, and 13 migrations by 9 deer in Phase 2b.

To evaluate whether the intensity of deer use (i.e. height of the UD) within migration routes changed in the development areas, we used the UD of migration routes estimated during Phase 1 as a reference and examined whether observed changes in the Dry Cow Creek and Wild Horse Basin were statistically different than those expected in a larger portion of the migration route. To do this, we designed a randomization procedure that estimated the expected change in deer use for a larger area (3 km buffer) surrounding both Dry Cow Creek and Wild Horse Basin development areas. For Dry Cow Creek, we randomly

selected 13, 2.6-km² units (equal to the size of the development area) from a larger sample of 51 and then calculated the percentage change in UD volume relative to Phase 1. This process was conducted 500 times and provided an estimate of the amount of change expected in any combination of 13, 2.6-km² units sampled from the larger 132-km² area. A similar process was repeated in Wild Horse Basin, except we randomly selected 6, 2.6-km² units from a larger sample of 21. We calculated 90% confidence intervals to test whether the changes observed in the development areas were more or less than expected based on the permutation results. Our randomization analysis used the three-dimensional structure or volume of UDs to detect changes in population-level migration use and is conceptually similar to the volume of intersection method described by Millsbaugh *et al.* (2004). We also calculated the change in the amount of migration surface area, as defined by the outer 99% contour of the population-level migration routes in the Dry Cow Creek and Wild Horse Basin during Phases 1 and 2. This simple, two-dimensional metric is useful for detecting change in the width of a migration route.

IDENTIFYING FUNCTIONAL ATTRIBUTES OF THE MIGRATORY LANDSCAPE

For temperate ungulates that migrate along elevation gradients, functional attributes of the migratory landscape can be generally characterized as either stopover habitat where animals spend most of their time, or the intervening movement corridors where animals travel quickly (Sawyer *et al.* 2009; Sawyer & Kauffman 2011). We defined migratory segments as either stopover habitat or movement corridors, although we note that future studies may use or reveal additional functional attributes, such as parturition sites (e.g. Singh *et al.* 2010; Barbknecht *et al.* 2011). Stopover sites were classified as the highest 25% quartile in the UD, whereas the 50–75% quartiles were considered movement corridors (Sawyer *et al.* 2009). At the individual level, we calculated the area of stopover habitat for each deer ($n = 43$) before and after development to assess whether this functional attribute was influenced by increased levels of development.

Results

IDENTIFYING THE POTENTIAL BARRIER

The Dry Cow Creek area was partially developed during Phase 1, with road and well pad densities of 0.56 km km^{-2} and 0.77 km^{-2} , respectively. However, by the spring of 2008 (Phase 2a), road and well pad densities increased to 1.07 km km^{-2} and 1.49 km^{-2} , respectively. Following construction in summer 2008 (Phase 2b), the road and well pad densities increased further to 1.92 km km^{-2} and 2.82 km^{-2} , respectively (Fig. 2). Compared to Dry Cow Creek, gas development in Wild Horse Basin was smaller in size and intensity. Road and well pad densities during Phase 1 were 0.83 km km^{-2} and 0.65 km^{-2} , respectively, and increased to 1.51 km km^{-2} and 1.86 km^{-2} during Phase 2 (Fig. 3).

CHANGES IN MIGRATORY BEHAVIOUR

At the individual level, movement rates of deer in the Dry Cow Creek development steadily increased from $1.06 \pm 0.26 \text{ km h}^{-1}$ (mean \pm SE) in Phase 1 to 1.68 ± 0.21 in Phase 2a, and 1.94 ± 0.18 in Phase 2b (Fig. 4). Movement rates in Phase 2b were higher than those observed in Phase 1 ($t_{11} = -2.68$, $P = 0.021$). Concurrently, movement rates of deer after they had moved through the development area steadily decreased from 1.25 ± 0.12 in Phase 1 to 0.79 ± 0.27 in Phase 2a, and 0.21 ± 0.05 in Phase 2b (Fig. 4). The rate of deer movement in undeveloped areas was lower in Phase 2b compared with Phase 1 ($t_{11} = 7.68$, $P < 0.001$). Of the 4 deer that collected data in both Phase 2a and 2b, three animals appeared to alter their routes in response to development by diverging from the previous year's path near the development boundary and then moving back to the path c. 3–4 km beyond the development (Fig. 5). Overall, the detours used by these animals bypassed approximately 8 km of their original migration route. At the population level, the intensity of deer use, as indicated by the UD volume, declined by 10% and 53% in Phases 2a and 2b, respectively (Fig. 6). The 53% decrease was statistically significant and coincided with road and well pad densities of 1.92 km km^{-2} and 2.82 km km^{-2} , respectively. Similarly, the surface area of migration routes in the Dry Cow Creek steadily decreased from 23.4 km^2 in Phase 1 to 21.5 km^2 in Phase 2a (-8%) and 15.4 km^2 in Phase 2b (-34%).

In contrast to the altered movement rates that followed development in the Dry Cow Creek, we did not detect any individual or population-level responses in the smaller and less concentrated development of Wild Horse Basin. Movement rates through the development area did not differ ($t_{17} = 0.56$, $P = 0.579$) between Phase 1 ($1.24 \pm 0.30 \text{ km}$

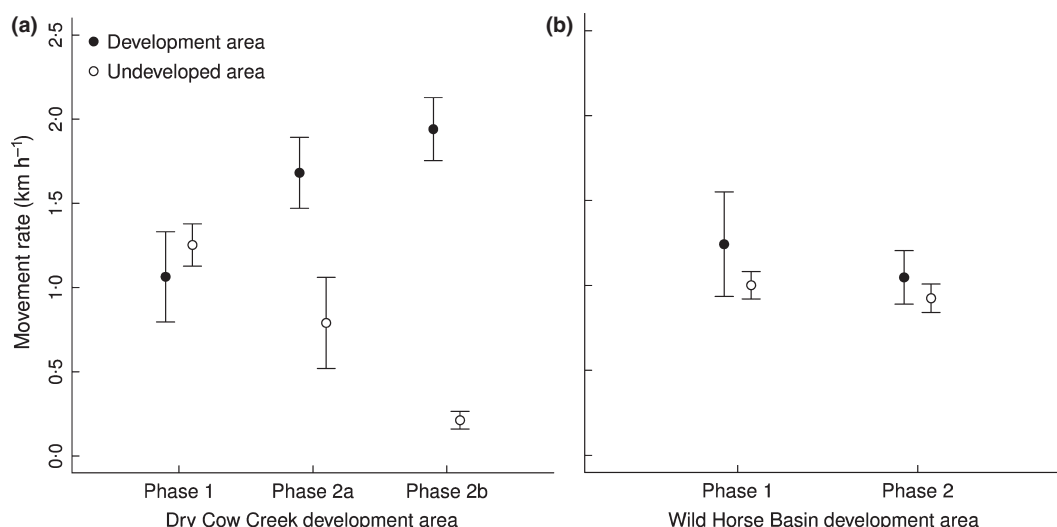


Fig. 4. (a) Movement rates (mean $\text{km h}^{-1} \pm \text{SE}$) of mule deer through the Dry Cow Creek development area during Phases 1, 2a and 2b. Movement rates through the developed area were higher during Phases 2a and 2b compared with Phase 1, whereas movement rates through undeveloped habitat decreased. (b) Movement rates of mule deer through the Wild Horse Basin development area during Phases 1 and 2. Movement rates through developed and undeveloped areas were similar in both phases.

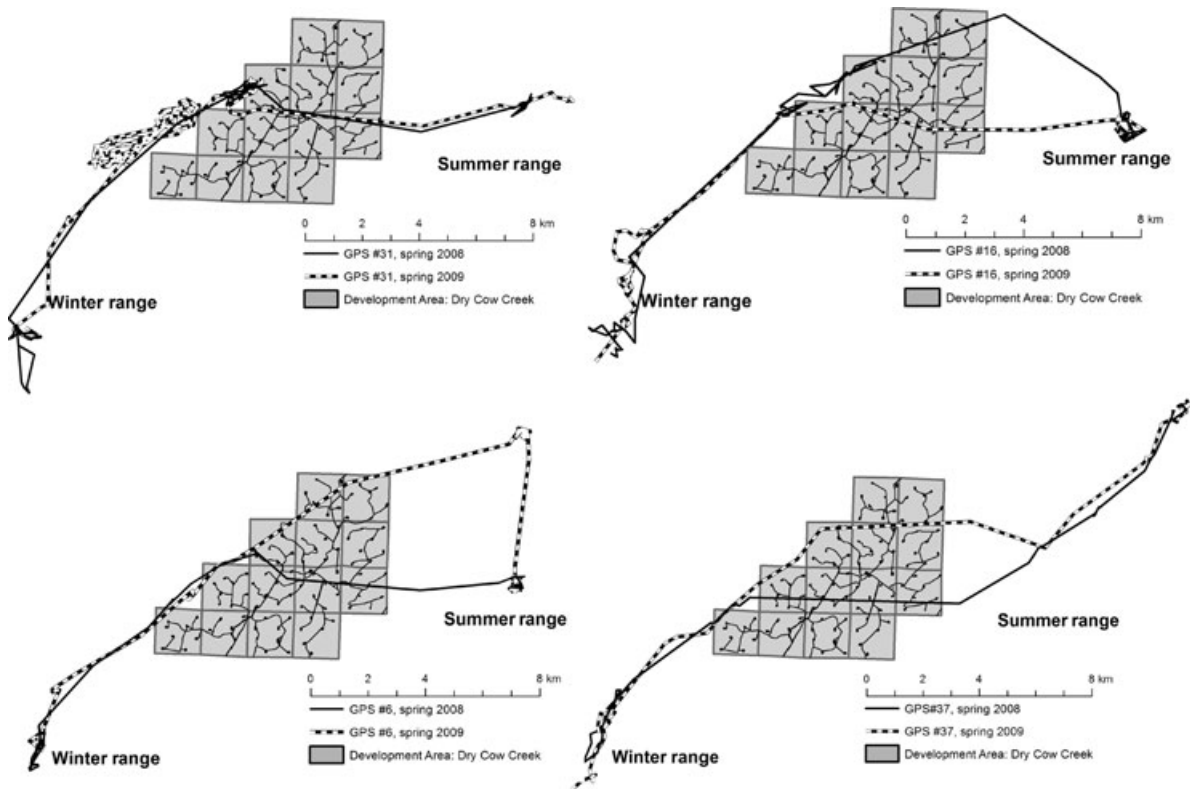


Fig. 5. Migration routes of four mule deer during Phase 2a and Phase 2b through the Dry Cow Creek development area. Deer #31 moved through the central portion of Dry Cow Creek in both Phases 2a and 2b, whereas Deer #16, #6 and #37 all show clear detours around or through different portions of the developed areas before and after development.

hr^{-1} ; mean \pm SE) and Phase 2 (1.05 ± 0.15 ; Fig. 4). Concurrently, movement rates outside of the development area also did not differ ($t_{17} = 0.66$, $P = 0.516$) between Phase 1 ($1.00 \pm 0.08 \text{ km hr}^{-1}$; mean \pm SE) and Phase 2 (0.92 ± 0.08 ; Fig. 4). At the population level, the intensity of deer use decreased by 23% in Phase 2, but was within the confidence intervals of the expected variance in deer use (Fig. 6). The surface area of migration route was similar between Phase 1 (10.9 km^2) and Phase 2 (12.1 km^2).

FUNCTIONAL ATTRIBUTES OF MIGRATORY LANDSCAPE

For individual deer migrating through Dry Cow Creek, the area of stopover habitat decreased as development increased, with an average of $1.63 \pm 0.43 \text{ km}^2$ (mean \pm SE) during Phase 1, $1.16 \pm 0.38 \text{ km}^2$ in Phase 2a and $0.66 \pm 0.19 \text{ km}^2$ in Phase 2b (Fig. 7). The area of stopover habitat used during Phase 2b was marginally lower than Phase 1 ($t_9 = 2.04$, $P = 0.07$). For individual deer migrating through Wild Horse Basin, the area of stopover habitat was similar ($t_{19} = -0.611$, $P = 0.548$) between Phase 1 ($1.30 \pm 0.34 \text{ km}^2$) and Phase 2 ($1.63 \pm 0.41 \text{ km}^2$; Fig. 7).

Discussion

Sustaining migratory ungulate populations in the face of widespread development and land-use change poses diffi-

cult conservation challenges across the globe (Bolger *et al.* 2008; Harris *et al.* 2009). Increased levels of development create a variety of barriers (e.g. roads, pipelines, fences) that are semi-permeable to ungulates; yet, we know little about how these types of barriers influence migratory behaviour or the persistence of migratory populations. We found that changes in migratory behaviour of two mule deer populations in western Wyoming varied with the size and intensity of semi-permeable barriers associated with gas development. In migration routes exposed to a larger, more concentrated development (i.e. Dry Cow Creek), mule deer use declined by 53% and movement rates nearly doubled ($1.06\text{--}1.94 \text{ km h}^{-1}$). The decline in deer use and accelerated movement rates reduced both the surface area of the migration route and area of stopover use. In contrast, we did not detect any changes in migratory behaviour through Wild Horse Basin, where the development area was smaller and infrastructure less concentrated. The intensity of deer use, surface area of the routes, movement rates of animals, and stopover use were similar before and after gas development. Presumably, the absence of any detectable response by migrating deer in this area was a function of permeability thresholds, due to either the lower level or smaller size of the development. Additionally, timing stipulations restricted development activities (i.e. drilling) in Wild Horse Basin between 1 November and 30 April – a time period that includes

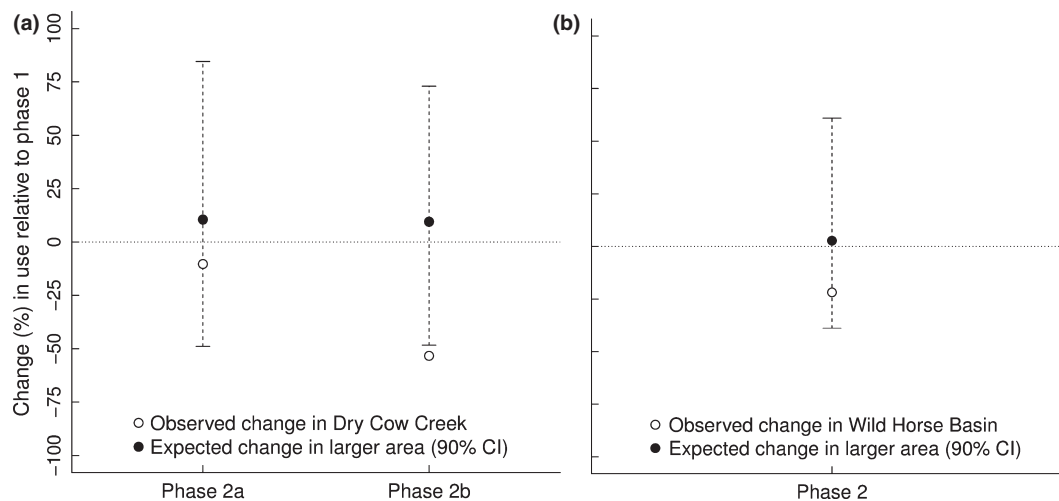


Fig. 6. (a) Change in population-level deer use in Dry Cow Creek development area during Phases 2a and 2b, relative to a larger 132-km² area and using Phase 1 as a reference level. (b) Change in population-level deer use in Wild Horse Basin development area during Phase 2, relative to a larger 54-km² area and using Phase 1 as a reference level.

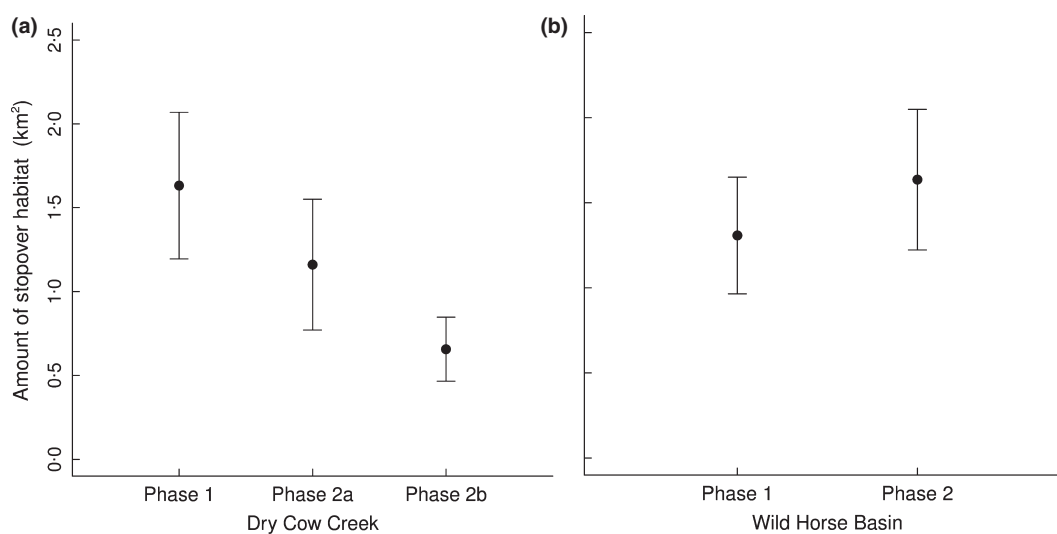


Fig. 7. (a) Area of stopover habitat (mean km² ± SE) used by mule deer in the Dry Cow Creek development area during Phases 1, 2a and 2b, and (b) Wild Horse Basin development area during Phases 1 and 2.

much of the spring and autumn migrations. Reducing traffic levels can reduce disturbance to mule deer (Sawyer, Kauffman & Nielson 2009), so these restrictions may have mitigated the potential barrier effects by minimizing disturbance to mule deer.

Our finding of individual and population-level responses to semi-permeable barriers makes clear that anthropogenic features can affect migration, even when connectivity between seasonal ranges is maintained. However, it is of interest whether these behavioural changes reduce the functionality of migration routes and ultimately, whether the functional loss could affect demography and persistence of migrants that use impacted routes. For example, stakeholders involved with this study have posed the question, 'Why does it matter if deer migrate more quickly through the development area?' Without a reasonable

answer to this question, agencies and industry are less motivated to modify, or attempt to mitigate, development plans that overlap with ungulate migration routes. Recent work suggests mule deer spend 95% of the migration period in stopovers, essentially using them to slow down their migration to exploit forage quality gradients created by phenological delays associated with elevation (Sawyer & Kauffman 2011). Our analyses suggest that development within a route can increase movement rates and alter migration route function by reducing stopover use. Although only 15% of the migration route in Dry Cow Creek was classified as stopover habitat, a 60% (1.63–0.66 km²) reduction in the size of these areas is concerning. Any behavioural change that impedes access to or discourages use of stopover habitat is likely to reduce the ability of animals to optimally forage and track vegetation

phenology. Whether such a functional loss has measurable demographic consequences is unknown, but given the importance of summer nutrition for body condition and reproduction (Cook *et al.* 2004; Parker, Barboza & Gillingham 2009; Tollefson *et al.* 2010), lost foraging opportunities during migration certainly have the potential to incur energetic and demographic costs. Further study, as has been done with avian taxa (e.g. Hoyer *et al.* 2012), is needed to link altered migratory behaviour by ungulates to fitness metrics (e.g. body condition, reproduction, survival).

Sawyer *et al.* (2009) suggest that semi-permeable barriers situated in movement corridors are less likely to impact migration route function than barriers in stopover areas, because animals do not rely on movement corridors as primary sources of forage. We caution, however, that changes in migratory behaviour within movement corridors have the potential to influence other, more subtle migration route functions. For example, it is possible that ungulates collect information on forage phenology while travelling through movement corridors to optimize the rate at which they access peak digestibility of forage (Sawyer & Kauffman 2011). Interestingly, our results suggest that when animals move more rapidly through developed areas, they tend to offset the quick movement by slowing down once they return to undeveloped habitat. This pattern is consistent with the hypothesis that increased movement rates create short-term phenological mismatches, and that animals attempt to correct for these mismatches by slowing down after moving through developed areas. Given the potential consequences of phenological mismatches (Post & Forchhammer 2008), this movement pattern warrants further research, especially in areas where development projects bisect long segments of migration routes. Of additional concern is that many migratory ungulates show high fidelity to migration routes (Berger, Cain & Berger 2006; Sawyer *et al.* 2009; Bunnefeld *et al.* 2011), and it is unknown how detours made along the route due to disturbance will influence movement rates and the ability of animals to track phenology. Certainly, when deer bypass 8 km of their traditional migration routes, like those in Dry Cow Creek, the functionality of that particular segment is effectively lost. Thus, there are a variety of mechanisms (i.e. increased movement rates and detouring) by which semi-permeable barriers may diminish the ability of migrants to track optimal forage conditions.

Most ungulate populations are partially migratory (Cagnacci *et al.* 2011; Hebblewhite & Merrill 2011), but the proportion of migratory animals is typically larger than the resident segment (Fryxell, Greever & Sinclair 1988; Bunnefeld *et al.* 2011). Our study was no exception, as only four of the 103 GPS-marked animals were resident. Recent studies suggest that the ratio of migratory to resident animals may shift when the benefits of migrating no longer exceed the benefits of a resident strategy (Hebblewhite & Merrill 2011). For example, elk populations have become increasingly resident in areas where

differential levels of predation on neonates and changes in habitat quality favour the resident strategy (Hebblewhite *et al.* 2006; Hebblewhite & Merrill 2011; Middleton *et al.* in press). Our work highlights the possibility that, like changes in predation or habitat quality, the effects of semi-permeable barriers on migration route function have the potential to reduce the benefits of migration and favour resident animals. Given that ungulate migrations generally occur along traditional routes that are learned and passed on from mother to young (McCullough 1985; Sweaner & Sandegren 1988; Nelson & Mech 1999), it may be difficult to restore migratory landscapes by removing barriers once migratory subpopulations have dwindled (but see Bartlam-Brooks, Bonyongo & Harris 2011). In general, ungulates that demonstrate strong fidelity to narrow, linear pathways (Berger, Cain & Berger 2006; Sawyer & Kauffman 2011) may be more vulnerable to barrier effects than those exhibiting more nomadic migratory patterns, such as wildebeest *Connochaetes taurinus* (Holdo, Holt & Fryxell 2009) and Mongolian gazelles *Procapra gutturosa* (Mueller *et al.* 2011). However, in contrast to populations that follow distinct migration routes, mitigating the potential effects of semi-permeable barriers for nomadic populations will be difficult because of their unpredictable movements across the landscape (Mueller *et al.* 2011).

Ideally, our study would have followed the same animals through the entire study period, such that changes in individual movements could be more closely examined. For example, the 4 animals that collected data during two phases revealed that increased levels of development may lead to individual animals detouring and bypassing entire segments of their traditional routes. Other work has found that increased levels of human disturbance may interact with environmental conditions to discourage older individuals from migrating (Singh *et al.* 2012). Thus, we suspect that evaluating individual movements through time would provide more insight into the mechanistic drivers of the behavioural changes we observed and reduce the amount of variation in the metrics of interest. For future studies, we recommend the same animals be marked through the entire study period so that individual and population-level movement patterns can be examined in more detail. Also critical to detecting changes in behaviour is the collection of baseline data before intensive development. In our case, had state and federal agencies not required both pre- and post-development study phases, changes in migratory behaviour would have gone undocumented.

CONCLUSIONS AND RECOMMENDATIONS

Managing migratory ungulates is especially difficult because of the long distances they move, often across a mix of land ownership and land-use practices. As energy development and other human disturbances expand, it is increasingly important to understand how migrating

ungulates respond to the semi-permeable barriers. Our study suggests that increased levels of gas development in migration routes may encourage detouring, increase movement rates, reduce the area of stopover use by individuals and reduce the overall amount of deer use and constrict the size of migration routes at the population level. The existence of such behavioural changes suggests that certain levels of development, while still allowing connectivity between seasonal ranges, may nevertheless reduce route functionality and the benefits of migration. Ultimately, demographic costs associated with barriers are the most desirable currency in which to measure the effects of development on migratory ungulates. In the absence of such data, quantifying behavioural changes and functional attributes of the migratory landscape before and after development provides an intuitive first step for understanding the consequences of semi-permeable barriers for the persistence of migratory ungulates.

Acknowledgements

This work was supported by the Bureau of Land Management, the Wyoming Game and Fish Department, the U.S. Geological Survey (Wyoming Landscape Conservation Initiative) and grants provided to Western Ecosystems Technology, Inc from Anadarko Petroleum Company. Any use of trade, product or firm names is for descriptive purposes only and does not imply endorsement by the U. S. Government. Comments from E. J. Milner-Gulland, G. Hopcraft, N. Korfanta and one anonymous reviewer improved the manuscript.

References

- Barbknecht, A.E., Fairbanks, W.S., Rogerson, J.D., Maichak, E.J., Scurlock, B.M. & Meadows, L.L. (2011) Elk parturition site selection at the local and landscape scales. *Journal of Wildlife Management*, **75**, 646–654.
- Bartlam-Brooks, H.L.A., Bonyongo, M.C. & Harris, S. (2011) Will reconnecting ecosystems allow long-distance mammal migrations to resume? A case study of a zebra *Equus burchelli* migration in Botswana. *Oryx*, **45**, 210–216.
- Benhamou, S. (2011) Dynamic approach to space and habitat use based on biased random bridges. *PLoS ONE*, **6**, e14592.
- Ben-Shahar, R. (1993) Does fencing reduce the carrying capacity for populations of large herbivores? *Journal of Tropical Ecology*, **9**, 249–253.
- Berger, J. (2004) The last mile: how to sustain long-distance migration in mammals. *Conservation Biology*, **18**, 320–331.
- Berger, J., Cain, S.L. & Berger, K.M. (2006) Connecting the dots: an invariant migration corridor links the Holocene to the present. *Biology Letters*, **22**, 528–531.
- Bolger, D.T., Newmark, W.D., Morrison, T.A. & Doak, D.F. (2008) The need for integrative approaches to understand and conserve migratory ungulates. *Ecology Letters*, **11**, 63–77.
- Bunnefeld, N., Börger, L., van Moorter, B., Rolandsen, C.M., Dettki, H., Solberg, E.J. & Ericsson, G. (2011) A model-driven approach to quantify migration patterns: individual, yearly and regional differences. *Journal of Animal Ecology*, **80**, 466–476.
- Bureau of Land Management (2007) *Record of Decision: Environmental Impact Statement for the Atlantic Rim Natural Gas Field Development Project*. Rawlins Field Office, Rawlins, Wyoming.
- Cagnacci, F., Focardi, S., Heurich, M., Stache, A., Hewison, A.J.M., Morellet, N., Kjellander, P., Linnell, J.D.C., Mysterud, A., Neteler, M., Delucchi, L., Ossi, F. & Urbano, F. (2011) Partial migration in roe deer: migratory and resident tactics are end points of a behavioural gradient determined by ecological factors. *Oikos*, **120**, 1790–1802.
- Carruthers, D.R. & Jakimchuk, R.D. (1987) Migratory movements of the Nelchina Caribou Herd in relation to the Trans-Alaska Pipeline. *Wildlife Society Bulletin*, **15**, 414–420.
- Cook, J.G., Johnson, B.K., Cook, R.C., Riggs, R.A., Delcurto, T., Bryant, L.D. & Irwin, L.L. (2004) Effects of summer-autumn nutrition and parturition date on reproduction and survival of elk. *Wildlife Monographs*, **155**, 1–61.
- Copeland, H.E., Doherty, K.E., Naugle, D.E., Pocewicz, A. & Kiesecker, J.M. (2009) Mapping oil and gas development potential in the US Intermountain West and estimating impacts to species. *PLoS ONE*, **4**, e7400. doi:10.1371/journal.pone.0007400.
- Dingle, H. (1996) *Migration: The Biology of Life on the Move*. Oxford University Press, New York.
- Dobson, A., Borner, M., Sinclair, A.R.E., et al. (2010) Road will ruin Serengeti. *Nature*, **467**, 272–273.
- Dyer, S.J., O'Neill, J.P., Wasel, S.M. & Boutin, S. (2002) Quantifying barrier effects of roads and seismic lines on movements of female woodland caribou in northeastern Alberta. *Canadian Journal of Zoology*, **80**, 839–845.
- Frair, J.L., Merrill, E.H., Beyer, H.L. & Morales, J.M. (2008) Thresholds in landscape connectivity and mortality risks in response to growing road networks. *Journal of Applied Ecology*, **45**, 1504–1513.
- Fryxell, J.M., Greever, J. & Sinclair, A.R.E. (1988) Why are migratory ungulates so abundant? *American Naturalist*, **131**, 781–798.
- Fryxell, J.M. & Sinclair, A.R.E. (1988) Causes and consequences of migration by large herbivores. *Trends in Ecology & Evolution*, **3**, 237–241.
- Harris, G., Thirgood, S., Hopcraft, J.G.C., Cromsigt, J.P.G.M. & Berger, J. (2009) Global decline in aggregated migrations of large terrestrial animals. *Endangered Species Research*, **7**, 55–76.
- Hebblewhite, M. & Merrill, E.H. (2007) Multi-scale wolf predation risk for elk: does migration reduce risk? *Oecologia*, **152**, 377–387.
- Hebblewhite, M. & Merrill, E.H. (2011) Demographic balancing of migrant and resident elk in a partially migratory population through forage-predation tradeoffs. *Oikos*, **120**, 1860–1870.
- Hebblewhite, M., Merrill, E.H., Morgantini, L.E., White, C.A., Allen, J. R., Bruns, E., Thurston, L. & Hurd, T.E. (2006) Is the migratory behavior of montane elk herds in peril? The case of Alberta's Ya Ya Tinda elk herd. *Wildlife Society Bulletin*, **35**, 1280–1294.
- Hilty, J.A., Lidicker Jr, W.Z. & Merenlender, A.M. (2006) *Corridor Ecology: The Science and Practice and Linking Landscapes for Biodiversity Conservation*. Island Press, Washington, USA.
- Hobbs, N.T. (1996) Modification of ecosystems by ungulates. *Journal of Wildlife Management*, **60**, 69–713.
- Holdo, R.M., Holt, R.D. & Fryxell, J.M. (2009) Opposing rainfall and plant nutritional gradients best explain the wildebeest migration in the Serengeti. *The American Naturalist*, **173**, 43–445.
- Holdo, R.M., Fryxell, J.M., Sinclair, A.R.E., Dobson, A. & Holt, R.D. (2011) Predicted impact of barriers to migration on the Serengeti wildebeest population. *PLoS ONE*, **6**, e16370.
- Horne, J.S., Garton, E.O., Krone, S.M. & Lewis, J.S. (2007) Analyzing animal movements using Brownian Bridges. *Ecology*, **88**, 2354–2363.
- Hoye, B.J., Hahn, S., Nolet, B.A. & Klaassen, M. (2012) Habitat use throughout migration: linking consistency, prior breeding success and future breeding potential. *Journal of Animal Ecology*, **81**, 657–666.
- Ito, T.Y., Miura, N., Lhagvasuren, B., Enkhbileg, D., Takatsuki, S., Tsunekawa, A. & Jiang, Z. (2005) Preliminary evidence of a barrier effect of a railroad on the migration of Mongolian gazelles. *Conservation Biology*, **19**, 945–948.
- Kranstauber, B., Kays, R., LaPoint, S.D., Wikelski, M. & Safi, K. (2012) A dynamic Brownian bridge movement model to estimate utilization distributions for heterogeneous animal movement. *Journal of Animal Ecology*, **81**, 738–746.
- McCullough, D.R. (1985) Long range movements of large terrestrial animals. *Contributions in Marine Science Supplement*, **27**, 444–465.
- McNaughton, S.J. (1985) Ecology of a grazing ecosystem: the Serengeti. *Ecological Monographs*, **55**, 260–294.
- Middleton, A.D., Kauffman, M.J., McWhirter, D.E., Cook, J.G., Cook, R.C., Nelson, A.A., Jimenez, M.D. & Klaver, R.W. Animal migration amid shifting patterns of phenology and predation: lessons learned from a Yellowstone elk herd. *Ecology*, in press.
- Millsbaugh, J.J., Gitzen, R.A., Kernohan, B.J., Larson, M.A. & Clay, C. L. (2004) Comparability of three analytical techniques to assess joint space use. *Wildlife Society Bulletin*, **32**, 148–157.
- Milner-Gulland, E.J., Fryxell, J.M. & Sinclair, A.R.E. (2011) *Animal Migration*. Oxford University Press, Oxford.
- Mueller, T., Olson, K.A., Dressler, G., Leimgruber, P., Fuller, T.K., Nicolson, C., Novaro, A.J., Bolgeri, M.J., Wattles, D., DeStefano, S., Calabrese, J.M. & Fagan, W.F. (2011) How landscape dynamics link

- individual- to population-level movement patterns: a multispecies comparison of ungulate relocation data. *Global Ecology and Biogeography*, **20**, 683–694.
- Nelson, M.E. & Mech, L.D. (1999) Twenty-year home range dynamics of a white-tailed deer matriline. *Canadian Journal of Zoology*, **77**, 1128–1135.
- Parker, K.L., Barboza, P.S. & Gillingham, M.P. (2009) Nutrition integrates environmental responses of ungulates. *Functional Ecology*, **23**, 57–69.
- Post, E. & Forchhammer, M.C. (2008) Climate change reduces reproductive success of an Arctic herbivore through trophic mismatch. *Philosophical Transactions of the Royal Society B*, **363**, 2369–2375.
- Sawyer, H. & Kauffman, M.J. (2011) Stopover ecology of a migratory ungulate. *Journal of Animal Ecology*, **80**, 1078–1087.
- Sawyer, H., Kauffman, M.J. & Nielson, R.M. (2009) Influence of well pad activity on the winter habitat selection patterns of mule deer. *Journal of Wildlife Management*, **73**, 1052–1061.
- Sawyer, H., Kauffman, M.J., Nielson, R.M. & Horne, J.S. (2009) Identifying and prioritizing ungulate migration routes for landscape-level conservation. *Ecological Applications*, **19**, 2016–2025.
- Singh, N.J., Grachev, I.A., Bekenov, A.B. & Milner-Gulland, E.J. (2010) Saiga antelope calving site selection is increasingly driven by human disturbance. *Biological Conservation*, **143**, 1770–1779.
- Singh, N.J., Börger, L., Dettki, H., Bunnefeld, N. & Ericsson, G. (2012) From migration to nomadism: movement variability in a northern ungulate across its latitudinal range. *Ecological Applications*, **22**, 2007–2020.
- Spinage, C.A. (1992) The decline of the Kalahari wildebeest. *Oryx*, **26**, 147–150.
- Sweanor, P.Y. & Sandegren, F. (1988) Migratory behavior of related moose. *Holarctic Ecology*, **11**, 190–193.
- Thirgood, S., Mosser, A., Tham, S., Hopcraft, G., Mwangomo, E., Mlengeya, T., Kilewo, M., Fryxell, J., Sinclair, A.R.E. & Borner, M. (2004) Can parks protect migratory ungulates? The case of the Serengeti wildebeest. *Animal Conservation*, **7**, 113–120.
- Tollefson, T.N., Shipley, L.A., Myers, W.L., Keisler, D.H. & Dasgupta, N. (2010) Influence of summer and autumn nutrition on body condition and reproduction in lactating mule deer. *Journal of Wildlife Management*, **74**, 974–986.
- Vors, L.S. & Boyce, M.S. (2009) Global declines of caribou and reindeer. *Global Change Biology*, **15**, 2626–2633.
- Whyte, I.J. & Joubert, S.C.J. (1988) Blue wildebeest population trends in the kruger national-park and the effects of fencing. *South African Journal of Wildlife Research*, **18**, 78–87.
- Williamson, D. & Williamson, J. (1984) Botswana's fences and the depletion of the Kalahari's wildlife. *Oryx*, **18**, 218–222.
- Wilmshurst, J.F., Fryxell, J.M., Farm, B.P., Sinclair, A.R.E. & Henschel, C.P. (1999) Spatial distribution of Serengeti wildebeest in relation to resources. *Canadian Journal of Zoology*, **77**, 1223–1232.

Received 2 July 2012; accepted 24 October 2012

Handling Editor: Nathalie Pettorelli

Risk-disturbance overrides density dependence in a hunted colonial rodent, the black-tailed prairie dog *Cynomys ludovicianus*

JONATHAN N. PAULI and STEVEN W. BUSKIRK

Department of Zoology and Physiology, University of Wyoming, 1000 E. University Avenue, Laramie, Wyoming 82071–3166, USA

Summary

1. Traditional understanding of how hunting affects vertebrate populations emphasizes competitive release and density dependence of vital rates, but more recent thinking has proposed complex non-lethal responses to hunting disturbance and predation risk. Colonial species have been proposed to be more vulnerable than dispersed, solitary species to disturbance and perceived risk from hunting. However, empirical comparisons of density dependence vs. risk disturbance in hunted species are few.

2. To compare density dependence with risk-disturbance effects of hunting on individuals and populations of a colonial species, we tested the response of black-tailed prairie dogs *Cynomys ludovicianus* to shooting in a before–after, treatment–control experiment. We subjected five colonies to a pulse of shooting, and compared individual and colony attributes to those of five control colonies, protected from shooting.

3. Surviving prairie dogs increased alert behaviours eightfold and reduced both above-ground activity and time spent foraging by 66%. Changes in behaviour lowered the body condition of surviving adults by 35%. Survivors of shooting, especially juveniles, exhibited elevated stress levels; faecal corticosterone concentrations increased by 80% among juveniles. Unexpectedly, overwinter survival rates did not increase in response to reduced prairie dog density. Colonies subjected to shooting experienced reproductive near-collapse the summer after shooting; pregnancy rates declined by 50% and reproductive output fell by 82%.

4. Risk-disturbance overwhelmed any possible density-dependent effects of shooting in prairie dogs, which exhibited additive mortality in response to hunting, and reproductive failure 1 year after shooting. Risk-disturbance was the predominant mechanism whereby individuals and colonies were affected by hunting.

5. *Synthesis and applications.* Because of their coloniality, prairie dogs possess certain life-history traits that predisposed them to be particularly susceptible to hunting-associated disturbances, which had cascading effects on population-level processes. Our findings contradict the general belief that small-bodied mammals quickly rebound from hunting exploitation via compensatory mortality and reproduction. Managers should consider measures to reduce recreational shooting intensity and duration in regions where black-tailed prairie dog colony growth and persistence is desired, yet allow shooting in areas where colonies conflict with landowner interests.

Key-words: Allee effect, coloniality, density dependence, hunting, prairie dog, risk-disturbance, Wyoming.

Journal of Applied Ecology (2007) **44**, 1219–1230

doi: 10.1111/j.1365-2664.2007.01337.x

Introduction

Ecologists increasingly realize that predators, including human hunters, can affect populations of prey in indirect but important ways. Under the risk of predation, animals may alter activity regimes, habitat-use patterns or other behaviours (Lima 1998). Such decision-making occurs when perceived risk of predation constrains other fitness-enhancing behaviours: animals trade-off various kinds of risk – such as starvation, being killed by a predator and not finding a mate – in order to maximize fitness (Lima 1998). These trade-offs have been postulated to result in cascades of effects that can extend to population size, demography and even the structure of ecosystems (Lima 1998; Ripple & Beschta 2004).

The predation–risk paradigm has been applied to animals responding to humans as the risk-disturbance hypothesis (Frid & Dill 2002). In response to hunting or other human disturbance, vertebrates may increase vigilance (Kilgo, Labisky & Fritzen 1998), alter foraging regimes (Roy & Woolf 2001), shift migration routes (Béchet *et al.* 2003) or use resource-poor habitats (Madsen 1998). Such behaviours have been shown to be capable of lowering body condition (Féret *et al.* 2003) or reducing reproductive output (Mainguy *et al.* 2002). Animals unable to avoid risk or disturbance may exhibit other symptoms such as physiological stress (Bateson & Bradshaw 1997). Clearly, these hypotheses and mechanisms have given us a fuller understanding of how animals with adaptive behaviours maximize fitness in the face of human disturbance, or suffer fitness losses in response to unavoidable perceived risk.

At the population level, positive growth ($\lambda > 1$) might counteract the effects of perceived risk, and has been attributed commonly to density dependence. Following hunting, populations limited by resources should enjoy higher survival and reproduction, and such density-dependent responses have been shown for a wide range of hunted taxa (Fowler 1987). Generally, reducing populations below carrying capacity should increase rates of somatic growth, survival and reproduction (Kokko 2001). Physiological stress may decline in response to reduced social interactions, and levels of intraspecific conflict are also expected to decrease. Therefore, human hunting and associated disturbance can be postulated to have contradictory effects on populations, primarily via the two mechanisms of risk-disturbance and density dependence.

Which of these mechanisms is more important might depend on degree of sociality and site fidelity. Gill, Norris & Sutherland (2001) hypothesized that vulnerability to disturbance should be related inversely to capacity to relocate to alternative habitats, and Fitz-Gibbon (1998) reviewed the mechanisms whereby some colonial species are more vulnerable to hunting disturbance than solitary ones. Colonial species cannot relocate to areas away from disturbance, tend to communicate threats to each other and are prone to social disruption (Stephens *et al.* 2002). Such hypothesized vulnerabilities

are additive to the fundamental one: that colonies represent high and predictable local concentrations of the species sought. Therefore, the risk-disturbance costs of hunting to surviving animals – perhaps extending to population processes – should be higher among colonial species than solitary, dispersed ones.

In the absence of human hunting, the ecological costs and benefits of colonial living have been well described (Hoogland & Sherman 1976; Hoogland 1979). Colonial animals benefit from cooperative breeding, shared vigilance, decreased predation, increased foraging efficiency and shared habitat enhancement and maintenance. Colony members also pay certain costs: increased disease transmission, intensified intraspecific competition for resources, easier detection by predators and heightened sensitivity to behavioural disturbances. It is unknown, however, how and to what degree an evolutionarily novel perturbation, such as modern human hunting, alters these costs and benefits of coloniality.

The black-tailed prairie dog *Cynomys ludovicianus*, a colonial sciurid rodent of the plains of North America, is an ideal model for examining responses to hunting in a colonial species. Studied intensively throughout their range, in part because of the ecologically pivotal role they play in grassland systems (Miller, Ceballos & Reading 1994), the natural history and coloniality of this species have been well documented (e.g. Hoogland 1995). Prairie dogs have only recently been subject to hunting and, in contrast to most other hunted taxa, are rarely killed for meat or fur but are used as targets by recreational shooters, who typically use high-velocity rifles effective at ≤ 500 m (Reeve & Vosburgh 2005). Most states impose no seasonal restrictions, harvest limits or licensing requirements on prairie dog shooting (Reeve & Vosburgh 2005), and a single shooter can shoot scores of prairie dogs in a single session (Vosburgh & Irby 1998). In recent years, recreational shooters have reported killing $> 2\,000\,000$ black-tailed prairie dogs year⁻¹ from three states combined (Reeve & Vosburgh 2005). Therefore, shooting is a widespread population influence across the range of *Cynomys*.

Although previous studies (Knowles 1982; Vosburgh & Irby 1998) have described some population and behavioural effects of shooting on prairie dogs, none has evaluated the relative importance of risk-disturbance vs. density-dependent effects. We experimentally tested the response of black-tailed prairie dogs to recreational shooting, and examined some mechanisms that might contribute to the effects hypothesized. We measured population and environmental attributes of 10 colonies, subjected one-half of the colonies to a pulse of shooting, and compared prairie dog attributes between shot and protected colonies over two consecutive summers. Specifically, we quantified shooting-related changes in abundance, survival, reproduction and demography of colonies subjected to vs. protected from shooting. We also compared the behaviours, growth rates and stress levels of the survivors on hunted colonies with those on protected colonies. We predicted effects of both risk-disturbance

and density dependence, but favoured the former because of the hypothesized vulnerability of colonial species. Specifically, we predicted that surviving prairie dogs would spend less time foraging and allocate more time to vigilance or hiding underground in burrows (Blumstein & Pelletier 2005), and recognized that such effects might cascade to body condition and population-level attributes. At the same time, we expected to observe increased overwinter survival on colonies subjected to shooting, and a pulsed increase in reproduction the following summer as a result of competitive release.

Materials and methods

STUDY SITE AND SELECTION OF COLONIES

We conducted fieldwork in summers 2003–04 on private lands around Thunder Basin National Grassland (TBNG), north-eastern Wyoming (43°45' N, 105°00' W). TBNG encompasses > 230 000 ha of federal land in a mosaic of public and private lands. The region is characterized by rolling hills of mixed-grass prairies and sagebrush steppe habitats, dissected locally into small areas of badlands. Dominant plant species included blue grama (*Bouteloua gracilis*), western wheatgrass (*Agropyron cristatum*), needle-and-thread grass (*Stipa comata*), big sagebrush (*Artemisia tridentata*) and pricklypear cactus (*Opuntia polyacantha*). Deciduous trees, primarily plains cottonwood (*Populus deltoides*) and willows (*Salix* spp.), occur along drainages and at some springs. Ponderosa pine (*Pinus ponderosa*) is common at higher elevations. TBNG has hot, dry summers and cold, dry winters. Mean monthly temperatures range from -7°C in January to 22.8°C in July. Average annual precipitation is about 33 cm, 70% falling during April–August (National Weather Service, Weather Station no. 487810).

We selected 10 black-tailed prairie dog colonies on private land near TBNG for study, excluding colonies that had experienced poisoning, recreational shooting or plague in the previous 10 years. Landowners agreed not to poison or allow shooting on study colonies during our study; we placed signs at each colony prohibiting shooting and monitored for compliance. All colonies were sufficiently isolated that they functioned independently; mean distance to another study colony was 6.5 km (minimum = 2.3 km; see Fig. S1 in Supplementary material) and no dispersals between colonies were detected during fieldwork.

EXPERIMENTAL DESIGN

To control for potentially confounding variation, colonies were paired based on colony area, vegetation attributes, density of prairie dogs and grazing regime in spring 2003 (Appendix S1, see Supplementary material). One colony from each pair was randomly assigned as the treatment and subjected to recreational shooting during midsummer 2003, while the other served as the control, with no shooting.

We entered standardized values (Krebs 1999) of colony area, vegetation, prairie dog density and grazing regime into Horn's (1966) equation to estimate pairwise similarity indices for the 10 colonies. Pairwise distances (PWD) were estimated as $\text{PWD}_{i,j} = 1 - \text{PWS}_{i,j}$, where $\text{PWS}_{i,j}$ is Horn's similarity index for colonies i and j . We generated a cluster tree using the unweighted pair group method with arithmetic averages (UPGMA). Colonies clustered into distinct clades when analysed with pairwise distances (Fig. S1). We paired colonies sharing the smallest pairwise distance and assigned randomly one colony from each pair as the treatment (Fig. S1). In spring 2004, a plague epizootic reduced prairie dog abundance on one control colony by 95% (Pauli *et al.* 2006). Therefore, data from this colony (5C) and its paired treatment colony (5T) were excluded from analyses requiring data from 2004.

Five volunteers participated in shooting prairie dogs on treatment colonies during 15 June–21 July 2003. Shooters used high-velocity rifles of various models and calibres, with variable-power telescopic sights and various ammunition types, most frequently the 0-223 (5.56 × 45 mm). Shooters (1–2) fired shots opportunistically from positions 15–150 m away from a colony perimeter. An observer recorded the number of prairie dogs killed, the number of rounds fired and the length of each session (Table 1). At the end of our treatment in July 2003, shooters had reduced prairie dog abundance on each treatment colony by our objective of 25–30%. Although invasive, such experimentation was necessary to rigorously quantify prairie dog responses to recreational shooting. Further, mortality rates incurred by our treatment were nominal compared to those on colonies subjected to unregulated shooting and poisoning, which is common on private lands in this region.

We recovered, mapped and determined the sex and age (using the premolar gap method of Cox & Franklin 1990) of 68% ($n = 324$) of shot prairie dogs. Following collection of data from carcasses, we returned them to the colony to mimic typical shooting events. We evaluated whether our treatment was dispersed evenly across each colony by transferring mapped locations of shot prairie dogs to ArcGIS (ESRI, Redlands, CA, USA) where we calculated dispersion using a nearest neighbour algorithm. Nearest-neighbour values of carcasses (1T = 2.33, 2T = 2.70, 3T = 1.15, 4T = 2.21, 5T = 1.79) were > 1.00 (all $Z > 4.93$ and $P < 0.001$) for all colonies, showing that animals were shot in a spatially uniform pattern (Krebs 1999).

BEHAVIOUR, BODY CONDITION AND STRESS

We randomly established 100 × 100 m marked grids on each colony and recorded prairie dog activity levels and behaviour within each of them twice daily [1.5 h after sunrise and 2 h before sunset, coinciding with peak above-ground activity (Powell *et al.* 1994)] for five consecutive days. Observations were conducted concurrently with trapping, and paired colonies were observed

Table 1. Number of rounds fired, time spent shooting and number of black-tailed prairie dogs killed by recreational shooters on treatment colonies, Thunder Basin National Grassland, Wyoming, June–July 2003. The percentage of the population killed was estimated by dividing the number killed by the estimated population size

Colony	No. rounds fired	Time spent shooting (person h ⁻¹)	No. killed	% of population killed
1T	323	8.0	67	31
2T	146	14.3	27	32
3T	697	27.6	209	30
4T	174	9.4	37	30
5T	502	21.2	138	25

simultaneously to control for temporal variation. Observations were postponed during periods of rain or wind speed > 40 km h⁻¹. Observations followed the protocol of scan sampling (Martin & Bateson 1993) and were taken from an elevated position with a 20 × telescope. Following a 20-min delay after assuming our position to allow behaviour to stabilize (Powell *et al.* 1994), we observed and recorded the number and behaviour of prairie dogs on the grid at 10-min intervals for 70 min. We categorized behaviour as: (1) foraging, (2) alert (vigilant or calling), (3) moving, (4) resting, (5) socializing or (6) other. Because few prairie dogs were observed socializing or moving, these two behaviours were pooled with the behavioural category ‘other’ for subsequent statistical analyses.

Above-ground counts and behaviour were each highly autocorrelated within observation periods (all $r^2 > 0.50$). Therefore, we averaged these two variables for each observation period. We expressed behavioural data as the percentage of total time spent in a particular behaviour type. We also expressed above-ground activity as the percentage of animals estimated to occur in a colony (prairie dogs ha⁻¹) above ground at a given time. We used a repeated-measures analysis of variance (RM-ANOVA) to test for the effect of recreational shooting on behaviours over the three trapping sessions and analysed within- and between-subject effects with *post-hoc* comparisons. Prior to each RM-ANOVA, we tested data for normality and homoscedasticity (Zar 1999), transforming non-conforming data using arcsine square-root transformations. In addition, we used Mauchly’s *W*-test statistic to test for assumed sphericity (von Ende 2001).

In May–June 2003 (predisturbance), we placed trapping grids randomly on each colony; these same grids were re-established in July–August 2003 and May–June 2004 (post-disturbance). Live traps (Model 203, Tomahawk Live Traps, Tomahawk, WI, USA) were arranged in a 9 × 9 grid, 15 m apart. Paired colonies were trapped simultaneously for 6 consecutive days. During each session, traps were set, wired open and prebaited for 24 h, then baited and set at sunrise and checked and closed at sunset of each day’s trapping. We marked captured prairie dogs with fingerling ear tags (National

Band and Tag, Newport, KY, USA) and, because we observed during preliminary studies that males tended to lose ear tags, also injected males with subcutaneous passive integrated transponders (Biomark, Boise, ID, USA). We measured weight and hind foot length of each captured animal, using a ratio of the two measurements (g cm⁻¹) as an index of body condition (Krebs & Singleton 1993). Values for an individual were averaged within a trapping session. We tested for changes in the body conditions of juveniles and adults over the study period and between control and treatment colonies with an RM-ANOVA.

We collected scat samples from trapped prairie dogs in 2003 for analysis of faecal corticosterone concentration, a metric of physiological stress (Harper & Austad 2000). Although trapping and handling elevates corticosterone levels, faecal corticosterone levels reflect stress experienced about 6–12 h before defecation (Harper & Austad 2000). Therefore, we collected scat samples only from animals that had been held in traps < 5 h, and that had not been captured the previous day. Scat samples were air-dried in the field and stored at –20 °C. Corticosterone was extracted from faeces using methods described by Monfort *et al.* (1998) and assayed using a radioimmunoassay kit (ICN Biomedical Inc., Costa Mesa, CA, USA) and scintillation counter. The inter- and intra-assay coefficients of variation were 10% and 8%, respectively, and the detection threshold of the assay was 2.6 ng g⁻¹. We analysed variation in corticosterone levels for two age groups (adults and juveniles), before and after shooting, and on control and treatment colonies with two-way ANOVAS.

DENSITY, DEMOGRAPHY AND VITAL RATES

We determined the reproductive status of adult females in May and June at first capture. Females were considered to have been reproductively active earlier in the year if their nipples were enlarged and turgid. We entered capture histories into program CAPTURE to estimate juvenile, adult female and adult male abundances for each colony and trapping session. We estimated densities by dividing abundance estimates by the effective area trapped (Wilson & Anderson 1985), using half the mean maximum distance moved (averaging the maximum distance between recaptures for captured animals) and adding it to the perimeter of the trapping grid. Separate values were calculated for each age–sex group (juvenile, adult female, adult male), colony and trapping session, and variances for densities were calculated using the approach of Otis *et al.* (1978). We analysed changes in the densities of each prairie dog age–sex group over the study period and between control and treatment colonies with RM-ANOVA.

We evaluated changes in the proportions of juveniles, adult males and adult females in each colony with log-likelihood ratios. We compared the proportion of males and females that were yearlings and adults, and the proportional number of reproductively active females

Table 2. Mean percentage and standard error of prairie dogs foraging, alert and resting on control (\bar{x}_C) and treatment (\bar{x}_T) colonies from three sampling periods, Thunder Basin National Grassland, Wyoming, 2003–04. Prior to testing, percentages were converted using the arcsine transformation (Zar 1999) and compared between control and treatment colonies for each sampling period with paired *t*-tests (d.f. for all tests = 39). Treatment colonies were subjected to intensive shooting in June–July 2003

Date	Foraging				Alert				Resting									
	\bar{x}_C	SE	\bar{x}_T	SE	<i>t</i>	<i>P</i>	\bar{x}_C	SE	\bar{x}_T	SE	<i>t</i>	<i>P</i>	\bar{x}_C	SE	\bar{x}_T	SE	<i>t</i>	<i>P</i>
May–June 2003 (pre-disturbance)	76.2	1.3	78.7	1.3	-1.23	0.23	7.4	1.2	3.7	1.0	1.65	0.11	6.0	1.1	6.8	1.5	0.60	0.55
July–August 2003 (post-disturbance)	91.1	1.0	66.1	1.2	14.5	<0.001	3.2	0.7	29.2	1.1	-22.1	<0.001	2.7	0.9	0.7	1.1	3.55	0.01
May–June 2004 (post-disturbance)	74.7	1.4	71.0	1.4	1.20	0.24	10.3	1.1	15.6	1.2	-3.45	0.001	1.1	1.1	0.7	1.1	0.85	0.40

between control and treatment colonies in 2003 and 2004 with log-likelihood ratios, corrected for continuity. Indices of reproductive rate (juveniles \times adult female⁻¹) were compared between control and treatment colonies in 2003 and 2004 with a paired *t*-test.

Using the robust design model in program MARK (White & Burnham 1999), we modelled apparent survival (\hat{S}_i), temporary immigration (γ_i') and emigration (γ_i''), conditional capture (\hat{p}_{ij}) and recapture (\hat{e}_{ij}) probabilities. Parameter estimates in MARK are maximum likelihood estimates with 95% confidence intervals. To remove the estimates of population size from the likelihood, we used Huggins' estimator (Huggins 1991). We developed a series of 12 *a priori* models based on a number of factors that we believed would influence one or more of the parameter estimates. Factors that we hypothesized would affect parameter estimation included the intervals between trapping occasions, the dates of capture, age and sex of animals, colony, colony pair and experimental status (control, treatment). Because we assumed that temporary immigration and emigration were not occurring on any of the colonies, the parameters γ_i' and γ_i'' were set at 0.

We ranked models of survival rate using Akaike's information criterion corrected for small samples (AIC_c; Burnham & Anderson 2002). We ranked AIC_c values relative to the model with the lowest AIC_c value. Comparisons among models were made using Δ AIC_c (Burnham & Anderson 2002). Normalized Akaike weights (w_i) were also computed for each model as an additional assessment of the strength of evidence for each model.

Results

BEHAVIOUR, BODY CONDITION AND STRESS

Behaviour and time spent above ground were not influenced by the time (morning or evening) of observations (all $F_{2,39} < 0.70$, $P > 0.65$); therefore, morning and evening observations were pooled. Before shooting, prairie dog behaviour did not differ between control and treatment colonies (Table 2); most behaviour was foraging (76%), followed by resting (7%) and alertness (5%). Shooting caused several behaviour types to differ (foraging: $F_{2,39} = 39.5$, $P < 0.001$; resting: $F_{2,39} = 2.90$, $P = 0.033$; alertness:

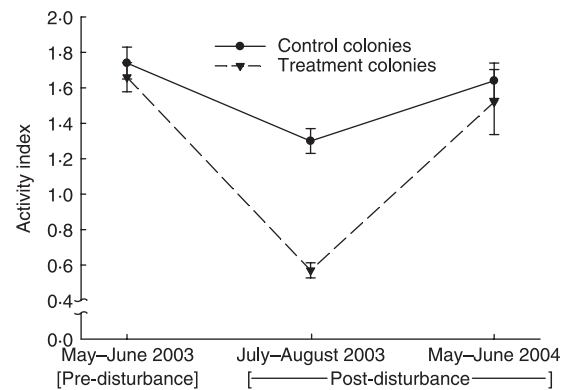


Fig. 1. Mean above-ground activity indices (± 1 SE) for black-tailed prairie dogs on control and treatment colonies, Thunder Basin National Grassland, Wyoming, 2003–04. Activity indices were calculated by dividing the number of prairie dogs above ground (prairie dogs ha⁻¹) by density estimates (prairie dogs ha⁻¹) for that colony. Prior to shooting, activity indices did not differ between control and treatment colonies. Following shooting in June–July 2003, activity on treatment colonies was lower than on control colonies. Above-ground activity on treatment colonies rebounded to control levels in 2004.

$F_{2,39} = 89.7$, $P < 0.001$) between control and treatment colonies. Shortly after shooting, alertness on treatment colonies increased to 29%, while on control colonies it decreased to 3% over the same period (Table 2). This coincided with reduced foraging (to 66%) and resting (to < 1%). By contrast, on control colonies foraging increased to 91%, while resting decreased slightly to 3% (Table 2). In 2004, the year after shooting, behaviours on treatment colonies returned to control levels for foraging (73%) and resting (1%), but alertness remained slightly elevated (Table 2). Time spent above ground was also affected by recreational shooting. Activity indices did not differ between control and treatment colonies prior to shooting ($t_{39} = 0.69$, $P = 0.50$), but diverged after shooting ($F_{2,39} = 5.95$, $P = 0.003$; Fig. 1). Above-ground activity on treatment colonies declined later the same summer by 66%, while corresponding values on control colonies declined by only 22% ($t_{39} = 8.72$, $P < 0.001$). The summer after shooting, above-ground activity recovered on treatment colonies, so that it did not differ ($t_{39} = 0.65$, $P = 0.52$; Fig. 1) from control-colony values.

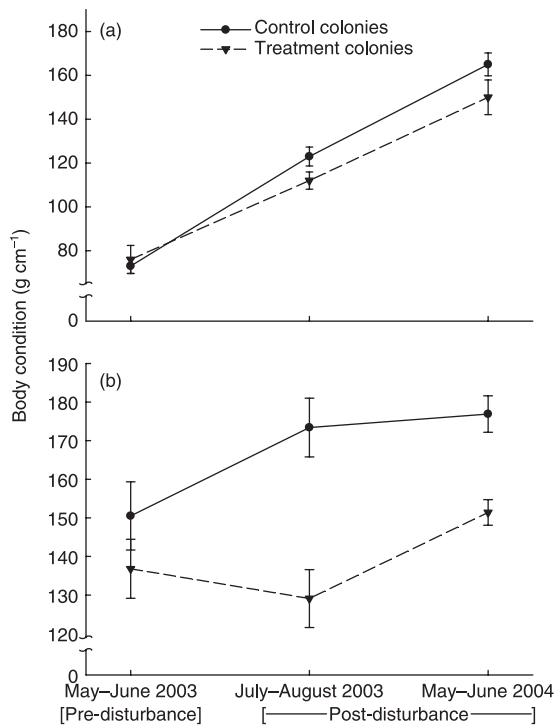


Fig. 2. Mean body condition (± 1 SE) for (a) juvenile and (b) adult black-tailed prairie dogs on control and treatment colonies over three sampling periods, Thunder Basin National Grassland, Wyoming, 2003–04. Treatment colonies were subjected to intensive recreational shooting in June–July 2003. On both control and treatment colonies, juveniles exhibited improved body condition over the course of the study. After shooting, adult body condition on treatment colonies declined, while adults on control colonies exhibited improved body condition. In 2004, adult body condition on treatment colonies partially recovered, but remained lower than on control colonies.

Recreational shooting did not affect body condition of surviving juveniles ($F_{2,56} = 2.37$, $P = 0.11$). On both control and treatment colonies, body condition of animals that were juveniles in 2003 increased approximately linearly from May–June 2003–04 ($F_{2,56} = 176.1$, $P < 0.001$; Fig. 2a). In contrast, recreational shooting altered the body condition of surviving adults ($F_{2,38} = 3.23$, $P = 0.050$; Fig. 2b). Before shooting, adult body condition did not differ between control and treatment colonies ($t_{18} = 0.46$, $P = 0.65$). After shooting, adult body condition decreased on treatment colonies and increased on control colonies, resulting in a 35% difference in body condition between control and treatment colonies later in the summer of shooting ($t_{18} = 2.78$, $P = 0.012$). By the following summer, body condition of adults on treatment colonies had recovered somewhat, to 17% below control values ($t_{18} = 3.00$, $P = 0.008$; Fig. 2b).

Stress responses to shooting depended on age (Fig. 3). For adults, recreational shooting did not affect stress; corticosterone levels were 10.8 ± 0.7 (± 1 SE) ng g^{-1} dry faeces for both control and treatment colonies, before and after shooting ($F_{1,27} < 0.001$, $P = 0.99$). Also for adults, corticosterone levels did not change from May–June 2003 to July–August 2003 ($F_{1,27} = 0.60$, $P = 0.45$). In

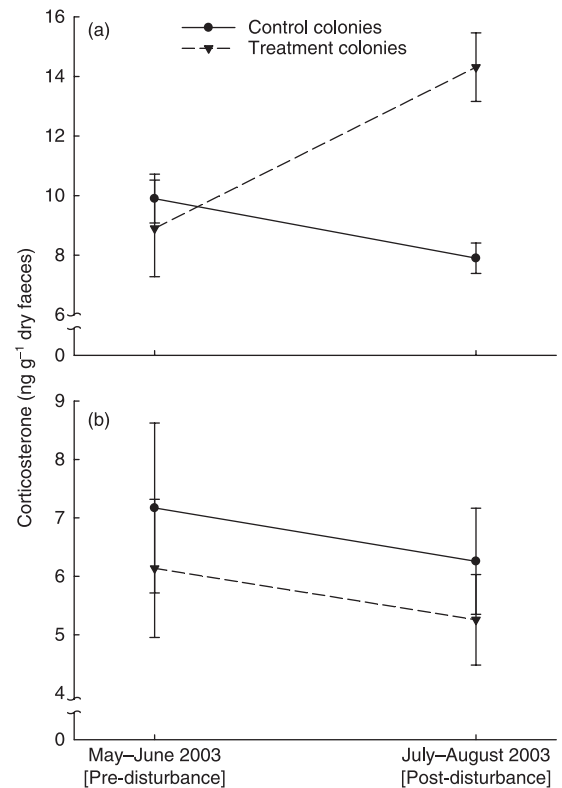


Fig. 3. Mean faecal corticosterone concentrations (± 1 SE) from (a) juvenile and (b) adult black-tailed prairie dogs on control and treatment colonies before and after treatment colonies were subjected to recreational shooting, Thunder Basin National Grassland, Wyoming.

contrast, juveniles exhibited elevated stress in response to shooting ($F_{1,41} = 7.42$, $P = 0.009$); before shooting, corticosterone levels did not differ between control and treatment colonies ($t_8 = 0.47$, $P = 0.65$). After shooting, however, values increased on treatment colonies to levels 80% higher than those on control colonies ($t_{33} = 4.95$, $P < 0.001$; Fig. 3).

DENSITY, DEMOGRAPHY AND VITAL RATES

Prior to shooting in 2003, prairie dog densities were similar between control and treatment colonies for juveniles ($t_3 = -0.33$, $P = 0.76$), adult females ($t_3 = 0.26$, $P = 0.81$) and adult males ($t_3 = 1.17$, $P = 0.33$; Fig. 4). Densities of all three age–sex groups on treatment and control colonies diverged following our shooting treatment (juveniles: $F_{2,7} = 11.83$, $P = 0.001$; adult females: $F_{2,7} = 4.54$, $P = 0.034$; adult males $F_{2,7} = 6.43$, $P = 0.013$; Fig. 4). After shooting in 2003, densities of juveniles born in 2003 did not differ between treatment and control colonies ($t_3 = -2.27$, $P = 0.11$), but 10 months later the 2004 cohort on treatment colonies was 85% smaller than that of the previous year, and 89% smaller than the 2004 cohort on control colonies ($t_3 = -4.96$, $P = 0.016$; Fig. 4a). Shooting reduced adult female densities by 40% in the short term, to values lower than those on control colonies at the same time ($t_3 = -3.40$, $P = 0.042$). Ten months later, densities of adult females on treatment

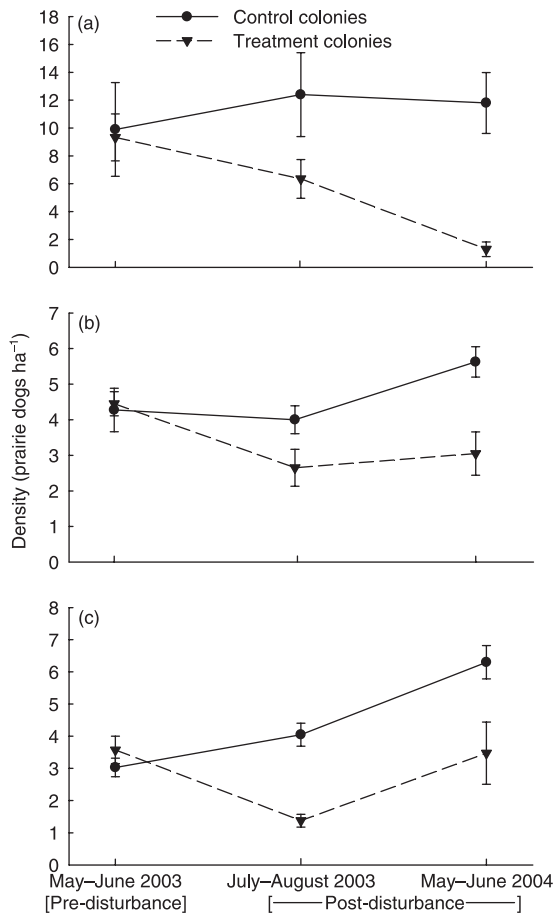


Fig. 4. Density estimates (± 1 SE) for (a) juvenile, (b) adult female and (c) adult male black-tailed prairie dogs on control and treatment colonies during three trapping occasions, Thunder Basin National Grassland, Wyoming, 2003–04. Treatment colonies were subjected to intensive recreational shooting in June–July 2003. For juveniles, shooting did not immediately reduce their densities, but 1 year later juvenile densities fell by 85%. Shooting reduced female densities by 40% in the short term and these densities remained lower 1 year later, in 2004. Adult males were most susceptible to shooting, exhibiting a 62% decline in densities, but were capable of rebounding to pre-disturbance and control-level densities in 2004.

colonies remained marginally lower than on control colonies ($t_3 = -2.99$, $P = 0.058$; Fig. 4b). For adult males the pattern differed; densities declined by 62% in short-term response to shooting ($t_3 = -9.05$, $P = 0.003$), but recovered by 10 months later, when densities did not differ from those on control colonies ($t_3 = -2.23$, $P = 0.11$; Fig. 4c). Changes in the density of each age-sex group is explained partially by the selectivity of shooting: juveniles represented 59% of animals shot, a value proportional to their abundance (Appendix S2), while adult females were the least vulnerable, composing only 15% of animals shot, but 22% of the colony members. Adult males were the most susceptible age-sex group, constituting 26% of animals shot, but only 19% of the animals on the colony (Appendix S2).

Age structure did not differ between control and treatment colonies before shooting ($G_2 = 2.46$, $P = 0.88$); juveniles dominated age-sex groups (53% of animals),

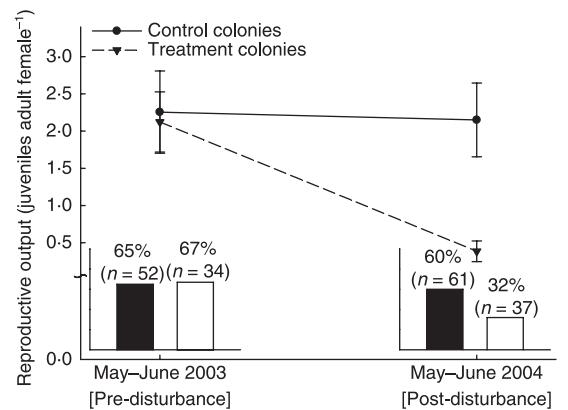


Fig. 5. Mean reproductive output (± 1 SE; main graph) and pregnancy rates (inset bar graphs) for adult female black-tailed prairie dogs on control (solid shading, inset bar graphs) and treatment colonies (open shading, inset bar graphs), Thunder Basin National Grassland, Wyoming, 2003–04. After shooting treatment colonies in June–July 2003, pregnancy rates fell by 50% and reproductive output fell by 82%.

followed by adult females (27%) and adult males (20%). Immediately after shooting, age structure did not differ between control and treatment colonies ($G_2 = 3.10$, $P = 0.21$). However, 10 months after shooting, age structure had changed ($G_2 = 11.7$, $P = 0.003$), with juveniles (born in 2004) less prevalent on treatment colonies (16% vs. 49% on control colonies). The proportion of the population that was adult female (39% treatment, 24% control) and adult male (45% treatment, 27% control) increased, reflecting the major decline in proportional abundance of juveniles.

Before shooting, age structure for prairie dogs > 1 year of age did not differ between control and treatment colonies for females ($G_1 = 2.47$, $P = 0.12$) or males ($G_1 = 0.14$, $P = 0.71$); 31% of adult females and 30% of adult males were yearlings. One year after shooting, the percentage of yearling females increased (53%), but did not differ between control and treatment colonies ($G_1 = 0.01$, $P = 0.94$). For males in 2004, however, age structure diverged between control and treatment colonies ($G_1 = 4.89$, $P = 0.027$); on control colonies, the percentage of yearlings increased to 54%, while on treatment colonies the percentage of yearlings more than doubled, to 76%.

In 2003, reproductive output was 2.2 juveniles \times adult female⁻¹, similar between control and treatment colonies ($t_3 = -0.41$, $P = 0.71$; Fig. 5). The summer following shooting reproductive output remained similar on control colonies, but fell by 82%, to 0.4 juveniles \times adult female⁻¹, on treatment colonies ($t_3 = 3.33$, $P = 0.044$; Fig. 5). This reduction reflected a concomitant decline in pregnancy rates, which did not differ between control and treatment colonies in 2003 ($G_1 < 0.01$, $P = 0.96$); 66% of females showed evidence of recent reproduction (Fig. 5). In 2004, however, 60% of females on control colonies gave birth, in comparison with just 32% of females on treatment colonies ($G_1 = 6.62$, $P = 0.010$; Fig. 5).

Table 3. Twelve a priori models from survival analyses using program MARK (White & Burnham 1999) for black-tailed prairie dogs from eight paired study colonies, Thunder Basin National Grasslands, captured in Wyoming, 2003 and 2004. We modelled capture (\hat{p}) and recapture probabilities (\hat{c}) from five variables: age (juvenile, adult), date (date of capture or recapture), session (primary trapping session), colony (the eight colonies) and treatment (control vs. treatment colonies). We modelled prairie dog survival (\hat{S}) from six variables: interval (period between trapping sessions), age (juvenile, adult), age/sex (juvenile, adult male, adult female), block (colony pairs), shot-acute (shooting effects on treatment colonies from May–August 2003) and shot-chronic (shooting effects on treatment colonies from August 2003 to May 2004). For all models temporary emigration (γ') and immigration (γ'') were set at 0 and therefore not depicted below. Models were ranked using Akaike's information criterion (AIC_c). Also provided with each model are Δ AIC_c (the difference between the best model's AIC_c and subsequent models), w_i (weight of evidence in favour of each model), K (number of parameters) and log(L) (maximum log-likelihood)

Model	AIC _c	Δ AIC _c	w_i	K	log(L)
\hat{S} (interval, age, shot-acute) \hat{p} (age,date*session, colony,treatment) \hat{c} (age,date*session,colony)	7952.6	0	0.410	29	-3945.9
\hat{S} (interval,age/sex,shot-acute) \hat{p} (age,date*session, colony,treatment) \hat{c} (age,date*session,colony)	7954.2	1.6	0.184	30	-3945.6
\hat{S} (interval,age,shot-acute,shot-chronic) \hat{p} (age,date* session,colony,treatment) \hat{c} (age,date*session,colony)	7954.6	2.0	0.151	30	-3945.8
\hat{S} (interval,age,block,shot-acute) \hat{p} (age,date* session,colony,treatment) \hat{c} (age,date*session,colony)	7955.5	2.9	0.095	32	-3944.0
\hat{S} (interval,age/sex,shot-acute,shot-chronic) \hat{p} (age,date *session,colony,treatment) \hat{c} (age,date*session,colony)	7956.3	3.7	0.064	31	-3945.5
\hat{S} (interval,age/sex,block,shot-acute) \hat{p} (age,date* session,colony,treatment) \hat{c} (age,date*session,colony)	7957.2	4.7	0.040	33	-3943.8
\hat{S} (interval,age,block,shot-acute,shot-chronic) \hat{p} (age,date* session,colony,treatment) \hat{c} (age,date*session,colony)	7957.2	4.7	0.040	33	-3943.8
\hat{S} (interval,age/sex,block,shot-acute,shot-chronic) \hat{p} (age,date* session,colony,treatment) \hat{c} (age,date*session,colony)	7959.1	6.5	0.016	34	-3943.6
\hat{S} (interval,age,block) \hat{p} (age,date*session,colony, treatment) \hat{c} (age,date*session,colony)	7968.4	15.8	0.000	31	-3951.6
\hat{S} (interval,age) \hat{p} (age,date*session,colony,treatment) \hat{c} (age,date*session,mix*adult,colony)	7968.0	15.4	0.000	28	-3954.7
\hat{S} (interval,age/sex,block) \hat{p} (age,date*session,colony,treatment) \hat{c} (age,date*session,colony)	7970.3	17.7	0.000	32	-3951.4
\hat{S} (interval,age/sex) \hat{p} (age,date*session,colony,treatment) \hat{c} (age,date*session,colony)	7969.8	17.2	0.000	29	-3954.5

Table 4. Survival estimates and associated standard errors for black-tailed prairie dog age–sex groups in Thunder Basin National Grassland, Wyoming, 2003–04. Modelling was performed in program MARK. Treatment colonies were subjected to a pulse of recreational shooting in June–July 2003; control colonies were protected from shooting. Shown, in order, are estimates from the three most competitive models ranked by AIC_c. For details on models and model ranking see Table 3

Age–sex group	June–August 2003				August 2003–June 2004			
	Control		Treatment		Control		Treatment	
	\hat{S}	SE	\hat{S}	SE	\hat{S}	SE	\hat{S}	SE
Model 1								
Juvenile	0.725	0.047	0.438	0.06	0.423	0.042	–*	–*
Adult	0.677	0.055	0.383	0.06	0.368	0.044	–*	–*
Model 2								
Juvenile	0.726	0.047	0.438	0.06	0.423	0.042	–*	–*
Adult male	0.647	0.069	0.351	0.07	0.337	0.057	–*	–*
Adult female	0.697	0.06	0.404	0.07	0.390	0.054	–*	–*
Model 3								
Juvenile	0.727	0.048	0.438	0.06	0.450	0.071	0.414	0.045
Adult	0.678	0.055	0.381	0.06	0.392	0.069	0.358	0.048

*Survival estimates and standard error did not differ from those computed for control colonies.

Variation in survival was best explained by a model (Table 3) that included parameters for trapping sessions, age–sex group and shooting (Δ AIC_c = 0.0, w_i = 0.410; Table 3). Estimates of survival from June to August

2003 on control colonies were 0.73 for juveniles and 0.68 for adults. Corresponding estimates for treatment colonies (juvenile: 0.44; adult: 0.38) were about 30% lower (Table 4). Overwinter survival (August 2003–

June 2004), however, was not affected by shooting; our best model estimated juvenile survival as 0.42 and adult survival as 0.37 for both control and treatment colonies (Table 4).

We found less, and approximately equal, support for two other models (Table 3) that, in addition to accounting for short-term effects of shooting, also (1) differentiated survival rates among sexes ($\Delta AIC_c = 1.6$, $w_i = 0.184$); and (2) accounted for shooting reducing overwinter survival on treatment colonies ($\Delta AIC_c = 2.0$, $w_i = 0.151$). However, both models appeared suboptimal because log-likelihood estimates remained similar even though the number of estimable parameter increased (Table 3). Further, uncertainty associated with the survival estimates from these suboptimal models resulted in considerable overlap between survival estimates, making their utility ambiguous (Table 4). As predicted, models that did not account for changes in survival from shooting fitted our data poorly (ΔAIC_c range = 15.8–17.2, all $w_i < 0.0001$; Table 3).

Discussion

Prairie dogs subjected to hunting exhibited risk-disturbance effects that overwhelmed and obscured any possible density-dependent ones. As observed in some other hunted vertebrates, surviving prairie dogs altered behaviour to reduce their vulnerability to shooting at the expense of other fitness-enhancing activities. They increased their alertness and decreased above-ground activity, time spent foraging and time spent resting above ground. In general, these behavioural responses were transient, returning to preshooting and control-colony levels the summer after shooting. However, alert behaviours remained slightly elevated on treatment colonies 10 months later. This persistently elevated alertness may reflect a confounding influence, however. Loughry (1992) showed that adult prairie dogs exhibited higher alertness than juveniles, so the increase in alertness we observed could be attributable to changes in age structure. Although qualitatively similar behavioural responses have been reported for hunted waterfowl (Madsen & Fox 1995; Féret *et al.* 2003) and ungulates (Kilgo *et al.* 1998), prairie dog responses to shooting were particularly dramatic.

Behavioural responses of prairie dogs to shooting may have been influenced by factors in addition to risk-avoidance. First, prairie dogs exhibit complex social networks and deaths due to shooting could have disrupted important social interactions and individual behaviour. Indeed, Shier (2006) showed that among translocated prairie dogs, group cohesion strongly affected foraging efficiency: intact family groups foraged more and were vigilant less than those composed of mostly unrelated individuals. Shier (2006) found that these changes extended to survival and reproductive output, as did we. Secondly, in contrast to more traditional forms of hunting, recreational shooting involves many rounds fired over hours (Table 1; Vosburgh & Irby 1998). Such a disturbance contributed presumably

to the dramatic behavioural responses we observed in prairie dogs. Thus, the behavioural sensitivity of prairie dogs to shooting seems a combination of their coloniality, low mobility and sensitivity to social disruption as well as the duration and intensity of the disturbance.

Reduced foraging and above-ground activity resulted in lowered body condition of surviving adult prairie dogs, which had body conditions 35% poorer than those of control animals. Although adult body condition on treatment colonies improved the following year, it remained 17% lower than on control colonies. So, effects of shooting on body condition via foraging persisted into the next growing season for adults. In contrast, shooting did not affect the body condition of surviving juveniles, for reasons that are suggested by our stress-response data. Unlike adults, juveniles exhibited a distinct stress-response to shooting, but no response of body condition. We hypothesize that juveniles tended to remain above ground and forage during and after shooting, unlike adults. Prairie dogs survive winter on somatic stores (Lehmer & Van Horne 2001), and with especially limited stores, juveniles must increase body mass rapidly in the first summer of life to survive their first winter (Rayor 1985). Therefore, juveniles may need to forage during times of increased risk, including during shooting, more than do adults, thereby exposing them to disturbance and resulting physiological stress (Lima 1998). Despite the adaptiveness of corticosterone in managing short-term stressors, chronically high levels can lead to reproductive failure, immune suppression and poor body condition (Sapolsky 1992). Therefore, chronically elevated corticosterone levels could cascade to reduced survival and recruitment.

Such cascades have been shown for other hunted species. Féret *et al.* (2003) and Roy & Woolf (2001) found that game birds subjected to intensive hunting increased the time spent flying and decreased the time spent foraging, with negative consequences for body condition. Hjeljord & Histøl (1999) found that the body mass of moose (*Alces alces*) in Norway was correlated negatively with hunting intensity. The changes reported by these studies were less dramatic than those we describe here for prairie dogs. Again, the coloniality and low mobility of prairie dogs and the intense and disruptive nature of recreational shooting seem to account for this difference.

Age-sex groups differed in their vulnerabilities and responses to shooting. Adult females were not particularly susceptible to shooting, but failed to recover to preshooting densities the summer after the treatment. Conversely, adult males were susceptible to shooting, but recovered to preshooting and control-colony densities the following summer. Juvenile proportional abundances were not affected in the short term, but were reduced by 85% the following summer. Juveniles, the dominant age group (54%) on colonies before shooting, were the smallest age group (< 20%) the following year. This delayed effect, of shooting on juvenile abundance the following summer, was mediated largely

by reproduction. On colonies subjected to shooting, reproductive output decreased by 82% from 2003 to 2004, whereas control colonies showed little change over the same period. The proportion of females producing young on treatment colonies declined by ~0.5. Knowles (1987) found a fairly constant 0.66 of females giving birth, similar to values from our control colonies.

Two mechanisms could explain reduced pregnancy and reproductive rates on treatment colonies. First, reduced body condition of surviving adult females could have reduced litter production. We found that the body condition of prairie dogs > 1 year of age decreased by 35% in response to shooting. Somatic stores are important for both survival and reproduction among ground squirrels (Woods & Armitage 2003), and female prairie dogs must accumulate sufficient reserves to reproduce successfully (Seabloom & Theisen 1990). At northern latitudes, black-tailed prairie dogs copulate in February–March (Hoogland 1995), when prairie dogs have the lowest body masses and are most stressed nutritionally (Lehmer & Van Horne 2001). Therefore, adult females on colonies subjected to shooting could have been physiologically unable to carry a litter to parturition, as a result of foraging opportunities lost (Knowles 1987). Secondly, changes in age–sex structure could have reduced the availability of reproductively competent males. The summer after shooting abundance of males rebounded, but ages were lower than before shooting; most mature males had been replaced by presumably immigrant, yearling males. As in larger mammals (Wielgus & Bunnell 1994), yearling male prairie dogs are less competent breeders than older males (Seabloom & Theisen 1990). This form of Allee effect, resulting from reduced availability to breeding-age females of competent adult males (Stephens & Sutherland 2000), could be an important consequence of male-biased hunting, particularly among polygynous, colonial animals for which females are mate-selective (Halliday 1983). Thus, nutritional deficiencies in adult females or the influx of yearling males after shooting – or both – might have contributed to the reproductive near-collapse that we observed the summer following shooting.

Surprisingly, shooting did not improve overwinter survival. If anything, we saw evidence for decreased survival among treatment animals. One top-ranked model accounted for reduced overwinter survival on treatment colonies, but those survival estimates overlapped considerably with values for control colonies. None the less, the importance of incorporating a parameter for shooting was apparent; models that accounted for an effect of shooting in reducing survival were competitive, while models that did not do so performed poorly.

COMPARATIVE ASPECTS

Generally, populations of small-bodied mammalian herbivores, such as lagomorphs (Rose 1977) and squir-

rels (Mosby 1969), recover quickly from hunting via density-dependent vital rates. In contrast, prairie dogs in our study showed no evidence of density dependence in overwinter survival or next-year natality. Rather, we found that for prairie dogs, hunting induced not only additive effects on survival, but also led to reproductive near-collapse the summer following the shooting. We attribute these remarkable effects to costs represented by shooting to a highly colonial species: shooting reduced the usual benefits and increased some of the common costs of coloniality. In unperturbed colonies, animals benefit from cooperative vigilance, allowing a relaxation of individual vigilance, which increases foraging efficiency and reduces susceptibility to predation (Hoogland 1981). Coloniality also facilitates adult reproduction and juvenile survival through enhanced mate-finding and cooperative breeding (Jennions & MacDonald 1994). These common benefits of coloniality, however, appeared to be reduced by recreational shooting. Intense, prolonged shooting increased alertness at the expense of foraging, so that a primary benefit of coloniality was reduced. Coloniality presumably facilitates reproduction by providing easy access to mates, and some communal care of neonates. However, the effects that we observed suggest a shift in the net fitness costs and benefits of coloniality.

Among costs, coloniality causes all above-ground animals to be disturbed by a single shot that is fired; animals that escape to below ground forego foraging opportunities. Under most circumstances, access to mature male mates is not limiting to female prairie dogs (Hoogland & Foltz 1982), but shooting selectively removed adult males, so that female access to competent mates may have been reduced. This may have contributed to the reproductive near-collapse the summer after shooting. Thus, coloniality appears to make prairie dogs more sensitive to hunting than other small mammalian herbivores.

CONSERVATION IMPLICATIONS

Prairie dogs are ecologically pivotal members of North American grassland systems (Miller *et al.* 1994), supporting predators, including the obligate prairie dog predator, the endangered black-footed ferret (*Mustela nigripes*). Prairie dog burrows also provide habitat structure for burrowing owls (*Athene cunicularia*), prairie rattlesnakes (*Crotalus viridis*) and various small mammals. Through herbivory, prairie dogs alter vegetation and cycle nutrients (Holland & Detling 1990). Therefore, the population biology of prairie dogs and the viability of their colonies have broad implications for North American grassland communities. Future research needs to elucidate the large-scale effects of shooting on prairie dog populations and colony viability as well as its effects on other species that depend on prairie dog colonies. Wildlife managers should consider measures to reduce recreational shooting intensity and duration in regions where black-tailed prairie dog colony growth

and persistence is desired, such as recovery sites for the black-footed ferret, yet allow shooting to continue in areas where colonies conflict with private-landowner interests.

Acknowledgements

This study was funded by the Bureau of Land Management, Wyoming State Office and through research grants from the Wildlife Heritage Foundation of Wyoming and the National Wildlife Federation. Logistical support was provided by the Casper, Wyoming Field Office of the Bureau of Land Management. E. A. Van Kirk, University of Wyoming, provided assistance in corticosterone analyses. S. J. Converse, G. C. White and D. R. Anderson, Colorado State University, provided advice on modelling prairie dog survival and abundance. We are grateful to field and laboratory technicians and volunteers: J. T. Schultz, R. D. Mares, M. B. Hamilton, S. Hansen, A. M. Roe, M. Carter, T. Roberts, B. P. Pauli and E. Donadio. J. L. Beck, K. G. Gerow, D. B. McDonald and G. K. Brown provided constructive comments on early drafts. Finally, we thank the ranching families and mining companies of north-eastern Wyoming, who granted us access to their land. In particular, we thank Antelope Coal Mine, Jacobs Ranch Mine, the Simmonses, Tracys, Gordons and Keelines. Permits for fieldwork were obtained through the Wyoming Game and Fish Department; all methods were approved by the Institutional Animal Care and Use Committee at the University of Wyoming and adhered to the ethical guidelines for the use of animals in research set forth by the British Ecological Society.

References

- Bateson, P. & Bradshaw, E.L. (1997) Physiological effects of hunting red deer (*Cervus elaphus*). *Proceedings of the Royal Society of London, Series B*, **263**, 1707–1714.
- Béchet, A., Giroux, J.-F., Gauthier, G., Nichols, J.D. & Hines, J.E. (2003) Spring hunting changes the regional movements of migrating greater snow geese. *Journal of Applied Ecology*, **40**, 553–564.
- Blumstein, D.T. & Pelletier, D. (2005) Yellow-bellied marmot hiding time is sensitive to variation in costs. *Canadian Journal of Zoology*, **83**, 363–367.
- Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference: a Practical Information-Theoretical Approach*, 2nd edn. Springer-Verlag, New York, NY.
- Cox, M.K. & Franklin, W.L. (1990) Premolar gap technique for aging live black-tailed prairie dogs. *Journal of Wildlife Management*, **54**, 143–146.
- von Ende, C.N. (2001) Repeated-measures analysis: growth and other time-dependent measures. *Design and Analysis of Ecological Experiments*, 2nd edn (eds S.M. Scheiner & J. Gurevitch), pp. 134–157. Oxford University Press, New York, NY.
- Féret, M., Gauthier, G., Béchet, A., Giroux, J.-F. & Hobson, K.A. (2003) Effect of a spring hunt on nutrient storage by greater snow geese in southern Quebec. *Journal of Wildlife Management*, **67**, 796–807.
- FitzGibbon, C. (1998) The management of subsistence harvesting: behavioral ecology of hunters and their mammalian prey. *Behavioral Ecology and Conservation Biology* (ed. T. Caro), pp. 449–473. Oxford University Press, New York, NY.
- Fowler, C.W. (1987) A review of density dependence in populations of large mammals. *Current Mammalogy* (ed. H. H. Genoways), pp. 401–441. Plenum Press, New York, NY.
- Frid, A. & Dill, L.M. (2002) Human-caused disturbance stimuli as a form of predation risk. *Conservation Ecology*, **6**, 11. Available at: <http://www.consecol.org/vol6/iss1/art11> (accessed 1 January 2005).
- Gill, J.A., Norris, K. & Sutherland, W.J. (2001) Why behavioural responses may not reflect the population consequences of human disturbance. *Biological Conservation*, **97**, 265–268.
- Halliday, T.R. (1983) The study of mate choice. *Mate Choice* (ed. P. Bateson), pp. 3–32. Cambridge University Press, Cambridge, UK.
- Harper, J.M. & Austad, S.N. (2000) Fecal glucocorticoids: a noninvasive method of measuring adrenal activity in wild and captive rodents. *Physiological and Biochemical Zoology*, **73**, 12–22.
- Hjeljord, O. & Histøl, T. (1999) Range–body mass interactions of a northern ungulate – a test of a hypothesis. *Oecologia*, **119**, 326–339.
- Holland, E.A. & Detling, J.K. (1990) Plant response to herbivory and below ground nitrogen cycling. *Ecology*, **71**, 1040–1049.
- Hoogland, J.L. (1979) Aggression, ectoparasitism, and other possible costs of prairie dog (Sciuridae, *Cynomys* spp.) coloniality. *Behaviour*, **69**, 1–35.
- Hoogland, J.L. (1981) The evolution of coloniality in white-tailed and black-tailed prairie dogs (Sciuridae: *Cynomys leucurus* and *C. ludovicianus*). *Ecology*, **62**, 252–272.
- Hoogland, J.L. (1995) *The Black-Tailed Prairie Dog: Social Life of a Burrowing Mammal*. University of Chicago Press, Chicago, IL.
- Hoogland, J.L. & Foltz, D.W. (1982) Variance in male and female reproductive success in a harem-polygynous mammal, the black-tailed prairie dog (Sciuridae: *Cynomys ludovicianus*). *Behavioral Ecology and Sociobiology*, **11**, 155–163.
- Hoogland, J.L. & Sherman, P.W. (1976) Advantages and disadvantages of bank swallow (*Riparia riparia*) coloniality. *Ecological Monographs*, **46**, 33–58.
- Horn, H.S. (1966) Measurement of 'overlap' in comparative ecological studies. *American Naturalist*, **100**, 419–424.
- Huggins, R.M. (1991) Some practical aspects of a conditional likelihood approach to capture experiments. *Biometrics*, **47**, 725–732.
- Jennions, M.D. & Macdonald, D.W. (1994) Cooperative breeding in mammals. *Trends in Ecology and Evolution*, **9**, 89–93.
- Kilgo, J.C., Labisky, R.F. & Fritzen, D.E. (1998) Influences of hunting on the behavior of white-tailed deer: implications for conservation of the Florida panther. *Conservation Biology*, **12**, 1359–1364.
- Knowles, C.J. (1982) *Habitat affinity, populations, and control of black-tailed prairie dogs on the Charles M. Russell National Wildlife Refuge*. MS Thesis, University of Montana.
- Knowles, C.J. (1987) Reproductive ecology of black-tailed prairie dogs in Montana. *Great Basin Naturalist*, **47**, 202–206.
- Kokko, H. (2001) Optimal and suboptimal use of compensatory responses to harvesting: timing of hunting as an example. *Wildlife Biology*, **7**, 141–150.
- Krebs, C.J. (1999) *Ecological Methodology*, 2nd edn. Addison-Wesley Longman, Inc., Menlo Park, CA.
- Krebs, C.J. & Singleton, G.R. (1993) Indices of condition for small mammals. *Australian Journal of Zoology*, **41**, 317–323.
- Lehmer, E.M. & Van Horne, B. (2001) Seasonal changes in lipids, diet, and body condition of free-ranging black-tailed prairie dogs (*Cynomys ludovicianus*). *Canadian Journal of Zoology*, **79**, 955–965.

- Lima, S.L. (1998) Stress and decision making under the threat of predation: recent developments from behavioral, reproductive, and ecological perspectives. *Advances in the Study of Behavior*, **27**, 215–290.
- Loughry, W.J. (1992) Ontogeny of time allocation in black-tailed prairie dogs. *Ethology*, **90**, 206–224.
- Madsen, J. (1998) Experimental refuges for migratory waterfowl in Danish wetlands. II. Tests of hunting disturbance effects. *Journal of Applied Ecology*, **35**, 398–417.
- Madsen, J. & Fox, A.D. (1995) Impacts of hunting disturbance on waterbirds – a review. *Wildlife Biology*, **1**, 193–207.
- Mainguy, J., Bêty, J., Gauthier, G. & Giroux, J.-F. (2002) Are body condition and reproductive effort of laying greater snow geese affected by the spring hunt? *Condor*, **104**, 156–161.
- Martin, P. & Bateson, P. (1993) *Measuring Behaviour: an Introductory Guide*, 2nd edn. Cambridge University Press, Cambridge, UK.
- Miller, B., Ceballos, G. & Reading, R. (1994) The prairie dog and biotic diversity. *Conservation Biology*, **8**, 677–681.
- Monfort, S.L., Mashburn, K.L., Brewer, B.A. & Creel, S.R. (1998) Evaluating adrenal activity in African wild dogs (*Lycaon pictus*) by fecal corticosteroid analysis. *Journal of Zoo and Wildlife Medicine*, **29**, 129–133.
- Mosby, H.S. (1969) The influence of hunting on the population dynamics of a woodlot grey squirrel population. *Journal of Wildlife Management*, **33**, 59–73.
- Otis, D.L., Burnham, K.P., White, G.C. & Anderson, D.R. (1978) Statistical inference from capture data on closed animal populations. *Wildlife Monographs*, **62**, 1–135.
- Pauli, J.N., Buskirk, S.W., Williams, E.S. & Edwards, W.H. (2006) A plague epizootic in the black-tailed prairie dog (*Cynomys ludovicianus*). *Journal of Wildlife Diseases*, **42**, 74–80.
- Powell, K.L., Robel, R.J., Kemp, K.E. & Nellis, M.D. (1994) Aboveground counts of black-tailed prairie dogs: temporal nature and relationship to burrow entrance density. *Journal of Wildlife Management*, **58**, 361–366.
- Rayor, L.S. (1985) Effects of habitat quality on growth, age of first reproduction, and dispersal in Gunnison's prairie dogs (*Cynomys gunnisoni*). *Canadian Journal of Zoology*, **63**, 2835–2840.
- Reeve, A.F. & Vosburgh, T.C. (2005) Recreational shooting of prairie dogs. *Conservation of the Black-Tailed Prairie Dog* (ed. J.L. Hoogland), pp. 139–156. Island Press, Washington, DC.
- Ripple, W.J. & Beschta, R.L. (2004) Wolves and the ecology of fear: can predation risk structure ecosystems? *Bioscience*, **54**, 755–766.
- Rose, G.B. (1977) Mortality rates of tagged adult cottontail rabbits. *Journal of Wildlife Management*, **41**, 511–514.
- Roy, C. & Woolf, A. (2001) Effects of hunting and hunting-hour extension on mourning dove foraging and physiology. *Journal of Wildlife Management*, **65**, 808–815.
- Sapolsky, R.M. (1992) Neuroendocrinology of the stress-response. *Behavioral Endocrinology* (eds J.B. Becker, S.M. Breedlove & D. Crews), pp. 287–324. Massachusetts Institute of Technology Press, Cambridge, MA.
- Seabloom, R.W. & Theisen, P.W. (1990) Breeding biology of the black-tailed prairie dog in North Dakota. *Prairie Naturalist*, **22**, 65–74.
- Shier, D.M. (2006) Effect of family support on success of translocated black-tailed prairie dogs. *Conservation Biology*, **20**, 1780–1790.
- Stephens, P.A., Frey-Roos, F., Arnold, W. & Sutherland, W.J. (2002) Sustainable exploitation of social species: a test and comparison of models. *Journal of Applied Ecology*, **39**, 629–642.
- Stephens, P.A. & Sutherland, W.J. (2000) Vertebrate mating systems, Allee effects and conservation. *Vertebrate Mating Systems* (eds M. Apollonio, M. Festa-Bianchet & D. Mainardi), pp. 186–213. World Scientific Publishing, London, UK.
- Vosburgh, T.C. & Irby, L.R. (1998) Effects of recreational shooting on prairie dog colonies. *Journal of Wildlife Management*, **62**, 363–372.
- White, G.C. & Burnham, K.P. (1999) Program MARK: survival estimation from populations of marked animals. *Bird Study*, **46** (Suppl.), 120–138.
- Wielgus, R.B. & Bunnell, F.L. (1994) Dynamics of a small, hunted brown bear *Ursus arctos* population in southwestern Alberta, Canada. *Biological Conservation*, **67**, 161–166.
- Wilson, K.R. & Anderson, D.R. (1985) Evaluation of two density estimators of small mammal population size. *Journal of Mammalogy*, **66**, 13–21.
- Woods, B.C. & Armitage, K.B. (2003) Effect of food supplementation on juvenile growth and survival in *Marmota flaviventris*. *Journal of Mammalogy*, **84**, 903–914.
- Zar, J.H. (1999) *Biostatistical Analysis*, 4th edn. Prentice Hall, Inc., Upper Saddle River, NJ.

Received 29 November 2006; final copy received 9 April 2007
Editor: E. J. Milner-Gulland

Supplementary material

The following supplementary material is available for this article.

Appendix S1. Attributes of 10 black-tailed prairie dog colonies, Thunder Basin National Grassland, Wyoming, summer 2003.

Appendix S2. Mean standardized selection indices and associated standard errors for prairie dog cohorts shot on treatment colonies, Thunder Basin National Grassland, Wyoming, 2003.

Fig. S1. (a) Locations of 10 black-tailed prairie dog colonies used in an experimental study of recreational shooting around Thunder Basin National Grassland (TBNG), Wyoming, 2003–2004. (b) Cluster tree depicting the relative similarity of colonies using the unweighted pair group method with arithmetic averages.

This material is available as part of the online article from: <http://www.blackwell-synergy.com/doi/full/10.1111/j.1365-2664.2007.01337.x>
(This link will take you to the article abstract.)

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Copyright of *Journal of Applied Ecology* is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.



Trophic Downgrading of Planet Earth

James A. Estes *et al.*
Science **333**, 301 (2011);
DOI: 10.1126/science.1205106

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 19, 2013):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/333/6040/301.full.html>

Supporting Online Material can be found at:

<http://www.sciencemag.org/content/suppl/2011/07/13/333.6040.301.DC1.html>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/content/333/6040/301.full.html#related>

This article **cites 63 articles**, 11 of which can be accessed free:

<http://www.sciencemag.org/content/333/6040/301.full.html#ref-list-1>

This article has been **cited by** 36 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/333/6040/301.full.html#related-urls>

This article appears in the following **subject collections**:

Ecology

<http://www.sciencemag.org/cgi/collection/ecology>

Trophic Downgrading of Planet Earth

James A. Estes,^{1*} John Terborgh,² Justin S. Brashares,³ Mary E. Power,⁴ Joel Berger,⁵ William J. Bond,⁶ Stephen R. Carpenter,⁷ Timothy E. Essington,⁸ Robert D. Holt,⁹ Jeremy B. C. Jackson,¹⁰ Robert J. Marquis,¹¹ Lauri Oksanen,¹² Tarja Oksanen,¹³ Robert T. Paine,¹⁴ Ellen K. Pikitch,¹⁵ William J. Ripple,¹⁶ Stuart A. Sandin,¹⁰ Marten Scheffer,¹⁷ Thomas W. Schoener,¹⁸ Jonathan B. Shurin,¹⁹ Anthony R. E. Sinclair,²⁰ Michael E. Soulé,²¹ Risto Virtanen,²² David A. Wardle²³

Until recently, large apex consumers were ubiquitous across the globe and had been for millions of years. The loss of these animals may be humankind's most pervasive influence on nature. Although such losses are widely viewed as an ethical and aesthetic problem, recent research reveals extensive cascading effects of their disappearance in marine, terrestrial, and freshwater ecosystems worldwide. This empirical work supports long-standing theory about the role of top-down forcing in ecosystems but also highlights the unanticipated impacts of trophic cascades on processes as diverse as the dynamics of disease, wildfire, carbon sequestration, invasive species, and biogeochemical cycles. These findings emphasize the urgent need for interdisciplinary research to forecast the effects of trophic downgrading on process, function, and resilience in global ecosystems.

The history of life on Earth is punctuated by several mass extinction events (2), during which global biological diversity was sharply reduced. These events were followed by novel changes in the evolution of surviving species and the structure and function of their ecosystems. Our planet is presently in the early to middle stages of a sixth mass extinction (3), which, like those before it, will separate evolutionary winners from losers. However, this event differs from those that preceded it in two fundamental ways: (i) Modern extinctions are largely being caused by a single species, *Homo sapiens*, and (ii) from its onset in the late Pleistocene, the sixth mass extinction has been characterized by the loss of larger-bodied animals in general and of apex consumers in particular (4, 5).

The loss of apex consumers is arguably humankind's most pervasive influence on the natural world. This is true in part because it has occurred globally and in part because extinctions are by their very nature perpetual, whereas most other environmental impacts are potentially reversible on decadal to millennial time scales. Recent research suggests that the disappearance of these animals reverberates further than previously anticipated (6–8), with far-reaching effects on processes as diverse as the dynamics of disease; fire; carbon sequestration; invasive species; and biogeochemical exchanges among Earth's soil, water, and atmosphere.

Here, we review contemporary findings on the consequences of removing large apex consumers from nature—a process we refer to as trophic downgrading. Specifically, we highlight the ecological theory that predicts trophic downgrading, consider why these effects have been difficult to observe, and summarize the key empirical evidence for trophic downgrading, much of which has appeared in the literature since the beginning of the 21st century. In

“What escapes the eye ... is a much more insidious kind of extinction: the extinction of ecological interactions”

Daniel H. Janzen (1)

so doing, we demonstrate the influence of predation and herbivory across global ecosystems and bring to light the far-reaching impacts of trophic downgrading on the structure and dynamics of these systems. These findings suggest that trophic downgrading acts additively and synergistically with other anthropogenic impacts on nature, such as climate and land use change, habitat loss, and pollution.

Foundations in Theory

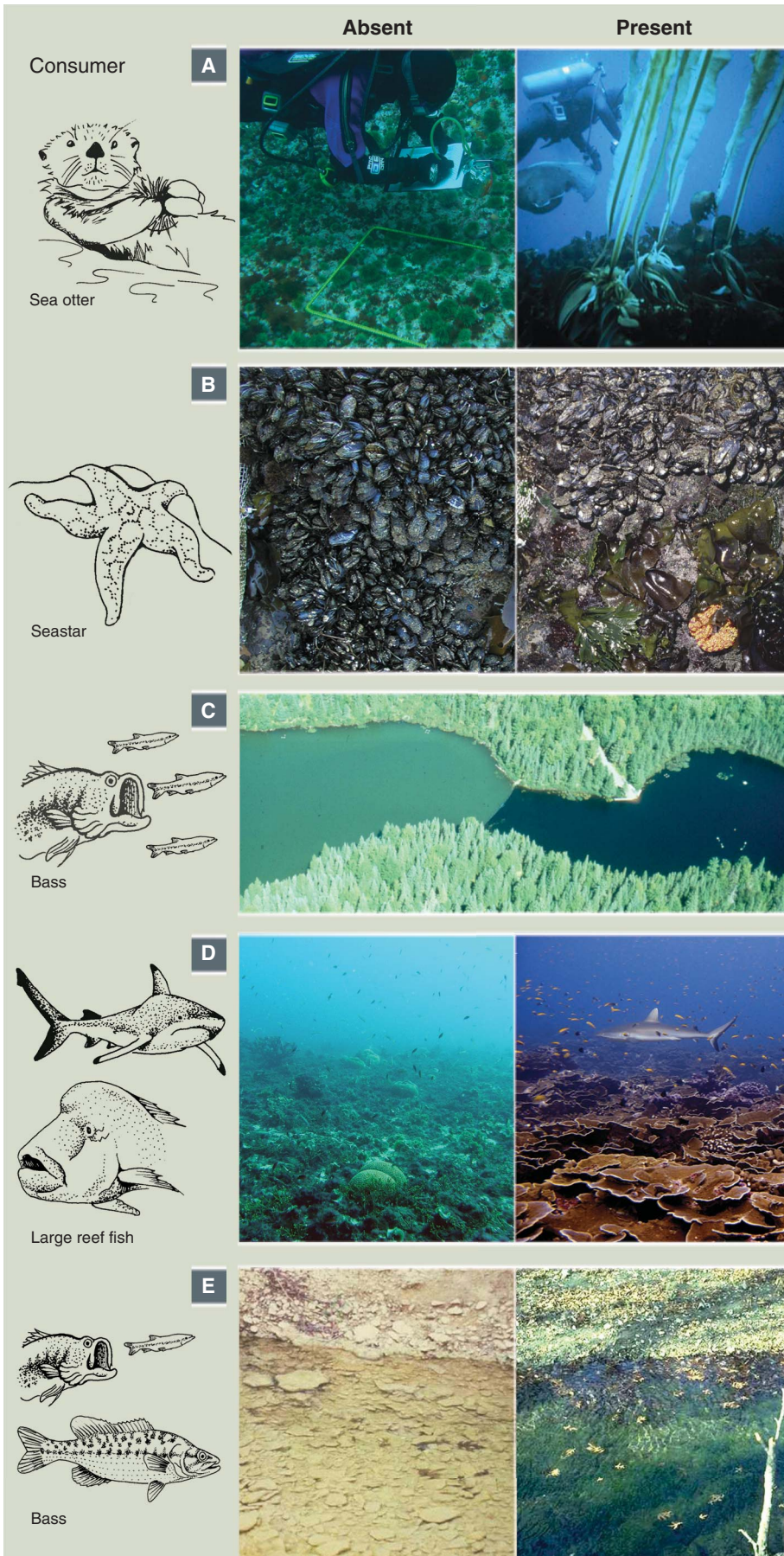
Ecological theory has long predicted that major shifts in ecosystems can follow changes in the abundance and distribution of apex consumers (9, 10). Three key elements of that theory provide the foundation for interpreting recurrent patterns suggestive of trophic downgrading in more recent empirical work across ecosystems. First is the idea that an ecosystem may be shaped by apex consumers, which dates back more than a century but was popularized in the 1960s (9). This concept was later formalized as the dynamic notion of “trophic cascades,” broadly defined as the propagation of impacts by consumers on their prey downward through food webs (11). Theoretical work on factors that control ecosystem state resulted in a second key advance, the recognition of “alternative stable states.” The topology of ecosystem dynamics is now understood to be nonlinear and convoluted, resulting in distinct basins of attraction.

Alternative stable states occur when perturbations of sufficient magnitude and direction push ecosystems from one basin of attraction to another (12). Tipping points (also known as thresholds or breakpoints), around which abrupt changes in ecosystem structure and function (a.k.a. phase shifts) occur, often characterize transitions between alternative stable states. Ecosystem phase shifts can also display hysteresis, a phenomenon in which the locations of tipping points between states differ with the directionality of change (13). A third key concept, connectivity, holds that ecosystems are built around interaction webs within which every species potentially can influence many other species. Such interactions, which include both biological processes (e.g., predation, competition, and mutualism) and physicochemical processes (e.g., the nourishing or limiting influences of water, temperature, and nutrients), link species together at an array of spatial scales (from millimeters to thousands of kilometers) in a highly complex network.

Taken together, these relatively simple concepts set the stage for the idea of trophic downgrading.

¹Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95060, USA. ²Center for Tropical Conservation, Nicholas School of the Environment and Earth Sciences, Post Office Box 90381, Duke University, Durham, NC 27708, USA. ³Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720, USA. ⁴Department of Integrative Biology, Valley Life Sciences Building, University of California, Berkeley, CA 94720, USA. ⁵Division of Biological Sciences, University of Montana, Missoula, MT 59812, USA and Wildlife Conservation Society, Bozeman, MT 59715, USA. ⁶Botany Department, University of Cape Town, Private Bag, Rondebosch 7701, South Africa. ⁷Center for Limnology, 680 North Park Street, University of Wisconsin, Madison, WI 53706, USA. ⁸School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195, USA. ⁹Department of Biology, Post Office Box 118525, University of Florida, Gainesville, FL 32611, USA. ¹⁰Center for Marine Biodiversity and Conservation, Scripps Institution of Oceanography, 9500 Gilman Drive, La Jolla, CA 92093, USA. ¹¹Department of Biology, University of Missouri–St. Louis, One University Boulevard, St. Louis, MO 63121, USA. ¹²Department of Biology, Section of Ecology, University of Turku, FI-20014 Turku, Finland and Department of Natural Sciences, Finnmark University College, N-9509 Alta, Norway. ¹³Department of Biology, Section of Ecology, University of Turku, FI-20014 Turku, Finland and Department of Ecology and Environmental Science, Umeå University, SE-90087 Umeå, Sweden. ¹⁴Department of Biology, Box 351800, University of Washington, Seattle, WA 98195, USA. ¹⁵School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY 11794, USA. ¹⁶Department of Forest Ecosystems and Society, 314 Richardson Hall, Oregon State University, Corvallis, OR 97331, USA. ¹⁷Aquatic Ecology and Water Quality Management Group, Department of Environmental Sciences, Wageningen University, Post Office Box 8080, 6700 DD Wageningen, Netherlands. ¹⁸Section of Evolution and Ecology and Center for Population Biology, 6328 Storer Hall, University of California, Davis, CA 95616, USA. ¹⁹Department of Zoology, University of British Columbia, 6270 University Boulevard, Vancouver, BC V6T 1Z4, Canada. ²⁰Centre for Biodiversity Research, 6270 University Boulevard, University of British Columbia, Vancouver, BC V6T 1Z4, Canada. ²¹Post Office Box 1808, Paonia, CO 81428, USA. ²²Department of Biology, University of Oulu, FI-90014 Oulu, Finland. ²³Department of Forest Ecology and Management, Faculty of Forestry, Swedish University of Agricultural Sciences, SE901-83 Umeå, Sweden.

*To whom correspondence should be addressed. E-mail: jastes@ucsc.edu



The loss of apex consumers reduces food chain length, thus altering the intensity of herbivory and the abundance and composition of plants in largely predictable ways (10). The transitions in ecosystems that characterize such changes are often abrupt, are sometimes difficult to reverse, and commonly lead to radically different patterns and pathways of energy and material flux and sequestration.

The Cryptic Nature of Trophic Downgrading

The omnipresence of top-down control in ecosystems is not widely appreciated because several of its key components are difficult to observe. The main reason for this is that species interactions, which are invisible under static or equilibrium conditions, must be perturbed if one is to witness and describe them. Even with such perturbations, responses to the loss or addition of a species may require years or decades to become evident because of the long generation times of some species. Adding to these difficulties is the fact that populations of large apex consumers have long been reduced or extirpated from much of the world. The irony of this latter situation is that we often cannot unequivocally see the effects of large apex consumers until after they have been lost from an ecosystem, at which point the capacity to restore top-down control has also been lost. Another difficulty is that many of the processes associated with trophic downgrading occur on scales of tens to thousands of square kilometers, whereas most empirical studies of species interactions have been done on small or weakly motile species

Fig. 1. Landscape-level effects of trophic cascades from five selected freshwater and marine ecosystems. **(A)** Shallow seafloor community at Amchitka Island (Aleutian archipelago) before (1971; photo credit: P. K. Dayton) and after (2009) the collapse of sea otter populations. Sea otters enhance kelp abundance (right) by limiting herbivorous sea urchins (left) (20). **(B)** A plot in the rocky intertidal zone of central California before (September 2001, right) and after (August 2003, left) seastar (*Pisaster ochraceus*) exclusion. *Pisaster* increases species diversity by preventing competitive dominance of mussels. [Photo credits: D. Hart] **(C)** Long Lake (Michigan) with largemouth bass present (right) and experimentally removed (left). Bass indirectly reduce phytoplankton (thereby increasing water clarity) by limiting smaller zooplanktivorous fishes, thus causing zooplankton to increase and phytoplankton to decline (26). **(D)** Coral reef ecosystems of uninhabited Jarvis Island (right, unfished) and neighboring Kiritimati Island (left, with an active reef fishery). Fishing alters the patterns of predation and herbivory, leading to shifted benthic dynamics, with the competitive advantage of reef-building corals and coralline algae diminished in concert with removal of large fish (66). **(E)** Pools in Brier Creek, a prairie margin stream in south-central Oklahoma with (right) and lacking (left) largemouth and spotted bass. The predatory bass extirpate herbivorous minnows, promoting the growth of benthic algae (67).

Downloaded from www.sciencemag.org on August 19, 2013

with short generation times that could be manipulated at small spatial scales. Although some influences of apex consumers (e.g., trophic cascades) seen in experiments scale up to systems with larger or more mobile species (14), others are harder to discern at small spatial and temporal scales (e.g., many of the indirect effects of trophic cascades on ecosystem processes described below). As a result, we have an incomplete and distorted picture of the influences of apex consumers across much of the natural world.

The Widespread Occurrence of Trophic Cascades

Despite these challenges, trophic cascades have now been documented in all of the world's major biomes—from the poles to the tropics and in terrestrial, freshwater, and marine systems (table S1). Top-down forcing and trophic cascades often have striking effects on the abundance and species composition of autotrophs, leading to regime shifts and alternative states of ecosystems (15). When the impacts of apex consumers are reduced or removed or when systems are examined over sufficiently large scales of space and time, their influences are often obvious (Figs. 1 and 2). Although purposeful manipulations have produced the most statistically robust evidence, “natural experiments” (i.e., perturbations caused by population declines, extinctions, reintroductions, invasions, and various forms of natural resource management) corroborate the essential role of top-down interactions in structuring ecosystems involving species such as killer whales (*Orcinus orca*) (16), lions (*Panthera leo*) (17), wolves (*Canis lupus*) and cougars (*Puma concolor*) (18), the great sharks (19), sea otters (*Enhydra lutris*) (20), diverse mesopredators (21), and megaherbivores (22). Although the extent and quality of evidence differs among species and systems, top-down effects over spatial scales that are amenable to experimentation have proven robust to alternative explanations (23).

The impacts of trophic cascades on communities are far-reaching, yet the strength of these impacts will likely differ among species and ecosystems. For example, empirical research in Serengeti, Tanzania, showed that the presence or absence of apex predators had little short-term effect on resident megaherbivores [elephant (*Loxodonta africana*), hippopotamus (*Hippopotamus amphibius*), and rhinoceros (*Diceros bicornis*)] because these herbivores were virtually invulnerable to predation (24). Conversely, predation accounted for nearly all mortality in smaller herbivores [oribi (*Ourebia ourebi*), Thompson's gazelle (*Eudorcas thomsonii*), and impala (*Aepyceros melampus*)], and these species showed dramatic increases in abundance and distribution after the local extinction of predators. Thus, top-down forcing in this system is more apparent in some species than others, at least when it is studied on relatively short time scales, although the aggregate ecological impact of apex consumers here, as elsewhere, remains great (24).

Other than the inclusion of top-down forcing, there is no rule of thumb on the interplay between apex consumers and autotrophs in intact ecosys-

tems. This is largely a consequence of natural variation in food chain length (10). In some cases, the influence of apex consumers is to suppress herbivory and to increase the abundance and production of autotrophs. The sea otter/kelp forest system in the North Pacific Ocean (20) (Fig. 1A) and the wolf/ungulate/forest system in temperate and boreal North America (25) (Fig. 2C) function in this manner. Apex consumers in other

systems reduce the abundance and production of autotrophs. The largemouth bass/planktivore/zooplankton/phytoplankton system in U.S. Midwestern lakes (26) (Fig. 1C) functions in such a manner.

Effects on Ecosystem Processes

Apart from small oceanic islands, all regions of our planet supported a megafauna before the

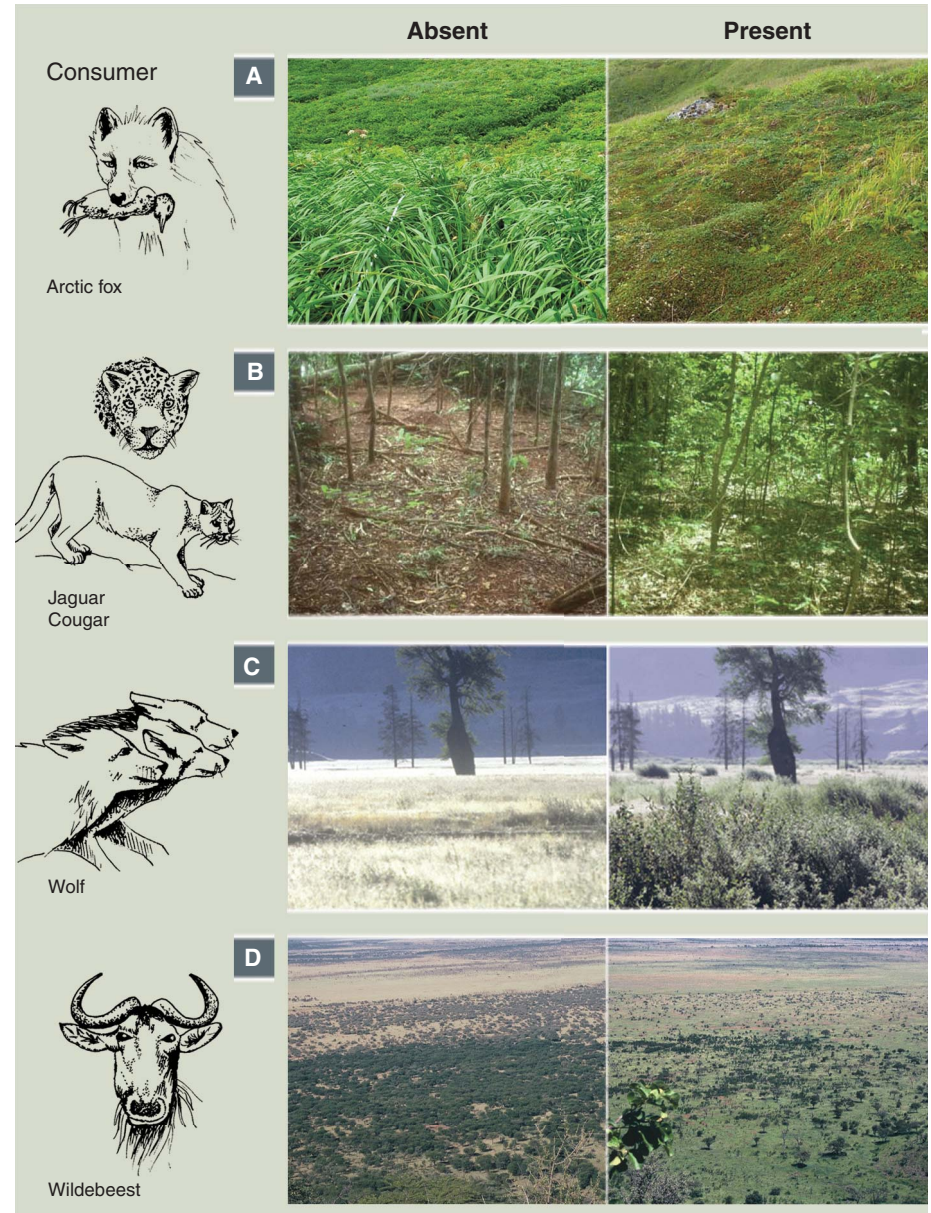


Fig. 2. Landscape-level effects of trophic cascades from four terrestrial ecosystems. **(A)** Upland habitat of islands with (right) and without (left) Arctic foxes in the Aleutian archipelago. Foxes drive terrestrial ecosystems from grasslands to tundra by limiting seabirds and thereby reducing nutrient inputs from sea to land (47). **(B)** Venezuelan forests on small islands of Lago Guri (left: jaguar, cougar, and harpy eagles absent) and mainland forest (right, predators present). A diverse herbivore guild erupted with the loss of predators from the island, thereby reducing plant recruitment and survival (68). **(C)** Riparian habitat near the confluence of Soda Butte Creek with the Lamar River (Yellowstone National Park) illustrating the stature of willow plants during suppression (left, 1997) from long-term elk browsing and their release from elk browsing (right, 2001) after wolf reintroductions of 1995 and 1996 (25). **(D)** Decline of woody vegetation in Serengeti after eradication of rinderpest (by early 1960s) and the recovery of native ungulates (by middle 1980s). Left, 1986; right, 2003 (69).

rise of *Homo sapiens* (4, 27). The apex consumers influence their associated ecosystems through top-down forcing and trophic cascades, which in turn often lead to myriad effects on other species and ecosystem processes (Figs. 3 and 4). Here, we describe some of the known or suspected indirect effects of losing these apex consumers.

Herbivory and wildfire. Wildfires burn up to 500 million ha of the global landscape annually, consuming an estimated 8700 Tg of dry plant biomass, releasing roughly 4000 Tg of carbon to the atmosphere, and costing billions of dollars in fire suppression and property loss (28). The frequency and extent of wildfire have been largely attributed to a warming and drying climate and fuel accumulation from protective wildland management practices. However, the global distribution and biomass of vegetation are poorly predicted by temperature and rainfall (29), and recent analyses suggest that interdependencies among predation (including disease), herbivory, plant communities, and fire may better explain the dynamics of vegetation. Such interdependencies are well illustrated in East Africa, where the introduction of rinderpest in the late 1800s decimated many native ungulate populations, including wildebeest (*Connochaetes taurinus*) and buffalo (*Syncerus caffer*). Reductions of these large herbivores caused an increase in plant biomass, which fueled wildfires during the dry season. Rinderpest was eliminated from East Africa in the 1960s through an extensive vaccination and control program. Because of this, wildebeest and buffalo populations had recovered to what was thought to be historically high levels by the early 1980s. The resulting increase in herbivory drove these systems from shrublands to grasslands, thus decreasing the fuel loads and reducing the frequency and intensity of wildfires (30) (Fig. 4). Other examples of the interplay between megafauna and wildfire are the increase in fire frequency after the late Pleistocene/early Holocene decline of megaherbivores in Australia (31) and the northeastern United States (32).

Disease. The apparent rise of infectious diseases across much of the globe is commonly attributed to climate change, eutrophication, and habitat deterioration. Although these factors are undoubtedly important, links also exist between disease and predation (33). For example, the reduction of lions and leopards from parts of sub-Saharan Africa has led to population outbreaks and changes in behavior of olive baboons (*Papio anubis*). The baboons, in turn, have been drawn into increasing contact with people because of their attraction to crops and other human food resources. The increased baboon densities and their expanded interface with human populations have led to higher rates of intestinal parasites in baboons and the humans who live in close proximity to them (17). A similar result, involving different species and processes, occurred in India, where the decline of vultures also led to increased health risks from rabies and anthrax (34). Further examples of the interplay between predation and disease exist for

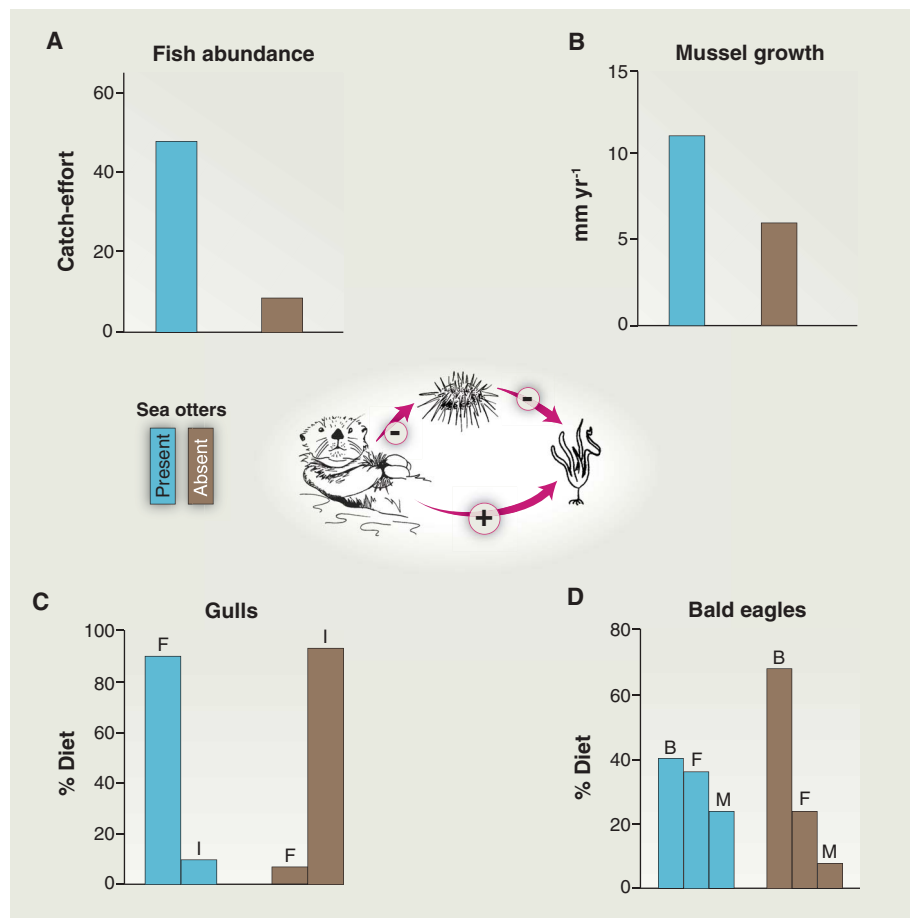


Fig. 3. Trophic cascade from sea otters to sea urchins to kelp (center) has myriad effects on other species and ecological processes. The increase in kelp enhances the abundance of kelp forest fish (A) (70). Enhanced kelp production increases the amount of particulate organic carbon in coastal ocean waters, thus increasing the growth rate of filter-feeding mussels (B) (71). The presence or absence of sea otters influences the diet of other consumers in the coastal ecosystems (C and D). In systems with abundant sea otters, Glaucous winged-gulls (*Larus glaucescens*) consume mostly fish (F), whereas in systems lacking sea otters, gulls consume mostly macroinvertebrates (I) (C) (72). When sea otters were abundant in the Aleutian archipelago, bald eagles (*Haliaeetus leucocephalus*) fed on fish (F), mammals (M), and birds (B) in roughly equal amounts. The loss of sea otters from this system led to a stronger reliance by the eagles on seabirds (D) (73). Blue bars from system with sea otters; brown bars from system without sea otters.

aquatic systems. The establishment of no-take marine reserves in the Channel Islands of southern California led to increases in the size and abundance of spiny lobsters (*Panulirus interruptus*) and declines in population densities of sea urchins, which are preyed on by the lobsters. The reduced urchin densities thwarted the spread of disease among individual sea urchins, which led to a lowered frequency of epidemics of sea urchin wasting disease within the reserves (35) (Fig. 4). In freshwater systems, the localized rise and fall of human malaria is associated with the impacts of predatory fishes on planktivores, which are in turn important consumers of mosquito larvae (36).

Physical and chemical influences. The influences of industrialization and agriculture on Earth's physical environments and geochemical processes are widely known. However, the contributing effects of changes in the distribution and abundance of apex consumers to the physical and chemical nature of our biosphere—the atmosphere,

soils, and water—are understudied and largely unappreciated. Even so, important connections between these entities have become apparent in the few instances where people have looked.

The atmosphere. Linkages between apex consumers and the atmosphere are known or suspected in freshwater, marine, and terrestrial ecosystems. Trophic cascades associated with the presence or absence of apex predatory fishes in lakes can affect phytoplankton density, in turn affecting the rate of primary production, the uptake rate of CO₂, and the direction of carbon flux between lakes and the atmosphere. Where apex predatory fishes are present in sufficient numbers, they reduce the abundance of smaller planktivorous minnows, thus releasing zooplankton from limitation by planktivores and increasing consumption rates of phytoplankton by zooplankton (Fig. 1B). This trophic cascade causes lakes to switch from net sinks for atmospheric CO₂ when predatory fishes are absent to net sources of atmospheric CO₂ when

these fishes are present (37) (Fig. 4). Similar processes occur in the oceans and on land. Industrial whaling during the 20th century transferred some 105 million tons of carbon from great whales to the atmosphere (38), and even today whale feces return various limiting nutrients from the aphotic to photic zones, thereby directly enhancing primary productivity (39, 40) and its influence on carbon flux and sequestration. From land, the demise of Pleistocene megaherbivores may have contributed to or even largely accounted for the reduced atmospheric methane concentration and the resulting abrupt 9°C temperature decline that defines the Younger-Dryas period (41).

Soils. Leaf-eating herbivores profoundly influence soils and their associated biota through altered plant allocation patterns of carbon and nutrients to the roots and rhizosphere, changing the quantity and quality of litter that plants return to the soil. Ungulate herbivores further influence soils through trampling, compaction, and the return of dung and urine. The collective influence of these processes is often an effect on species composition of the vegetation and altered successional pathways (42, 43). Predators of these herbivores and the trophic cascades they set in motion reverse these belowground effects (44). For example, the reintroduction of wolves to Yellowstone National Park has reduced the positive indirect effects of ungulates on soil nitrogen mineralization and potentially the nitrogen supply for plant growth (45). In contrast, introduced rats (46) and arctic foxes (Fig. 4) (47) have reduced soil fertility and plant nutrition on high-latitude islands by disrupting seabirds and their sea-to-land nutrient subsidies, with striking effects on plant community composition.

Water. Large consumers influence the composition and quality of both fresh and salt water through a variety of mechanisms. For example, the collapse of large demersal fish led to a 20% reduction in silica supply to pelagic diatoms in the Baltic Sea (48). In rivers, mass spawning by salmon suspends sediments, thus increasing downstream sediment transport (49) (Fig. 4). This flushing of stream bed sediments by the spawning fish and the increased circulation of fresh water through the gravel interstices of the stream bed have positive feedbacks on salmon populations by increasing oxygen for incubating eggs and fry and decreasing the frequency with which bed-mobilizing floods kill salmon in these early life stages (50). Similarly, in terrestrial systems wolves protect riparian trees and shrubs from over-browsing by large ungulates, in turn shading and cooling the adjacent streams, reducing stream bank erosion, and providing cover for fish and other aquatic life (51, 52).

Invasive species. A common feature of many successful invasive species is that they have left behind their natural predators and freed themselves from top-down control (53). Likewise, the loss of native predators leaves ecosystems more vulnerable to invasion by nonnative species (54). There are many examples of hypersuccessful invasions due to the absence or loss of top-down

control in aquatic and terrestrial systems. The experimental exclusion of native birds from small areas in Hawaii resulted in an up to 80-fold increase in nonnative spider density (55) (Fig. 4). Other examples include the spread of the invasive brown tree snake (*Boiga irregularis*) on the otherwise vertebrate predator-free island of Guam (56), the facilitating influence of reduced fish predation on the invasion of zebra mussels (*Dreissena polymorpha*) in the Mississippi River (57), and reduced abundance and spread of the introduced European green crab (*Carcinus maenas*) by predation from native blue crabs (*Callinectes sapidus*) in eastern North America (58).

Biodiversity. Earth's biodiversity (defined here as both species diversity and the associated func-

tional diversity) is increasingly confined to formal protected areas. Although the establishment of protected areas mitigates certain threats to biodiversity—habitat loss and fragmentation, over-exploitation, and the spread of invasive species—when large apex consumers are missing, protected areas often fail to function as intended. The link between apex consumers and species diversity can occur via a number of interaction pathways, for example, by blocking competitive exclusion [predatory seastars in the rocky intertidal (59)], mesopredator release [coyotes (*Canis latrans*) maintaining small vertebrate species in chaparral habitats (Fig. 4) (60)], and indirect habitat effects [e.g., the loss of small vertebrates from over-grazed and degraded riparian habitats after the

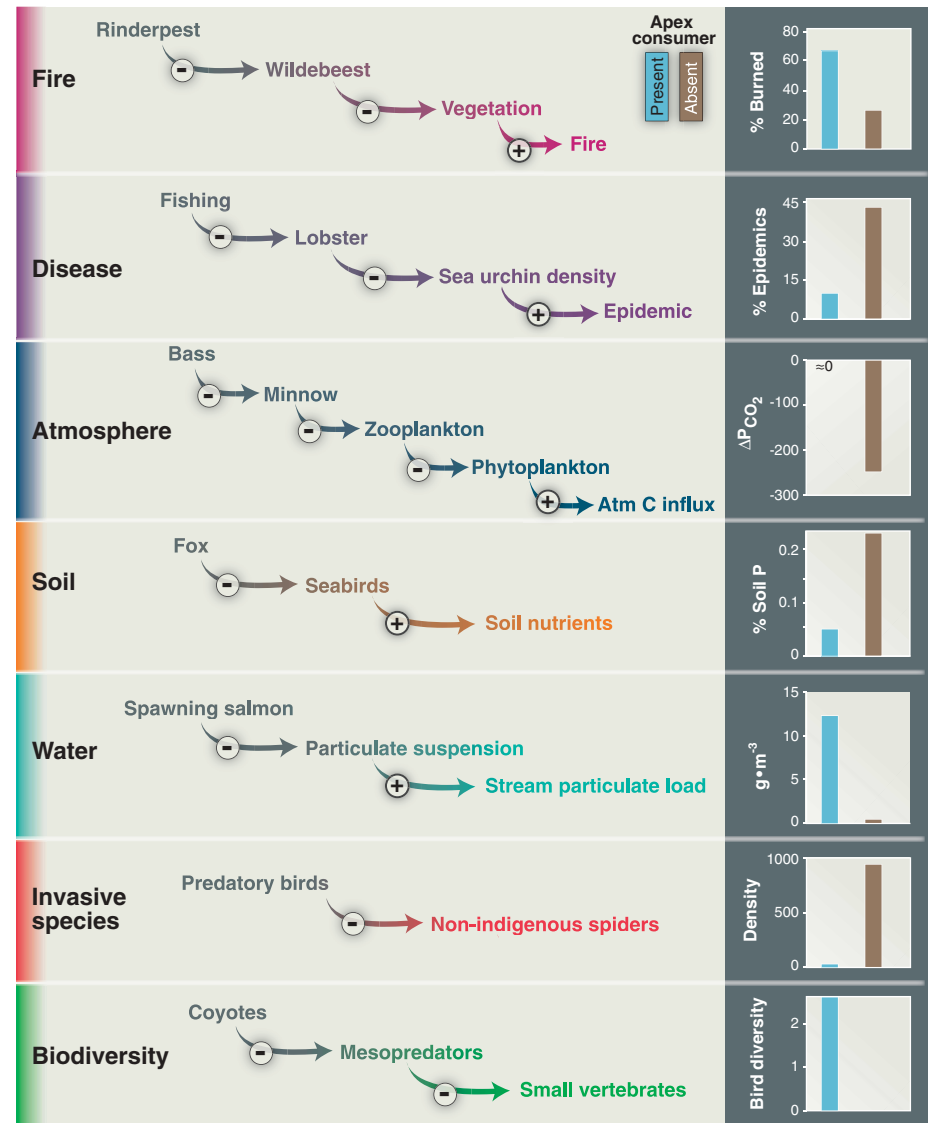


Fig. 4. Examples of the indirect effects of apex consumers and top-down forcing on diverse ecosystem processes, including wildfires (30); disease (35); composition of atmosphere (37), soil (47), and fresh water (49); invadability by exotic species (55); and species diversity (60). Interaction web linkages by which these processes are connected to apex consumers are shown in the center. Magnitude of effect is shown in graphs on right. Blue bars are data from systems containing the apex consumer; brown bars are data from systems lacking the apex consumer. Data replotted from original sources (cited above), except raw data on native bird diversity in chaparral habitats provided by K. Crooks.

loss of cougars (61) or wolves and grizzly bears (*Ursus arctos*) (62) from temperate and boreal forests of western North America].

Tree recruitment failure and the eventual transformation of forests to heaths and grasslands because of increased ungulate herbivory illustrates the influence of large apex consumers on functional diversity. This process is most clearly seen by contrasting areas where apex consumers have been absent for differing lengths of time. In North America, where wolves and other large carnivores were not extirpated until the early 20th century, the effects of their loss on plants is evident only as the recruitment failure of the younger trees. Because of the longevity of adult trees, the older individuals persist in what superficially appears to be a normally functioning forest ecosystem. These effects are best known from various U.S. National Parks, where the loss of large predators a few decades ago has left a characteristic signal of reduced tree growth rate (63) or recruitment failure (64) in the dominant tree species. A longer time horizon can be obtained from the Canadian island of Anticosti, where white-tailed deer (*Odocoileus virginianus*) have persisted in the absence of predators for more than a century, causing the successive elimination of saplings of less and less palatable trees and shrubs and increasing graminoid dominance in the understory (65). The Scottish island of Rùm, from which wolves have been absent for 250 to 500 years, provides a view of the likely final outcome of predator loss and elevated herbivory in many temperate forests. Rùm has transitioned over this same period from a forested environment to a treeless island.

These examples support the conclusion that disruptions of trophic cascades due to the decline of predation constitute a threat to biodiversity from within for which the best management solution is likely the restoration of effective predation regimes.

A Paradigm Shift in Ecology

The accumulation of theoretical and empirical evidence calls for an altered perspective on top-down forcing in ecosystem dynamics. Many practicing ecologists still view large animals in general, and apex consumers in particular, as ecological passengers riding atop the trophic pyramid but having little impact on the structure below. The influences of these animals, although acknowledged in particular cases, are generally regarded as anomalous, occurring in some systems but not in many others. This perception has generally led to the requirement of independent study and confirmation for each species and system before the null hypothesis that they serve no important ecological role can be rejected. We argue that the burden of proof be shifted to show, for any ecosystem, that consumers do (or did) not exert strong cascading effects.

Conclusions

Unanticipated changes in the distribution and abundance of key species have often been attributed in some unspecified manner to the “complexity of nature.” We propose that many of the

ecological surprises that have confronted society over past centuries—pandemics, population collapses of species we value and eruptions of those we do not, major shifts in ecosystem states, and losses of diverse ecosystem services—were caused or facilitated by altered top-down forcing regimes associated with the loss of native apex consumers or the introduction of exotics. Our repeated failure to predict and moderate these events results not only from the complexity of nature but from fundamental misunderstandings of their root causes. Except for controlling predators to enhance fish, wild game, and livestock, resource managers commonly base their actions on the assumption that physical processes are the ultimate driver of ecological change. Bottom-up forces are ubiquitous and fundamental, and they are necessary to account for the responses of ecosystems to perturbations, but they are not sufficient. Top-down forcing must be included in conceptual overviews if there is to be any real hope of understanding and managing the workings of nature.

References and Notes

1. D. H. Janzen, *Nat. Hist.* **83**, 48 (1974).
2. D. M. Raup, J. J. Sepkoski Jr., *Science* **215**, 1501 (1982).
3. D. B. Wake, V. T. Vredenburg, *Proc. Natl. Acad. Sci. U.S.A.* **105** (suppl. 1), 11466 (2008).
4. F. A. Smith *et al.*, *Ecology* **84**, 3403 (2003).
5. J. Ray, K. H. Redford, R. Steneck, J. Berger, Eds., *Large Carnivores and the Conservation of Biodiversity* (Island, Washington, DC, 2005).
6. E. T. Duffy, *Ecol. Lett.* **6**, 680 (2003).
7. D. R. Strong, K. T. Frank, *Annu. Rev. Environ. Resour.* **35**, 1 (2010).
8. O. J. Schmitz, D. Hawlena, G. C. Trussell, *Ecol. Lett.* **13**, 1199 (2010).
9. N. G. Hairston, F. E. Smith, L. B. Slobodkin, *Am. Nat.* **94**, 421 (1960).
10. S. D. Fretwell, *Oikos* **50**, 291 (1987).
11. R. T. Paine, *J. Anim. Ecol.* **49**, 667 (1980).
12. R. C. Lewontin, *Brookhaven Symp. Biol.* **22**, 13 (1969).
13. M. Scheffer, S. Carpenter, J. A. Foley, C. Folke, B. Walker, *Nature* **413**, 591 (2001).
14. O. J. Schmitz, *Oecologia* **145**, 224 (2005).
15. J. Terborgh, J. A. Estes, Eds., *Trophic Cascades: Predators, Prey, and the Changing Dynamics of Nature* (Island, Washington, DC, 2010).
16. J. A. Estes, M. T. Tinker, T. M. Williams, D. F. Doak, *Science* **282**, 473 (1998).
17. J. Brashares, L. R. Prugh, C. J. Stoner, C. W. Epps, in (15), pp. 221–240.
18. W. J. Ripple, T. P. Rooney, R. L. Beschta, in (15), pp. 141–162.
19. R. A. Myers, J. K. Baum, T. D. Shepherd, S. P. Powers, C. H. Peterson, *Science* **315**, 1846 (2007).
20. J. A. Estes, D. O. Duggins, *Ecol. Monogr.* **65**, 75 (1995).
21. L. R. Prugh *et al.*, *Bioscience* **59**, 779 (2009).
22. R. N. Owen-Smith, *Megaherbivores: The Influence of Very Large Body Size on Ecology* (Cambridge Univ. Press, Cambridge, 1992).
23. O. J. Schmitz, P. A. Hambäck, A. P. Beckerman, *Am. Nat.* **155**, 141 (2000).
24. A. R. E. Sinclair, S. Mduma, J. S. Brashares, *Nature* **245**, 288 (2003).
25. W. J. Ripple, R. L. Beschta, *For. Ecol. Manage.* **184**, 299 (2003).
26. S. R. Carpenter *et al.*, *Ecol. Monogr.* **71**, 163 (2001).
27. F. A. Smith *et al.*, *Science* **330**, 1216 (2010).
28. J. S. Levine, T. Bobbe, T. Ray, R. G. Witt, N. Singh, *Wildland Fires and the Environment: A Global Synthesis* [United Nations Environment Programme Division of Environmental Information, Assessment, and Early Warning Technical Report (UNEP/DEIAEW/TR) 99.1, Nairobi, Kenya, 1999].
29. W. Bond, in (15), pp. 275–286.
30. R. M. Holdo *et al.*, *PLoS Biol.* **7**, e1000210 (2009).
31. T. F. Flannery, *The Future Eaters: An Ecological History of the Australasian Lands and People* (Griffin, Adelaide, Australia, 1994).
32. J. L. Gill, J. W. Williams, S. T. Jackson, K. B. Lininger, G. S. Robinson, *Science* **326**, 1100 (2009).
33. R. S. Ostfeld, R. D. Holt, *Front. Ecol. Environ.* **2**, 13 (2004).
34. A. Markandya *et al.*, *Ecol. Econ.* **67**, 194 (2008).
35. K. D. Lafferty, *Ecol. Appl.* **14**, 1566 (2004).
36. T. H. Mather, T. T. That, *Environmental Management for Vector Control in Rice Fields* [Food and Agriculture Organization (FAO) Irrigation and Drainage Papers no. 41.1, Rome, 1984].
37. D. E. Schindler, S. R. Carpenter, J. J. Cole, J. F. Kitchell, M. L. Pace, *Science* **277**, 248 (1997).
38. A. J. Pershing, L. B. Christensen, N. R. Record, G. D. Sherwood, P. B. Stetson, *PLoS One* **5**, e12444 (2010).
39. S. Nicol *et al.*, *Fish Fish.* **11**, 203 (2010).
40. J. Roman, J. J. McCarthy, *PLoS One* **5**, e13255 (2010).
41. F. A. Smith, S. M. Elliott, S. K. Lyons, *Nat. Geosci.* **3**, 374 (2010).
42. R. D. Bardgett, D. A. Wardle, *Aboveground-Belowground Linkages: Biotic Interactions, Ecosystem Processes, and Global Change* (Oxford Univ. Press, Oxford, 2010).
43. S. A. Zimov *et al.*, *Am. Nat.* **146**, 765 (1995).
44. A. E. Dunham, *Oikos* **117**, 571 (2008).
45. D. A. Frank, *Oikos* **117**, 1718 (2008).
46. T. Fukami *et al.*, *Ecol. Lett.* **9**, 1299 (2006).
47. D. A. Croll, J. L. Maron, J. A. Estes, E. M. Danner, G. V. Byrd, *Science* **307**, 1959 (2005).
48. T. Katz *et al.*, *Global Biogeochem. Cycles* **23**, GB4032 (2009).
49. J. W. Moore *et al.*, *Ecology* **88**, 1278 (2007).
50. D. R. Montgomery, J. M. Buffington, N. P. Peterson, D. Schuett-Hames, T. P. Quinn, *Can. J. Fish. Aquat. Sci.* **53**, 1061 (1996).
51. R. J. Naiman, K. H. Rogers, *Bioscience* **47**, 521 (1997).
52. R. L. Beschta, W. J. Ripple, *Earth Surf. Process. Landf.* **31**, 1525 (2006).
53. C. S. Kolar, D. M. Lodge, *Trends Ecol. Evol.* **16**, 199 (2001).
54. N. O. L. Carlsson, O. Sarnelle, D. L. Strayer, *Front. Ecol. Environ.* **7**, 525 (2009).
55. D. S. Gruner, *Ecology* **85**, 3010 (2004).
56. J. A. Savidge, *Ecology* **68**, 660 (1987).
57. M. R. Bartsch, L. A. Bartsch, S. Gutreuter, *J. N. Am. Benthol. Soc.* **24**, 168 (2005).
58. C. E. deRivera, G. M. Ruiz, A. H. Hines, P. Jivoff, *Ecology* **86**, 3364 (2005).
59. R. T. Paine, *Am. Nat.* **103**, 91 (1969).
60. K. R. Crooks, M. E. Soulé, *Nature* **400**, 563 (1999).
61. W. J. Ripple, R. L. Beschta, *Biol. Conserv.* **133**, 397 (2006).
62. J. Berger, P. D. Stacey, L. Bellis, M. P. Johnson, *Ecol. Appl.* **11**, 947 (2001).
63. B. E. McLaren, R. O. Peterson, *Science* **266**, 1555 (1994).
64. R. L. Beschta, W. J. Ripple, *Biol. Conserv.* **142**, 2401 (2009).
65. C. Cabascon, D. Pothier, *For. Ecol. Manage.* **253**, 112 (2007).
66. S. A. Sandin *et al.*, *PLoS One* **3**, e1548 (2008).
67. M. E. Power, W. J. Matthews, A. J. Stewart, *Ecology* **66**, 1448 (1985).
68. J. Terborgh *et al.*, *Science* **294**, 1923 (2001).
69. A. R. E. Sinclair *et al.*, *Conserv. Biol.* **21**, 580 (2007).
70. S. E. Reisewitz, J. A. Estes, C. A. Simenstad, *Oecologia* **146**, 623 (2006).
71. D. O. Duggins, C. A. Simenstad, J. A. Estes, *Science* **245**, 170 (1989).
72. D. B. Irons, R. G. Anthony, J. A. Estes, *Ecology* **67**, 1460 (1986).
73. R. G. Anthony, J. A. Estes, M. A. Ricca, A. K. Miles, E. D. Forsman, *Ecology* **89**, 2725 (2008).

Acknowledgments: Financial and logistic support provided by the Institute for Ocean Conservation Science, Defenders of Wildlife, and the White Oak Plantation; research support by the NSF (various U.S. authors), Natural Sciences and Engineering Research Council of Canada (A.R.E.S.), and Nordforsk (L.O., T.O., R.V.). C. Eisenberg, G. Mittelbach, J. Moore, D. Reznick, O. Schmitz, and T. Williams commented on the manuscript, and E. Smith assisted in manuscript preparation.

10.1126/science.1205106

Defensive Plant-Ants Stabilize Megaherbivore-Driven Landscape Change in an African Savanna

Jacob R. Goheen^{1,2,3,5,*} and Todd M. Palmer^{3,4,5}

¹Department of Zoology and Physiology and Department of Botany, University of Wyoming, Laramie, WY 82071, USA

²Department of Zoology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

³Mpala Research Centre, P.O. Box 555, Nanyuki, Kenya

⁴Department of Biology, University of Florida, Gainesville, FL 32611, USA

Summary

Tree cover in savanna ecosystems is usually regarded as unstable, varying with rainfall, fire, and herbivory [1–4]. In sub-Saharan Africa, elephants (*Loxodonta africana*) suppress tree cover, thereby maintaining landscape heterogeneity by promoting tree-grass coexistence. In the absence of elephants, tree encroachment may convert savannas into closed-canopy woodlands [5, 6]; when elephants increase in abundance, intensified browsing pressure can transform savannas into open grasslands [5–8]. We show that symbiotic ants stabilize tree cover across landscapes in Kenya by protecting a dominant tree from elephants. In feeding trials, elephants avoided plants with ants and did not distinguish between a myrmecophyte (the whistling-thorn tree [*Acacia drepanolobium*]) from which ants had been removed and a highly palatable, nonmyrmecophytic congener. In field experiments, elephants inflicted severe damage on whistling-thorn trees from which ants had been removed. Across two properties on which elephants increased between 2003 and 2008, cover of whistling-thorn did not change significantly inside versus outside large-scale elephant exclusion fences; over the same period of time, cover of nonmyrmecophytes differed profoundly inside versus outside exclusion fences. These results highlight the powerful role that symbioses and plant defense play in driving tree growth and survival in savannas, ecosystems of global economic and ecological importance.

Results and Discussion

Within African savannas, elephants are powerful drivers of landscape-level habitat heterogeneity, capable of inflicting intensive and extensive damage to woody plants [4–8]. Trees and shrubs employ various mechanisms to reduce such catastrophic herbivory, including vigorous resprouting of broken stems, the development of heavy buttresses, growth to large sizes, and the production of a variety of chemical and spinescent defenses that may reduce palatability [9]. In African savanna ecosystems, including the Laikipia plateau in central Kenya, many trees in the widespread genus *Acacia* are defended with spines and digestibility-reducing secondary compounds (tannins). Despite such defenses, these plants often suffer intense bouts of elephant herbivory that may

severely damage or kill mature trees (Figure 1). In striking contrast, a co-occurring congener defended by ant bodyguards (*Acacia drepanolobium*) is seldom browsed by elephants and occurs in dense near-monocultures (800–1100 individuals/hectare) throughout many portions of its range [10–13]. We hypothesized that the protective ant symbionts of *A. drepanolobium* serve as an effective defense against elephant herbivory and thus play a strong role in buffering this species from fluctuations in abundance in the face of variation in elephant numbers.

We investigated the role of symbiotic ants in determining levels of tree cover following manipulated and natural changes in numbers of large mammalian herbivores in central Kenya. Our study sites in Laikipia, Kenya (see Figure S1 available online) are underlain by one of two well-defined soil types: black clayey vertisols of volcanic origin (hereafter “clayey soils”), comprising ~35% of Laikipia, and red sandy aridosols derived from quartzite (hereafter “sandy soils”), comprising the remaining 65% of Laikipia [14]. Although elephant abundances are similar between clayey and sandy soils [15], each soil type harbors a distinctive community of woody plants. On clayey soils, *A. drepanolobium* occurs in virtual monoculture, typically accounting for ≥95% of the overstory vegetation [16]. *Acacia drepanolobium* is a myrmecophyte (ant-plant), providing both housing (swollen thorn domatia) and food (extrafloral nectar) for symbiotic ants. Four species of ants (*Crematogaster mimosae*, *C. nigriceps*, *C. sjostedti*, and *Tetraponera penzigi*) compete for exclusive access to host plants and protect host trees (to varying degrees) by swarming, biting, and stinging intruders [17]. Tree communities occurring on sandy soils are more diverse, with the 3–5 most common woody plant species typically accounting for no more than 80% of the canopy in a given locale. *Acacia drepanolobium* is virtually absent from sandy soils, constituting <0.1% of the overstory.

Between 1992 and 2002, elephant abundances throughout the Laikipia ecosystem increased approximately 5-fold [18] (P. Omondi, personal communication) and continued to increase over the course of our study (unpublished data; W. Giesen, personal communication; Figure S2). To assess the impact of increased elephant densities on tree assemblages, we quantified changes in tree cover both inside and outside of plots excluding megaherbivores (elephants and less-common giraffe [*Giraffa camelopardalis*]) on sandy and clayey soils at the Lewa Wildlife Conservancy in central Kenya (37°41'E, 0°2'N, Figure S1). Changes in tree cover were determined by comparing high-resolution (60 cm) Quickbird satellite images (Digital Globe) acquired in 2003 and 2008. Between 1992 and 2002, six double-strand, electrified fences were erected on Lewa to exclude megaherbivores from parcels of land while allowing other wildlife species to freely pass beneath the 2 m high fence strands. Four fences were established in sandy soil, and two fences were established in clayey soil (Table S1). Hereafter, we refer to megaherbivore enclosures as elephant enclosures, because elephants (and not giraffes) were responsible for the vast majority of differences arising from megaherbivore browsing on both clayey and sandy soils (Tables S2 and S3). Control plots paired with

*Correspondence: jgoheen@uwyo.edu

⁵These authors contributed equally to this work



Figure 1. Elephant Herbivory on *Acacia* spp.

Recent catastrophic herbivory by elephants on the nonmyrmecophyte *Acacia mellifera* (foreground) surrounded by unbrowsed individuals of the myrmecophyte *A. drepanolobium*.

elephant exclosures were sampled from 200 m wide buffer strips around each exclosure (see “Satellite Imagery and Aerial Photos” in [Experimental Procedures](#)).

To increase our sample sizes within clayey soil habitats, we further quantified changes in tree cover at the Kenya Long-term Exclusion Experiment (KLEE) at the Mpala Research Centre in central Kenya (36°52'E, 0°17'N, [Figure S1](#)) using a single Quickbird satellite image from June 2003 (Digital Globe) and a high-resolution (30 cm) aerial photograph (Ramani Communications) from December 2007. Established in 1995, KLEE consists of three replicate blocks, each of which contains two plots accessible to all wildlife, two plots accessible to all wildlife except megaherbivores, and two plots from which all wildlife are excluded. KLEE occurs entirely on clayey soil.

Between 2003 and 2008, tree cover diverged significantly between elephant exclosures and control plots on sandy soil at Lewa ($F_{1,6} = 9.27$, $p = 0.02$). Absolute tree cover increased by 6.0% in response to elephant exclusion (from 25.3% in 2003 to 31.3% in 2008; [Figure 2](#)) while simultaneously decreasing by 8.3% within control plots (from 24.7% in 2003 to 16.6% in 2008; [Figure 2](#)). The decline in tree cover in control plots coincided with an ~2.5-fold increase in elephant densities at Lewa ([Figure S2](#)). In 2003, tree cover within elephant exclosures and control plots on clayey soils did not differ between Lewa and KLEE, nor did change in tree cover between 2003 and 2008 differ significantly between Lewa and KLEE (see “Analysis of Remotely-Sensed Imagery” in [Supplemental Experimental Procedures](#)). Thus, we pooled sites in our analysis for tree cover within elephant exclosures on clayey soils. Between 2003 and 2008 (2007 at KLEE), tree cover did not change significantly between elephant exclosures and control plots on *A. drepanolobium*-dominated clayey soil at Lewa and KLEE (exclosures: 26.5%–23.9%; controls: 22.5%–22.8%; $F_{1,8} = 0.90$, $p = 0.37$; [Figure 2](#)), nor did the change in tree cover inside versus outside elephant exclosures differ significantly from zero, despite increasing elephant numbers on both Lewa and Mpala ([Figure S2](#)).

To explore whether differential change in tree cover was due to ants or other factors associated with sandy versus clayey

soils, we conducted ground surveys for the incidence of browse on clayey soils. Ground surveys revealed that elephants preferred to browse on nonmyrmecophytes ([Tables S2 and S3](#)), thereby reducing tree cover of subordinate (i.e., nonmyrmecophytic) woody plant species (multivariate analysis of variance [MANOVA] for megaherbivore effect on subordinate species: Wilks' $\lambda_{9,4} = 0.003$, $p < 0.0001$; $p < 0.01$ for univariate F tests on five most abundant nonmyrmecophytes; [Figure 3](#)). Further, and consistent with analysis of remotely sensed data, ground surveys confirmed relatively low levels of browsing on *A. drepanolobium* by elephants ([Tables S2 and S3](#)) and nonsignificant impacts of elephants on tree cover of *A. drepanolobium* ($p = 0.27$; [Figure 3](#)). Other (nonelephant) browsers reduced cover of a single subordinate species (*Rhus natalensis*; Wilks' $\lambda_{9,4} = 0.05$, $p = 0.03$; univariate F test for *R. natalensis*: $p = 0.01$).

To directly establish whether plant defense by *Acacia* ants influenced elephant browsing of host *A. drepanolobium* trees, we conducted a 12 month in situ ant removal experiment 2.5 km east of KLEE. We reduced ant abundances on host plants by removing approximately 100%, 60%, or 30% of existing colony members on individual trees, and then we assessed levels of elephant damage on these trees relative to unmanipulated plants after a 1 year period. The level of

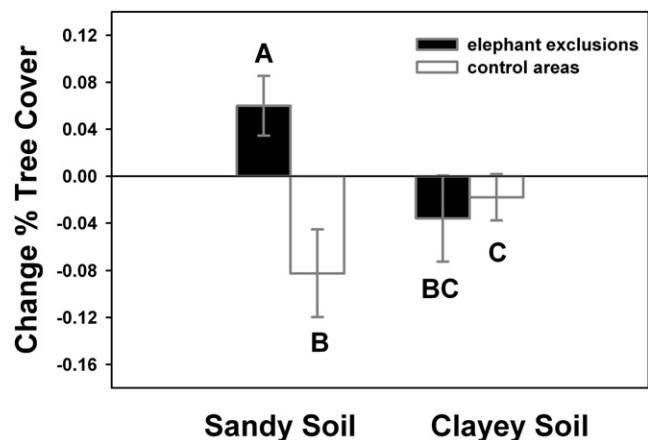


Figure 2. Landscape Change on Lewa Downs Conservancy and Mpala
Change in tree cover in elephant exclosures (black bars) and paired controls (white bars), 2003–2008. Virtually all trees on sandy soil are nonmyrmecophytes; ~95% of trees on clayey soil are the myrmecophyte *A. drepanolobium*. Means ($\pm 95\%$ confidence intervals) from sandy soil represent averages across four exclosure plots and their paired controls from Lewa. Means from clayey soil represent averages across six exclosure plots and their paired controls at KLEE and two exclosure plots and their paired controls at Lewa. Letters associated with bars represent statistically significant differences between groups ($p < 0.05$).

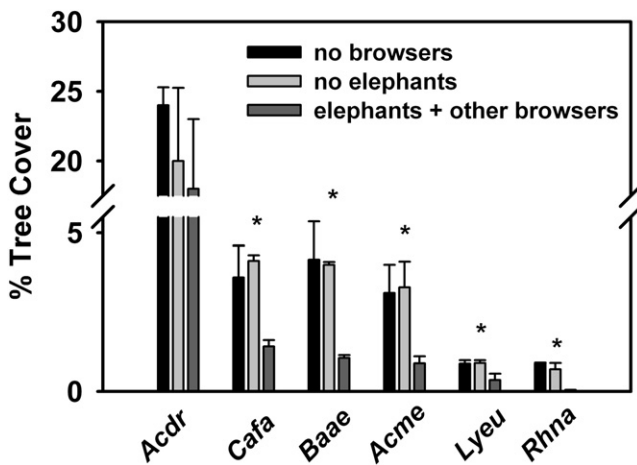


Figure 3. Differences in Tree Cover as a Function of Herbivore Treatment
Percent tree cover of the myrmecophyte *A. drepanolobium* (Acdr) and nonmyrmecophytes *Cadaba farinosa* (Cafa), *Balanites aegyptiaca* (Baae), *A. mellifera* (Acme), *Lycium europaeum* (Lyeu), and *Rhus natalensis* (Rhna) by herbivore treatment at KLEE in 2008. Black bars (\pm standard error of the mean) represent plots from which all browsers have been excluded, light gray bars represent plots from which only elephants have been excluded, and dark gray bars represent plots accessible to all browsers. * $p < 0.01$ is statistically significant between plots.

elephant browsing on host plants was significantly and negatively related to ant abundances on host plants (negative binomial regression: $\chi^2_1 = 28.41$, $p < 0.01$; Figure 4).

To further investigate whether protection by ant symbionts was the causal mechanism underlying observed patterns of landscape change, we conducted free-choice feeding trials on six 8-year-old elephants at the Sheldrick Wildlife Trust Reintegration Centre in Tsavo National Park, Kenya. We presented elephants with four groups comprised of ~ 20 1.5–2 m branches: (1) *Acacia drepanolobium* control, (2) *A. drepanolobium* ant removal, (3) *A. mellifera* control, and (4) *A. mellifera* ant addition. Browse surveys on sandy soils from Lewa demonstrated that elephants neither prefer nor avoid *A. mellifera* relative to other tree species on sandy soil (Table S4); thus, *A. mellifera* represents a typical nonmyrmecophyte from the perspective of an elephant. Groups of branches were spaced 10 m apart from each other, and their position in the elephant corral was determined randomly. Elephants were equally likely to feed on *A. drepanolobium* and *A. mellifera* in the absence of ants (Cox regression: $\beta = 0.33$, $p = 0.57$), indicating that, without its ants, *A. drepanolobium* is fundamentally palatable to elephants. Similarly, elephants avoided branches of both tree species if ants were present (Cox regression: $\beta = 3.61$, $p < 0.01$), demonstrating that symbiotic ants can deter elephant herbivory when alternative food plants are available (Figure 5).

Elephants are known to avoid swarming attack by other hymenopterans (bees [19]). The efficacy of ant defense may result from a combination of high densities of ants on host plants (up to 90,000 workers on some trees [20]), the species of ant occupant (*C. mimosae* and *C. nigriceps* swarm equally aggressively in response to disturbance [21] and occupy $\sim 70\%$ and 80% of trees at KLEE [22] and Lewa [“Lewa Browse Surveys” in Supplemental Experimental Procedures]), and the tendency of ants to attack areas of thin skin and mucous membranes by biting down and holding fast with their

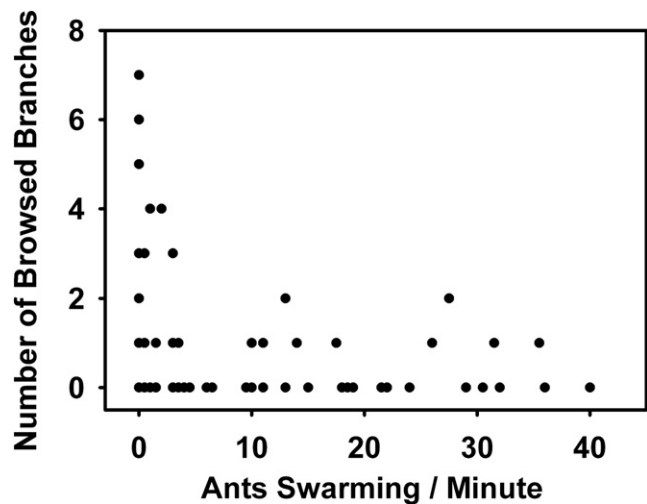


Figure 4. Responses of Elephant Browsing to In Situ Ant Removal from Trees

The number of *A. drepanolobium* branches browsed by elephants as a function of ant activity levels on trees from the ant-removal experiment ($\chi^2_1 = 28.41$, $p < 0.01$).

mandibles. Further, elephants are unique in that their nostrils are located away from their mouths at the apex of their feeding apparatus (trunk), rendering them vulnerable to swarming insects. In contrast, giraffes use their long, prehensile tongues to swipe away ants from their muzzles (unpublished data). Thus, despite their thick dermis, elephants are highly sensitive

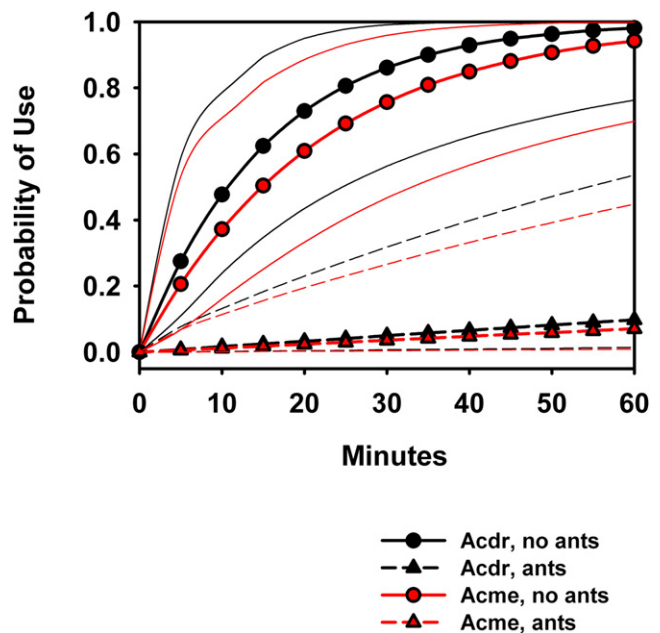


Figure 5. Free-Choice Feeding Trials with Elephants and *Acacia* spp.

Best-fitting Cox regression models as a function of food type: control *A. drepanolobium* (black triangles), ant-removal *A. drepanolobium* (black circles), control *A. mellifera* (red circles), ant-addition *A. mellifera* (red triangles). Solid and dashed thin lines represent 95% confidence intervals for probability of use of branches with and without ants, respectively. Elephants preferred branches without ants ($\beta = 3.61$, $p < 0.01$) but did not distinguish between tree species ($\beta = 0.33$, $p = 0.57$). The points in the graph are fitted from the model.

around their eyes and on the inner membranes of their trunks [23]; attack by scores of biting ants probably serves as a strong deterrent.

Classic experiments by Janzen [24] and subsequent studies by others [25–28] have elegantly demonstrated the capacity of symbiotic ants to prevent or greatly reduce herbivory and competition for light, thereby promoting the growth and survival of individual host plants. Ours is the first study to demonstrate that ant-plant symbioses can stabilize landscape structure at larger spatial scales by protecting adult trees from catastrophic herbivory. Because ants reduce palatability of their host trees, selective browsing on nonmyrmecophytes is at least partly responsible for creating the virtual monocultures of *A. drepanolobium* that typify black clayey soils in Laikipia [27, 29] and other regions of East Africa [10–13, 28]. A major challenge for the future is elucidating why *A. drepanolobium* is restricted to clayey vertisols, such that a diversity of nonmyrmecophytes thrives on other soil types. We hypothesize that characteristics of sandy soils (e.g., particle size, nutrient content, infiltration, etc.) favor nonmyrmecophytes and interact strongly with browsing to promote segregation of trees (*A. drepanolobium* and other myrmecophytes versus nonmyrmecophytes) across soil types (e.g., see [30]). In light of this, we expect that, on clayey soils, other determinants of savanna structure (i.e., rainfall and fire) or mortality factors affecting early life stages of trees (e.g., seeds, seedlings) will override browsing as drivers of tree populations, where plant defenses nullify elephants as important agents of mortality on adult trees [31].

Savannas typically are envisaged as unstable or disequilibrium systems in which climatic variability or disturbances generate the tree-grass mixtures that typify these ecosystems [1–3]. In Africa, browsing and killing of trees by elephants is often a critical force underlying the coexistence of trees and grasses [4–8]. Throughout much of their historic range, however, declines in elephant populations have triggered extensive increases in tree numbers, shifting open savannas to closed-canopy woodlands [8]. Elsewhere, and typically in response to confinement within protected areas, elephants have become “compressed,” have overexploited trees, and have shifted savannas toward structurally simplified grasslands [8, 32]. Our study highlights the stabilizing effect that ant symbionts can confer on tree cover over expansive spatial scales. In sum, our experiments show that ant symbionts protect against catastrophic herbivory, effectively buffering a dominant tree against top-down control by megaherbivores. Because tree cover strongly regulates a host of ecosystem processes, including carbon storage, fire-return intervals, food web dynamics, nutrient cycling, and soil-water relations in our system [33, 34] and others [35, 36], these tiny bodyguards likely exert powerful indirect effects at very large spatial and temporal scales. As elephants and other large mammals in Africa exhibit chronic declines in some habitats and overabundance in others, identifying the ecological consequences of such landscape change remains an important challenge for wildlife managers in the future.

Experimental Procedures

Satellite Imagery and Aerial Photos

At Lewa, control plots paired with elephant enclosures were sampled from 200 m wide buffer strips around each enclosure, subject to the constraint that the buffer strip occurred entirely within Lewa boundaries. When an elephant enclosure abutted a neighboring property, we expanded the width of buffer strips to compensate for the area not sampled in that property.

At KLEE, elephant enclosures consisted of the central hectare (ha) within each 4 ha fence.

KLEE Browse Surveys

From July 2007 to September 2007 at KLEE, we recorded canopy breadth, height, and diameter at breast height (DBH) on all individuals of the subordinate woody species (i.e., nonmyrmecophytes; $n = 721$). We paired each of the 721 individual trees with the nearest neighboring *A. drepanolobium*, subject to the constraint that the diameter of the *A. drepanolobium* was within 5 cm of the subordinate individual with which it was paired, and we recorded canopy breadth, height, DBH, and incidence of browsing. We used MANOVA to test for the effects of megaherbivores, wildlife, and cattle on percent tree cover of *A. drepanolobium* and the five most common, subordinate woody species: *A. mellifera*, *Balanites aegyptiaca*, *Cadaba farinosa*, *Lycium europaeum*, and *Rhus natalensis*. In addition, we included replicate as a fixed effect (random effects are extremely difficult to implement and interpret in MANOVA) in our analysis, because tree cover at KLEE increases from north to south. For each individual in the six plots accessible to megaherbivores ($n = 332$), we recorded the incidence of browsing by megaherbivores (elephant and giraffe). We ignored elephant browsing >1 year old, as evidenced by chalky, dull-colored wood. We used log-linear models to calculate odds ratios of browse by elephant and giraffe on each of the five most common nonmyrmecophytes (Table S3).

Ant-Removal Experiment

We removed ant colonies from host plants by inundating the host plant with smoke generated by burning dry grass in a bucket underneath the tree. *Crematogaster mimosae* displays an evacuation behavior when inundated by the smoke from burning grass, in which workers carry the majority of brood, eggs, pupae, winged reproductives, and queens from swollen thorn domatia into cracks in the soil at the base of the host plant over the course of 45–60 min. During smoke inundation, we made a rough approximation of the total number of workers on each tree. We then imposed treatments immediately following smoke inundation, either completely barring ants from recolonizing trees by applying a Tanglefoot sticky barrier at the base of the host plant (full ant removal) or allowing approximately 1/3, 2/3, or the entire colony to recolonize the host plant prior to applying a Tanglefoot barrier (for the 2/3 removal, 1/3 removal, and control treatments, respectively). On control trees, two small (6 cm) dead branches were wired to trees across the sticky barrier to allow ants to move freely across the sticky barrier.

Following the imposition of treatments, we then assayed trees for relative levels of ant defense at 6 and 12 months by disturbing a randomly chosen swollen thorn on each of two branches per tree (one in each of the north and south cardinal directions) and counting the number of workers swarming onto the tip of the uppermost spine of the disturbed swollen thorn over a 30 s period. Two researchers conducted these assays so that swollen thorns on two separate branches could be disturbed simultaneously. Overall levels of ant activity on trees were calculated as the average of these activity assays. Ant activity at 12 months was significantly correlated with ant activity at 6 months ($r = 0.64$, $p < 0.0001$), and our treatments were effective in generating a range of ant activities (Table S5). We then resurveyed each tree at the end of 12 months for browse damage inflicted by both elephants and nonelephant browsers.

Free-Choice Feeding Trials

We cut 1.5–2.0 m branches of *A. drepanolobium* and *A. mellifera* from trees at the periphery of Tsavo National Park near the town of Voi. At the time of collection, all *A. drepanolobium* in our feeding trials were inhabited by the ant *C. nigriceps*. We removed ants and swollen thorns from *A. drepanolobium* with wire cutters. Branches were transported to the holding corral and were presented to elephants within 2 hr of collection. Ants were added to *A. mellifera* by immersing branches for 5 min in a metal drum into which the ants and swollen thorns from the ant removal branches had been collected. A single observer, positioned ~20 m from groups of branches, recorded data on the number, type, and order of foods taken over a 1 hr period. Groups of branches were randomly placed 10 m apart within 10 m of the entrance of the holding corral.

Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures, two figures, and five tables and can be found with this article online at doi:10.1016/j.cub.2010.08.015.

Acknowledgments

A. Hassan, F. Kobia, J. Lemboi, S. Lima, and A. Lopez provided assistance in the field. G. Aike, G. Chege, N. Georgiadis, M. Kinnaird, J. Kinoti, J. Sauni, D. Sheldrick, and C. Thouless provided logistical support. W. Giesen, M. Graham, and T. O'Brien shared unpublished data. T.M. Anderson, D.J. Augustine, R.M. Holdo, R.M. Pringle, A.R.E. Sinclair, and T.P. Young provided comments. We thank T.P. Young for continued collaboration through KLEE. This work was funded by the American Society of Mammalogists, the Natural Science and Engineering Research Council of Canada, the Smithsonian Tropical Research Institute, and the U.S. National Science Foundation. KLEE was built and maintained through support from the National Geographic Society, National Science Foundation, U.S. Fish and Wildlife Service, and the Smithsonian Institution. This research was carried out in accordance with permit MOEST 13/001/3417 from the Republic of Kenya.

Received: May 24, 2010

Revised: July 14, 2010

Accepted: August 9, 2010

Published online: September 2, 2010

References

- Scholes, R.J., and Archer, S.R. (1997). Tree-grass interactions in savannas. *Annu. Rev. Ecol. Syst.* 28, 517–544.
- Sankaran, M., Ratnam, J., and Hanan, N.P. (2004). Tree-grass coexistence in savannas revisited: Insights from an examination of assumptions and mechanisms invoked in existing models. *Ecol. Lett.* 7, 480–490.
- Sankaran, M., Hanan, N.P., Scholes, R.J., Ratnam, J., Augustine, D.J., Cade, B.S., Gignoux, J., Higgins, S.I., Le Roux, X., Ludwig, F., et al. (2005). Determinants of woody cover in African savannas. *Nature* 438, 846–849.
- Holdo, R.M. (2007). Elephants, fire, and frost can determine community structure and composition in Kalahari Woodlands. *Ecol. Appl.* 17, 558–568.
- Dublin, H.T., Sinclair, A.R.E., and McGlade, J. (1990). Elephants and fire as causes of multiple stable states in the Serengeti-Mara woodlands. *J. Anim. Ecol.* 59, 1147–1164.
- Laws, R.M. (1970). Elephants as agents of habitat and landscape change in East Africa. *Oikos* 21, 1–15.
- Cumming, D.H.M., Fenton, M.B., Rautenbach, I.L., Taylor, R.D., Cumming, G.S., Cumming, M.S., Dunlop, J.M., Ford, G.S., Hovorka, M.D., Johnston, D.S., et al. (1997). Elephants, woodlands and biodiversity in southern Africa. *S. Afr. J. Sci.* 93, 231–236.
- Western, D., and Maitumo, D. (2004). Woodland loss and restoration in a savanna park: A 20-year experiment. *Afr. J. Ecol.* 42, 111–121.
- Sheil, D., and Salim, A. (2004). Forest tree persistence, elephants, and stem scars. *Biotropica* 36, 505–521.
- Herlocker, D. (1975). *Woody Vegetation of the Serengeti National Park* (College Station, TX: Texas A&M University).
- Young, T.P., and Isbell, L.A. (1991). Sex differences in giraffe feeding ecology: Energetic and social constraints. *Ethology* 87, 79–89.
- Dalle, G.T. (2004). *Vegetation ecology, rangeland condition, and forage resources evaluation in the Borana lowlands, southern Oromia, Ethiopia*. PhD thesis, Universitat Gottingen, Gottingen, Germany.
- Andrews, P., and Bamford, M. (2008). Past and present vegetation ecology of Laetoli, Tanzania. *J. Hum. Evol.* 54, 78–98.
- Ahn, P.M., and Geiger, L.C. (1987). *Soils of Laikipia District* (Nairobi, Kenya: National Agricultural Research Laboratories).
- Khaemba, W.M., Stein, A., Rasch, D., De Leeuw, J., and Georgiadis, N.J. (2001). Empirically simulated study to compare and validate sampling methods used in aerial surveys of wildlife populations. *Afr. J. Ecol.* 39, 374–382.
- Young, T.P., Stubblefield, C.H., and Isbell, L.A. (1997). Coexistence among obligate acacia ants. *Oecologia* 109, 98–107.
- Palmer, T.M., Young, T.P., Stanton, M.L., and Wenk, E. (2000). Short-term dynamics of an acacia ant community in Laikipia, Kenya. *Oecologia* 123, 425–435.
- Thouless, C.R., and Sakwa, J. (1995). Shocking elephants: Fences and crop raiders in Laikipia District, Kenya. *Biol. Conserv.* 72, 99–107.
- Vollrath, F., and Douglas-Hamilton, I. (2002). African bees to control African elephants. *J. Am. Vet. Med. A. Naturwissenschaften* 89, 508–511.
- Hocking, B. (1970). Insect associations with the swollen thorn acacias. *Trans. R. Ent. Soc. London* 122, 211–255.
- Palmer, T.M., and Brody, A.K. (2007). Mutualism as reciprocal exploitation: African plant-ants defend foliage but not reproductive structures of an African ant-plant. *Ecology* 88, 3004–3011.
- Palmer, T.M., Stanton, M.L., Young, T.P., Goheen, J.R., Pringle, R.M., and Karban, R. (2008). Breakdown of an ant-plant mutualism follows the loss of large herbivores from an African savanna. *Science* 319, 192–195.
- Jacobson, E.R., Sundberg, J.P., Gaskin, J.M., Kollias, G.V., and O'Banion, M.K. (1986). Cutaneous papillomas associated with a herpesvirus-like infection in a herd of captive African elephants. *J. Am. Vet. Med. Assoc.* 189, 1075–1078.
- Janzen, D.H. (1966). Coevolution of mutualism between ants and acacias in Central America. *Evolution* 20, 249–275.
- Bronstein, J.L. (1998). The contribution of ant-plant protection studies to our understanding of mutualism. *Biotropica* 30, 150–161.
- Heil, M., and McKey, D. (2003). Protective ant-plant interactions as model systems in ecological and evolutionary research. *Annu. Rev. Ecol. Syst.* 34, 425–454.
- Madden, D., and Young, T.P. (1992). Symbiotic ants as an alternative defense against giraffe herbivory in spinescent *Acacia drepanolobium*. *Oecologia* 91, 235–238.
- Stapley, L. (1998). The interaction of thorns and symbiotic ants as an effective defence mechanism of swollen-thorn acacias. *Oecologia* 115, 401–405.
- Young, T.P., Okello, B., Kinyua, D., and Palmer, T.M. (1998). KLEE: A long-term multi-species herbivore exclusion experiment in Laikipia, Kenya. *Afr. J. Range Forage Sci.* 14, 94–102.
- Fine, P.V.A., Mesones, I., and Coley, P.D. (2004). Herbivores promote habitat specialization by trees in Amazonian forests. *Science* 305, 663–665.
- Goheen, J.R., Palmer, T.M., Keesing, F., Riginos, C., and Young, T.P. (2010). Large herbivores facilitate savanna tree establishment via diverse and indirect pathways. *J. Anim. Ecol.* 79, 372–382.
- Whyte, I., van Aarde, R., and Pimm, S. (1998). Managing the elephants of Kruger National Park. *Anim. Conserv.* 1, 77–83.
- Pringle, R.M., and Fox-Dobbs, K. (2008). Coupling of canopy and understory food webs by ground-dwelling predators. *Ecol. Lett.* 11, 1328–1337.
- Riginos, C., Grace, J.B., Augustine, D.J., and Young, T.P. (2009). Local versus landscape-scale effects of savanna trees on grasses. *J. Ecol.* 97, 1337–1345.
- Belsky, A.J., Amundson, R.G., Duxbury, J.M., Riha, S.J., Ali, A.R., and Mwonga, S.M. (1989). The effects of trees on their physical, chemical, and biological environments in a semi-arid savanna in Kenya. *J. Appl. Ecol.* 26, 1005–1024.
- Holdo, R.M., Sinclair, A.R.E., Dobson, A.P., Metzger, K.L., Bolker, B.M., Ritchie, M.E., and Holt, R.D. (2009). A disease-mediated trophic cascade in the Serengeti and its implications for ecosystem C. *PLoS Biol.* 7, e1000210.

- Anderson, D. M., Hallegraeff, G. M. & Cembella, A. D.) 267–292 (NATO Advanced Study Institute Series, Springer, Heidelberg, 1998).
16. Bates, S. S. in *The Physiological Ecology of Harmful Algal Blooms* (eds Anderson, D. M., Hallegraeff, G. M. & Cembella, A. D.) 405–426 (NATO Advanced Study Institute Series, Springer, Heidelberg, 1998).
 17. Gulland, F. et al. *Unusual Marine Mammal Mortality Event—Domoic Acid Toxicity in California Sea Lions (*Zalophus californianus*) Stranded Along the Central California Coast, May–October 1998* (NOAA Technical Memo NMFS-OPR-8, National Marine Fisheries Service, US Department of Commerce, Silver Spring, Maryland, 1999).
 18. Sutherland, R. J., Hoising, J. M. & Whisaw, I. Q. Domoic acid, an environmental toxin, produces hippocampal damage and severe memory impairment. *Neurosci. Lett.* **120**, 221–223 (1990).
 19. Strain, S. M. & Tasker, R. A. Hippocampal damage produced by systemic injections of domoic acid in mice. *Neuroscience* **44**, 343–352 (1991).
 20. Dakshinamurti, K., Sharma, S. K., Sundaram, M. & Wanatabe, T. Hippocampal changes in developing postnatal mice following intrauterine exposure to domoic acid. *J. Neurosci.* **13**, 4486–4495 (1993).
 21. Antonelis, G. A., Fiscus, C. H., DeLong, R. H. Spring and summer prey of California sea lions, *Zalophus californianus*, at San Miguel Island, California, 1978–1979. *Fish. Bull.* **82**, 67–76 (1984).
 22. Ochoa, J. I., Sierra-Beltrán, A. P., Oláiz-Fernández, G. & Del Villar Ponce, L. M. Should mollusc toxicity in Mexico be considered a public health issue? *J. Shellfish Res.* **17**, 1671–1673 (1998).
 23. Sakamoto, C., Freiderich, G. E. & Codispoti, L. A. *MBARI Procedures for Automated Nutrient Analyses Using a Modified Alpkem Series 300 Rapid Flow Analyzer* (MBARI Technical Report 90-2, Monterey Bay Aquarium Research Institute, Moss Landing, California, 1990).
 24. Quilliam, M. A., Xie, M. & Hardstaff, W. R. A rapid extraction and clean-up procedure for the liquid chromatographic determination of domoic acid in unsalted seafood. *J. AOAC Int.* **78**, 543–554 (1995).
 25. Hatfield, C. L., Wekell, J. C., Gauglitz, E. J. Jr & Barnett, H. J. Salt clean-up procedure for the determination of domoic acid by HPLC. *Nat. Toxins* **2**, 206–211 (1994).
 26. McClain, C. R. et al. Science quality seaWiFS data for global biosphere research. *Sea Technol.* **39**, No. 9 (1998).
 27. Luna, L. G. *Manual of Histologic Staining Methods of the Armed Forces Institute of Pathology*, 3rd edn (McGraw-Hill, New York, 1968).

Acknowledgements

We thank the many volunteers who assisted The Marine Mammal Center (TMMC), members of the Monterey Bay Sanctuary Beach COMBERS programme, the crew of the RV *Pt. Lobos*, R. Kudela for providing the processed SeaWiFS image, and P. Walz for nutrient analyses. This work was supported in part by NOAA/NOS operational funds, NOAA Saltonstall-Kennedy grant programme, the many supporters of TMMC, and The David and Lucile Packard Foundation. All authors contributed equally to the work.

Correspondence and requests for materials should be addressed to C.S. (e-mail: scholin@mbari.org).

Dynamic biogeography and conservation of endangered species

Rob Channell* & Mark V. Lomolino†

* Department of Biological Sciences, Fort Hays State University, Hays, Kansas 67601, USA

† Oklahoma Biological Survey, Oklahoma Natural Heritage Inventory, and Department of Zoology, University of Oklahoma, Norman, Oklahoma 73019, USA

As one moves from the core to the periphery of a species' geographical range, populations occupy less favourable habitats and exhibit lower and more variable densities^{1–4}. Populations along the periphery of the range tend to be more fragmented and, as a result, are less likely to receive immigrants from other populations. A population's probability of extinction is directly correlated with its variability and inversely correlated with density and immigration rate^{5–9}. This has led to the prediction that, when a species becomes endangered, its geographical range should contract inwards, with the core populations persisting until the final stages of decline^{2,10}. Convinced by these logical but untested deductions, conservation biologists and wildlife managers have been instructed to avoid the range periphery when planning conservation strategies or allocating resources for endangered species^{11–13}. We have analysed range contraction in 245 species from a broad range of taxonomic groups and geographical regions. Here we report that observed patterns of

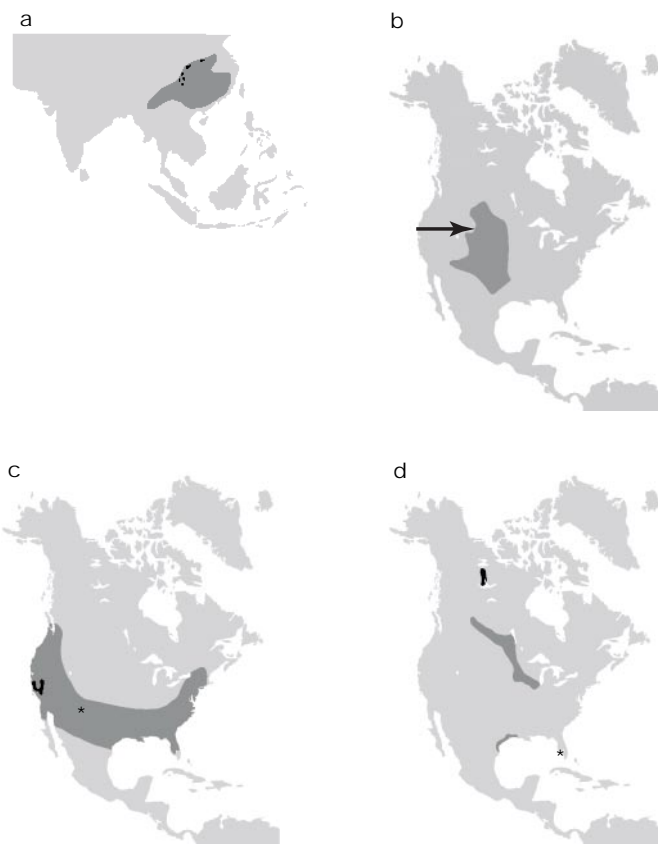


Figure 1 Patterns of range contraction in four endangered species. **a**, Giant panda, *Ailuropoda melanoleuca*; **b**, black-footed ferret, *Mustela nigripes*; **c**, California condor, *Gymnogyps californianus*; **d**, whooping crane, *Grus americana*. Historical range is in grey, extant range is in black or indicated by an arrow, and asterisks mark the locations of recent re-introduction sites for the California condor and the whooping crane.

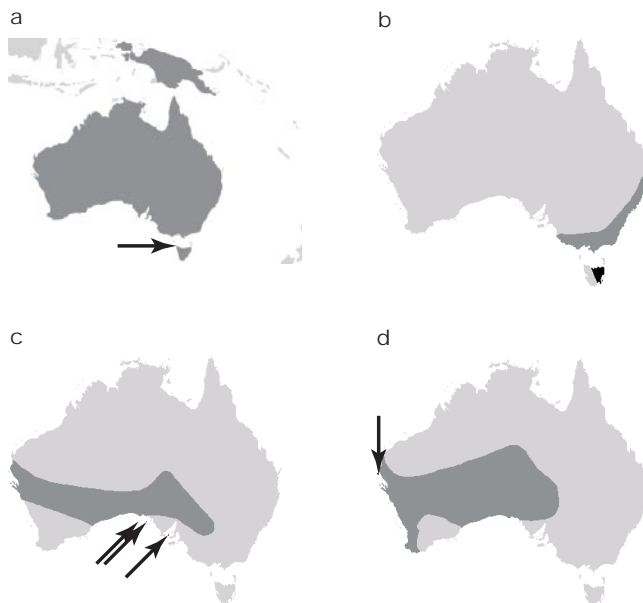


Figure 2 Patterns of range contraction in four species whose historical range included islands as well as much larger areas on the Australian mainland. **a**, Tasmanian tiger, *Thylacinus cynocephalus*; **b**, Tasmanian bettong, *Bettongia gaimardi*; **c**, greater stick-nest rat, *Leporillus conditor*; **d**, Shark Bay mouse, *Pseudomys fieldi*. Historical range in grey, and extant or final range is in black or indicated by an arrow.

Table 1 Number of species studied from different taxonomic groups and geographical regions

	North America	Australia	Eurasia	South America	Africa	Islands	Subtotal
Birds	12	6	19	2	3	45	87
Mammals	8	36	30	5	20	1	100
Reptiles	1		2	1		1	5
Amphibians	3		1				4
Fishes	1		1				2
Mollusks	1		1			20	22
Arthropods	2		1		1		4
Plants	4					17	21
Subtotal	32	42	55	8	24	84	245

See ref. 24.

range contraction do not support the above predictions and that most species examined persist in the periphery of their historical geographical ranges.

Table 1 shows the number of species studied and their geographical distribution. We found that 240 (98%) of the 245 species maintained populations in at least a portion of their peripheral range. Furthermore, 167 (68%) maintained a greater than expected portion of their range in the periphery, not the core ($P < 0.001$, binomial test). In fact, remnant populations of 91 species occurred exclusively in the periphery of their historical range, whereas populations of only five species persisted solely in the core of their historical range ($P < 0.001$, binomial test). We detected no significant difference in the patterns of range contraction between birds and mammals (63 (72%) of 87 birds and 70 (70%) of 100 mammals exhibiting greater persistence along the periphery). Most species, including some of the flagship species of conservation biology (Fig. 1), persist along the edge of their range.

Consistent with contemporary theory in ecology^{6,7,9}, persistence was greater for populations occupying larger patches of their historical range. On the mainland, 12 (75%) of 16 species persisted in larger patches of their historical range, whereas 15 (83%) of 18 insular species persisted in larger patches. However, if a species' historical range included both mainland and insular sites, population persistence was highest on the islands, despite their smaller size (23 [68%] of 34 species exhibited greater than expected persistence on islands; $P = 0.029$, binomial test; Fig. 2).

We found two additional patterns that seem contrary to the general tendency for greater persistence along the range periphery—Africa and the Hawaiian Islands. Africa was the only continent with an adequate sample size whose species failed to exhibit a significant peripheral bias in persistence (14 (58%) of 24 species persisted in the periphery; $P = 0.271$, binomial test). In contrast, 42 (78%) of 54 Eurasian species, 34 (81%) of 42 Australian species and 26 (81%) of 32 North American species persisted in their range peripheries ($P < 0.001, 0.001, 0.001$, respectively, binomial tests). In a similar fashion, whereas 11 (92%) of the 12 species we studied from New Zealand, and all of the 6 species from the Mariana Islands (including Guam) persisted more in the periphery than expected by chance, only 43% of the 54 Hawaiian species exhibited a peripheral bias.

These apparently exceptional results and the more general tendency for persistence along the periphery indicate that range contraction is strongly influenced by anthropogenic extinction forces (for example, habitat degradation, biocides and introduced species) which render historical density patterns irrelevant. Populations that persist the longest are those last affected by the contagion-like spread of extinction forces; that is, those along the edge of the range, on an isolated and undisturbed island, or at high elevations. African species failed to show any peripheral bias in range decline because, instead of moving across species' geographical ranges like a contagion, humans having a significant ecological effect became established in many places across the continent before the earliest record of historical extinctions. We actually predicted this result for Africa, based largely on Martin's^{14,15} explanation for the absence of a

post-Pleistocene collapse of the African megafauna: large mammals and birds shared a long evolutionary and ecological history with prehistoric humans. The 'exceptional' pattern for Hawaiian species is also entirely consistent with the above hypothesis concerning the contagion-like spread of extinction forces. Polynesians and, later, Europeans colonized most of the beach front and lowlands of these islands, and then spread, along with their commensals, upward. Persistent populations of Hawaiian species are either those that can cope with these anthropogenic disturbances, or those whose final populations remain in the least disturbed and most isolated sites; that is, in the montane areas. In short, the geography of recent extinctions is largely the geography of humanity. Thus, our ability to understand patterns in recent extinctions and to predict those of future ones depends to a very large degree on our ability to reconstruct and predict the spatial dynamics of humans and associated extinction forces.

These results have strong implications for conservation biology. Although they may have represented suboptimal habitats in historical times, areas along the range periphery and on remote islands and mountain ranges often provide valuable opportunities for conserving endangered species^{16,20}. We find it very encouraging, therefore, that a number of recent conservation programmes have broadened their options by including peripheral sites for re-introductions and areas to search for undiscovered populations of endangered species (asterisks in Fig. 1c, d). Although once viewed as the land of the living dead^{21,22}, sites along the range periphery may now hold great promise for conserving endangered species and biological diversity in general. □

Methods

We obtained range maps for 245 species from the literature or through personal correspondence with authorities (see Supplementary Information). We include only those species with maps available for both historical and extant ranges (or final site in the case of extinct species), and with extant ranges that were less than 25% of the species' historical distribution. We digitized the range maps into Idrisi, a geographical information system²³. For each species, we first located the centre, which was the point within the species' historical range that was most distant from all edges of the range. The distance from this point to the nearest edge was then calculated. We defined the region that was within half of this distance to an edge as periphery and the remaining portion of the range as central. We then calculated an index of centrality (C), which is a measure of the proportion of the extant or final range that fell within the central region of the historical range.

First, we calculated the area of the extant range expected to occur within the central region (C_{EE}) as follows:

$$C_{EE} = \left(\frac{C_H}{T_H} \right) T_E,$$

where T_E is the total area of the extant (or final) range; T_H is the total area of the historical range; and C_H is the area of the central region of the historical range. We then calculated C as follows. If $C_{EO} \leq C_{EE}$, where C_{EO} is the area of the extant range observed within the historical central region, then

$$C = \left(\frac{C_{EO}}{C_{EE}} \right) 0.5$$

If $C_{EO} > C_{EE}$, then

$$C = 0.5 + \left[0.5 \left(\frac{C_{EO} - C_{EE}}{T_E} \right) \right].$$

The index of centrality (C) ranged from 0, where the extant range fell completely outside the central portion of the historical range, to 1, where the extant range fell completely within the central portion of the historical range. We designated species with C values greater than 0.5 as 'central species', and those species with C values less than 0.5 as 'peripheral species'. We then used a binomial test to determine whether the ratio of central to peripheral species differed significantly from 1 : 1.

We used maps for species with multiple patches in their historical range to test whether persistence was higher for populations inhabiting larger patches. We first assigned patches to one of two size categories ('large' or 'small'), based on their area relative to the median patch size. If a species had an odd number of patches in its historical range, the median-sized patch was excluded from the analysis. For each species, we counted the number of large and small patches maintaining persistent populations (P_1 and P_s , respectively). We counted the number of species (S_1) for which P_1 was greater than P_s and the number of species (S_s) where P_s was greater than P_1 . Species with ties ($P_1 = P_s$) were excluded from analysis. We used a binomial test to determine whether the ratio of S_1 to S_s differed significantly from 1 : 1. This analysis was done for 124 continental and 44 insular species²⁴.

To compare the relative persistence of mainland and island patches, we first calculated the total area of all of the historical patches (A_{TH}) and the area of the historical mainland patches (A_{MH}) for 44 species. We multiplied A_{MH}/A_{TH} by the total number of persisting patches (P_{TP}) to generate the expected number of patches persisting on the mainland. If the number of patches persisting on the mainland (P_{MP}) was greater than expected, we classified the species as a mainland species, otherwise it was classified as an island species. There were no ties ($P_{MP} =$ expected number of patches). We tested whether the ratio of mainland species and island species differed significantly from 1 : 1 using a binomial test.

Received 23 April; accepted 21 October 1999.

1. Brown, J. H. On the relationship between abundance and distribution of species. *Am. Nat.* **124**, 255–279 (1984).
2. Lawton, J. H. In *Extinction Rates* (eds Lawton, J. H. & May, R. M.) 147–163 (Oxford Univ. Press, 1995).
3. Brown, J. H., Mehlman, D. W. & Stevens, G. C. Spatial variation in abundance. *Ecology* **76**, 2028–2043 (1995).
4. Gaston, K. J. Patterns in the geographical ranges of species. *Biol. Rev.* **65**, 105–129 (1990).
5. MacArthur, R. H. & Wilson, E. O. The theory of island biogeography. *Monogr. Popul. Biol.* **1**, 1–203 (1967).
6. Pimm, S. L., Jones, H. L. & Diamond, J. On the risk of extinction. *Am. Nat.* **132**, 757–785 (1988).
7. Tracy, C. R. & George, T. L. On the determinants of extinction. *Am. Nat.* **139**, 102–122 (1992).
8. Brown, J. H. & Kodric-Brown, A. Turnover rates in insular biogeography: effects of immigration on extinction. *Ecology* **58**, 445–449 (1977).
9. Goel, N. S. & Richter-Dyn, N. *Stochastic Models in Biology* (Academic, New York, 1974).
10. Brown, J. H. *Macroecology* (Univ. Chicago Press, 1995).
11. Wolf, C. M., Griffith, B., Reed, C. & Temple, S. A. Avian and mammalian translocations: update and reanalysis of 1987 survey data. *Conserv. Biol.* **10**, 1142–1153 (1996).
12. Griffith, B., Scott, J. M., Carpenter, J. W. & Reed, C. Translocation as a species conservation tool: status and strategy. *Science* **245**, 477–480 (1989).
13. Pearl, M. in *Conservation Biology: The Theory and Practice of Nature Conservation Preservation and Management* (eds Fielder, P. L. & Jain, S. K.) 297–320 (Chapman & Hall, New York, 1992).
14. Martin, P. S. in *Quaternary Extinctions* (eds Martin, P. S. & Klein, R. G.) 354–403 (Univ. Arizona Press, Tucson 1984).
15. Diamond, J. M. in *Quaternary Extinctions* (eds Martin, P. S. & Klein, R. G.) 824–862 (Univ. Arizona Press, Tucson, 1984).
16. Burbidge, A. A. & McKenzie, N. L. Patterns in the modern decline of western Australia's vertebrate fauna: causes and conservation implications. *Biol. Conserv.* **50**, 143–198 (1989).
17. Short, J., Bradshaw, S. D., Giles, J., Prince, R. I. T. & Wilson, G. R. Reintroduction of macropods (Marsupialia: Macropodoidea) in Australia—A review. *Biol. Conserv.* **62**, 189–204 (1992).
18. Bibby, C. J. Recent past and future extinctions in birds. *Phil. Trans. R. Soc. Lond. B* **344**, 35–40 (1994).
19. Franklin, J. & Steadman, D. W. The potential for conservation of Polynesian birds through habitat mapping and species translocation. *Conserv. Biol.* **5**, 506–521 (1991).
20. Towns, D. R. & Daugherty, C. H. Patterns of range contractions and extinctions in the New Zealand herpetofauna following human colonisation. *N.Z. J. Zool.* **21**, 325–339 (1994).
21. Stevens, G. in *Systematics, Ecology, and the Biodiversity Crisis* (ed. Eldredge, N.) 40–58 (Columbia Univ. Press, New York, 1992).
22. Wolf, C. M., Griffith, B., Reed, C. & Temple, S. A. Avian and mammalian translocations: update and reanalysis of 1987 survey data. *Conserv. Biol.* **10**, 1142–1153 (1996).
23. Eastman, J. R. *Idrisi for Windows, Version 1.0* (Clark Labs for Cartographic Technology and Geographic Analysis, Worcester, MA, 1995).
24. Channell, R. A geography of extinction: patterns in the contraction of geographic ranges. Thesis, Univ. Oklahoma (1998).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial offices of Nature.

Acknowledgements

We thank A. Baynes, J. H. Brown, N. Czaplowski, B. Danielson, T. Franklin, M. Kaspari, B. Maurer, K. Pandora, D. Perault, K. Perez, S. Pimm, G. A. Smith and C. Vaughn for advice and comments on this paper, and J. M. Scott, D. Steadman and L. Carbyn for information on the distribution of several species. R.C. was supported by the Department of Zoology, University of Oklahoma, while conducting this research, and M.V.L. was supported by grants from the US National Science Foundation.

Correspondence and requests for materials should be addressed to R.C. (e-mail: Rchannel@FHSU.edu).

Reduced vas deferens contraction and male infertility in mice lacking P2X₁ receptors

K. Mulryan*, D. P. Gitterman*, C. J. Lewis*, C. Vial*, B. J. Leckie†, A. L. Cobb‡, J. E. Brown‡, E. C. Conley§, G. Buell||#, C. A. Pritchard¶ & R. J. Evans*

* Department of Cell Physiology & Pharmacology, Medical Sciences Building, † Department of Medicine, ‡ Transgenic Unit, Biomedical Services, § Department of Pathology and Centre for Mechanisms of Human Toxicity & ¶ Department of Biochemistry, University of Leicester, Leicester LE1 9HN, UK || Glaxo-Wellcome Biomedical Research Institute, 14 chemin des Aulx, Plan-les-Ouates, 1228 Geneva, Switzerland # Present address: Serono Pharmaceuticals, 14 chemin des Aulx, Plan-les-Ouates, 1228 Geneva, Switzerland

P2X₁ receptors for ATP are ligand-gated cation channels, present on many excitable cells including vas deferens smooth muscle cells^{1–5}. A substantial component of the contractile response of the vas deferens to sympathetic nerve stimulation, which propels sperm into the ejaculate, is mediated through P2X receptors¹. Here we show that male fertility is reduced by ~90% in mice with a targeted deletion of the P2X₁ receptor gene. Male mice copulate normally—reduced fertility results from a reduction of sperm in the ejaculate and not from sperm dysfunction. Female mice and heterozygote mice are unaffected. In P2X₁-receptor-deficient mice, contraction of the vas deferens to sympathetic nerve stimulation is reduced by up to 60% and responds to P2X

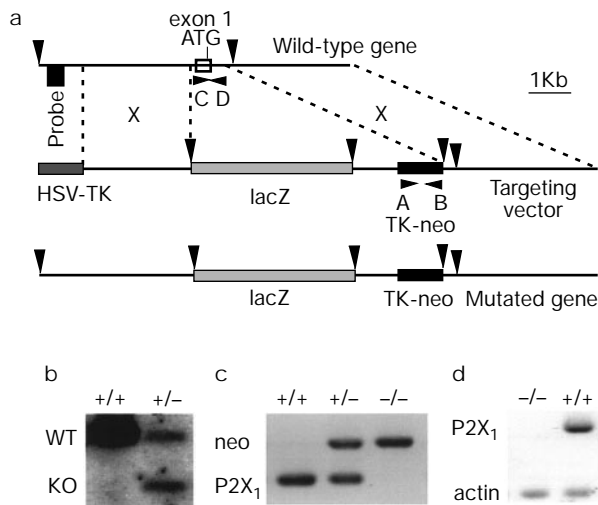


Figure 1 Generation of P2X₁-receptor-deficient mice. **a**, Genomic maps of the wild-type gene, targeting vector and mutated gene. BamHI sites (indicated by arrows) and the probe used for detection of the homologous recombination events by Southern analysis are shown. Polymerase chain reaction (PCR) primers used for genotyping of mouse-tail DNA are indicated (A–D). **b**, Southern blot analysis of tail genomic DNA from +/+ and -/- animals. Genomic DNA was digested with BamHI and hybridized with the probe indicated in **a** which detects a 4.8-kb band in +/+ DNA and a 3.7-kb band in -/- DNA. WT, wild-type; KO, knock-out. **c**, PCR genotyping of mouse-tail DNA. Primers A, B, C and D were used in one PCR reaction to genotype mouse-tail genomic DNA. Primers A and B amplify a 519-bp product from the *ned³* gene, whereas primers C and D amplify a 317-bp product from the deleted region of the P2X₁ receptor gene. **c**, RT-PCR analysis. A PCR product of 442 bp from the P2X₁-receptor gene was amplified from bladder complementary DNA from a +/+ animal but not from bladder cDNA of a -/- animal. As a control, amplification of 199-bp product from the actin gene was detected in both samples.



The Evolution of Maximum Body Size of Terrestrial Mammals
Felisa A. Smith *et al.*
Science **330**, 1216 (2010);
DOI: 10.1126/science.1194830

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 19, 2013):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/330/6008/1216.full.html>

Supporting Online Material can be found at:

<http://www.sciencemag.org/content/suppl/2010/11/22/330.6008.1216.DC1.html>

This article **cites 21 articles**, 8 of which can be accessed free:

<http://www.sciencemag.org/content/330/6008/1216.full.html#ref-list-1>

This article has been **cited by** 20 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/330/6008/1216.full.html#related-urls>

This article appears in the following **subject collections**:

Paleontology

<http://www.sciencemag.org/cgi/collection/paleo>

The Evolution of Maximum Body Size of Terrestrial Mammals

Felisa A. Smith,^{1*} Alison G. Boyer,² James H. Brown,¹ Daniel P. Costa,³ Tamar Dayan,⁴ S. K. Morgan Ernest,⁵ Alistair R. Evans,⁶ Mikael Fortelius,⁷ John L. Gittleman,⁸ Marcus J. Hamilton,¹ Larisa E. Harding,⁹ Kari Lintulaakso,⁷ S. Kathleen Lyons,¹⁰ Christy McCain,¹¹ Jordan G. Okie,¹ Juha J. Saarinен,⁷ Richard M. Sibly,¹² Patrick R. Stephens,⁸ Jessica Theodor,¹³ Mark D. Uhen¹⁴

The extinction of dinosaurs at the Cretaceous/Paleogene (K/Pg) boundary was the seminal event that opened the door for the subsequent diversification of terrestrial mammals. Our compilation of maximum body size at the ordinal level by sub-epoch shows a near-exponential increase after the K/Pg. On each continent, the maximum size of mammals leveled off after 40 million years ago and thereafter remained approximately constant. There was remarkable congruence in the rate, trajectory, and upper limit across continents, orders, and trophic guilds, despite differences in geological and climatic history, turnover of lineages, and ecological variation. Our analysis suggests that although the primary driver for the evolution of giant mammals was diversification to fill ecological niches, environmental temperature and land area may have ultimately constrained the maximum size achieved.

For the first 140 million years of their evolutionary history, mammals were small and occupied a fairly narrow range of body sizes and niches (1, 2). Although diverse feeding adaptations evolved by the middle Mesozoic, and larger mammals may have preyed on small dinosaurs (3, 4), their body size range extended only from ~3 to 5 g to ~10 to 15 kg (4, 5). This restricted range almost certainly constrained the ecological roles of early mammals in paleocommunities. For example, herbivory was probably limited; allometric, anatomical, and physiological constraints set a lower threshold of ~5 kg for ruminant herbivores (6). The Cretaceous/Paleogene (K/Pg) mass extinction, which eliminated non-avian dinosaurs as well as many vertebrate, plant, and invertebrate taxa, was followed by a wholesale reorganization of ecological communities

(7). It marked the onset of rapid morphological, ecological, and phylogenetic diversification in terrestrial mammals that led to an expansion in mass by four orders of magnitude and the occupation of a full range of ecological roles (8).

Here we analyze maximum size of terrestrial mammals across different continents, taxonomic groups, phylogenetic lineages, and feeding guilds. We compiled and analyzed data on the maximum body size of each taxonomic order in each sub-epoch on each continent over their entire evolutionary history (9). Information about body mass was obtained for fossil taxa from primary sources or estimated directly from taxon-specific allometric regressions based on measurements of teeth or limbs (table S1). Because of taphonomic considerations, we focused on the maximum size achieved by each order; it tends to be reported in the literature and is robustly related to the overall body size distribution and hence to the mean and median body size (10). Fossil ages were standardized using the midpoint for each Cenozoic sub-epoch on the Gradstein geological time scale (11). Diversity estimates were extracted from the Paleobiology Database (12), using the range-through option for each interval of time. We conducted simulations to assess the potential effect of sampling on the probability of detecting the largest mammal; including as few as 10% of fossil sites yielded nearly 100% probability of recovering the largest mammal on a continent (fig. S1).

The data show that the pattern of body size evolution was similar across continents, lineages, and trophic groups. Globally, and on each continent, maximum body mass increased rapidly during the early Cenozoic (Fig. 1). By the late Eocene [42.9 million years ago (Ma)], maximum body mass was three orders of magnitude larger than at the beginning of the Cenozoic. Our results are consistent with a previous analysis of North American mammals (5, 8). The upper limit of ~17 tons was reached in the early Oligocene of

Eurasia, with the evolution of *Indricotherium transouralicum* (Perissodactyla) and again in the Miocene by several *Deinotherium* species (Proboscidea) in Eurasia and Africa (Fig. 1B; fig. S2); North America never supported a mammal of this size. Strikingly, the overall pattern was not driven by a single taxon or an individual continent. At one time or another, six different orders and three of the four continents contained the largest mammal. Because of the current paucity of data for South America, body mass values for this continent should be considered an underestimate; nonetheless, results illustrate the same general trends. Contrary to earlier suggestions (13–15), increases in body mass were not driven by increasing generic or ordinal diversity: Mammals were not consistently larger when they were more diverse (9) (fig. S3).

We tested two hypotheses for the evolution of maximum body size. The first is a simple growth model, in which maximum body size (M) evolves following a geometric Brownian motion, that is, an unconstrained random walk on the logarithmic scale. This model implicitly assumes that niche space is uniformly distributed. Under a random walk, M is predicted to increase as a power law of the form $\log M = M_0 t^\gamma$, where M_0 is initial maximum body size, t is time, and $\gamma = 1/2$, so that maximum body size increases as the square root of time (15).

The second model has growth saturating over time, reflecting limits of resources or physiological, allometric, biomechanical, or ecological constraints, such as the slower life histories of larger mammals. Thus, the initial change in body mass M with time is proportional to body mass (that is, $\frac{dM}{dt} \propto M$) and increases at some intrinsic rate α . However, as maximum body size evolves, the evolutionary possibilities for increasing size are progressively exhausted. Consequently, the rate of change is also proportional to the availability of open niche space, which is captured by the difference between asymptotic (K) and current log body mass [that is, $\log(K) - \log(M)$], or $\log\left(\frac{K}{M}\right)$. Combining these ecological and evolutionary growth dynamics yields the Gompertz equation $\frac{dM}{dt} = \alpha M \log\left(\frac{K}{M}\right)$, a sigmoidal growth model often used in time series analyses. The integrated

form is $\log M = \log K - \log\left(\frac{K}{M_0}\right)e^{-\alpha t}$, where M_0 is initial maximum body size. The Gompertz model is more biologically plausible than the random walk model, because it captures both the multiplicative nature of body size evolution and the saturating effects of exponentially decreasing niche space availability at larger body sizes.

We compared model fits using corrected Akaike information criteria (AICc). The results suggested that the random walk was not an appropriate model (Table 1). Although a power function provided a reasonable fit to the data, the fitted exponent γ was 0.25, significantly less than the predicted value of 0.50. Moreover, after the initial growth phase, the residuals were not normally distributed. This was probably because maximum

¹Department of Biology, MSC03 2020, University of New Mexico, Albuquerque, NM 87131, USA. ²Department of Ecology and Evolutionary Biology, Yale University, Box 208106, New Haven, CT 06520, USA. ³Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, Santa Cruz, CA 95064, USA. ⁴Department of Zoology, Tel-Aviv University, Tel Aviv 69978, Israel. ⁵Department of Biology and Ecology Center, Utah State University, Logan, UT 84322, USA. ⁶School of Biological Sciences, Monash University, VIC 3800, Australia. ⁷Department of Geosciences and Geography, Institute of Biotechnology, Post Office Box 64, FIN-00014 University of Helsinki, Finland. ⁸Odum School of Ecology, 140 East Green Street, University of Georgia, Athens, GA 30602, USA. ⁹Landscape Ecology, Tvistevägen 48, Department of Ecology and Environmental Science, Umeå University, Umeå, Sweden SE-90187. ¹⁰Smithsonian Institution, Post Office Box 37012, MRC 121, Washington, DC 20013–7012, USA. ¹¹Department of Ecology and Evolutionary Biology, CU Natural History Museum, Campus Box 265, University of Colorado at Boulder, Boulder, CO 80309–0265, USA. ¹²School of Biological Sciences, Harborne Building, University of Reading, Reading, UK. ¹³Department of Biological Sciences, 2500 University Drive North West, University of Calgary, Calgary, Alberta T2N 1N4, Canada. ¹⁴Department of Atmospheric, Oceanic, and Earth Sciences, George Mason University, Fairfax, VA, USA.

*To whom correspondence should be addressed. E-mail: fasmith@unm.edu

body size approaches a plateau as opposed to increasing monotonically. The Gompertz model provided a much better fit to the data throughout the time series and yielded the lowest AICc (Table

1 and fig. S2). The inflection point between the growth phase and the saturating phase occurred during the late Eocene at 42.9 Ma, at a body mass of 4850 kg.

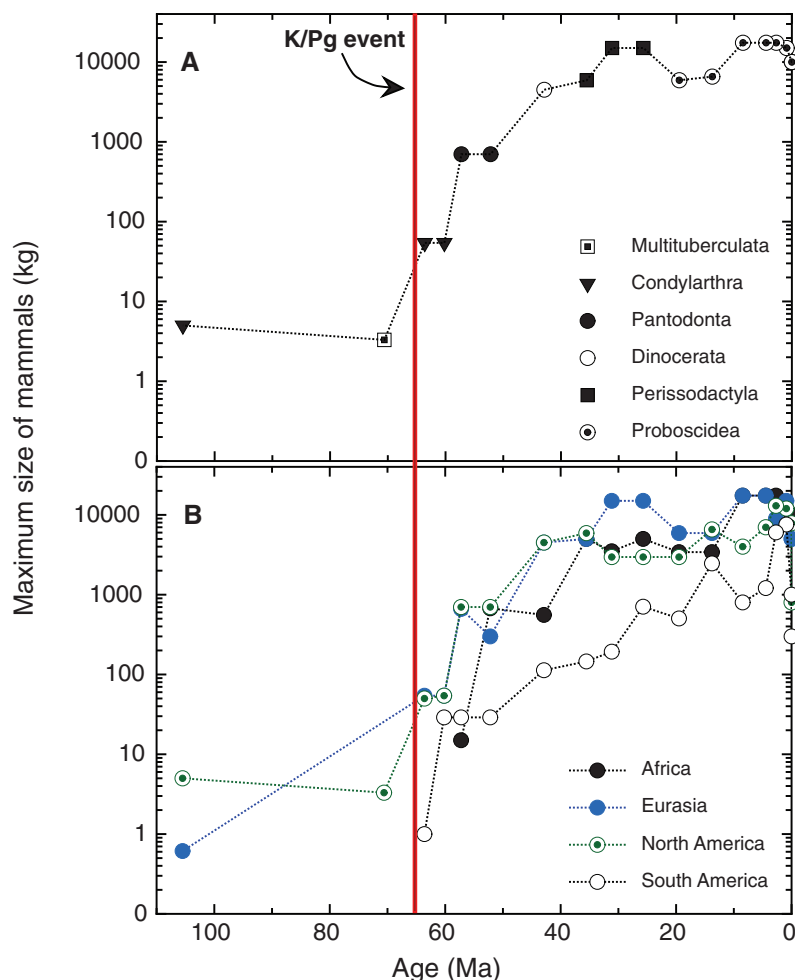


Fig. 1. Maximum body mass of terrestrial mammals over time and space. (A) Maximum body mass over time examined globally at the sub-epoch level over the past 110 million years. (B) Maximum body mass for the largest continents (South America, North America, Africa, and Eurasia) over the same time interval. The overall trend is not driven by a single taxonomic order or an individual continent; six different orders and three of the four continents depicted have at one time or another housed the largest mammal. Data for Australia (not shown) and South America were particularly difficult to obtain because of limited material and/or collecting; thus, estimates for these continents should be considered underestimates. Data are binned at the resolution of sub-epochs using the Gradstein time scale (12).

Table 1. Model fits for global, continental and trophic level body size trajectories. The power law is of the form $\log M = c_0 t^\gamma$ and the Gompertz equation $\log M = \log K - \log\left(\frac{K}{M_0}\right)e^{-\alpha t}$.

Model	Parameters	AICc	R ² value	P value
All data				
Power law	$c_0 = 1.504, \gamma = 0.25$	9.3	0.92	<0.001
Gompertz	$K = 13182.57, M_0 = 6.92, \alpha = 0.08$	8.2	0.94	<0.001
Eurasia				
Gompertz	$K = 15977.18, M_0 = 25.14, \alpha = 0.05$	—	0.83	<0.001
Africa				
Gompertz	$K = 12900.31, M_0 = 0.44, \alpha = 0.06$	—	0.86	<0.001
North America				
Gompertz	$K = 6675.75, M_0 = 8.78, \alpha = 0.07$	—	0.85	<0.001
Carnivores				
Gompertz	$K = 710.56, M_0 = 14.62, \alpha = 0.10$	—	0.76	<0.001

The Gompertz model also provided good fits for the trajectories of maximum body size on each continent (Table 1 and fig. S2). Fifteen different lineages, representative of different archaic and modern orders (such as Proboscidea, Perissodactyla, Artiodactyla, Dinocerata, Pantodonta, Condylarthra, Xenarthra, etc.) evolved similar maximum size at different times and on different continents. These results show that the sigmoidal or saturating trajectory of maximal size evolution for Cenozoic mammals in North America (5, 8) occurred independently in multiple lineages on all the large continents. These results support the interpretation that similar niches were available to and filled by comparably sized giant mammals on each continent after 35 to 40 Ma. Because these niches were occupied by multiple different lineages at different times and on different continents, the patterns suggest that large mammals convergently evolved to fill similar ecological roles. Consistent with this idea, the largest mammals after the beginning of the Cenozoic were always herbivores. These patterns are also congruent with arguments relating the maximum body size of contemporary herbivorous mammals to constraints of diet and digestive physiology (16).

Carnivorous mammals showed similar saturating trajectories but attained smaller maximum sizes than coexisting mega-herbivores (Fig. 2). Large mammal-eating mammals were effectively absent in the early Paleocene; instead, terrestrial crocodiles, snakes, and large lizards were the dominant carnivores (17). Once carnivorous mammal guilds began to diversify, however, they showed a similar trajectory to that of the herbivores—also well fit by a Gompertz function (Table 1). Although carnivores and herbivores started from a similar size immediately after the K/Pg, after ~30 million years the largest carnivores approached an asymptotic maximum about one order of magnitude smaller than that of the largest herbivores (Fig. 2). As with herbivores (Fig. 1A), the carnivores convergently evolved similar maximum sizes in different lineages: the archaic orders Creodonta and Mesonychia, and the modern order Carnivora. Although the duration of these clades overlapped, there was turnover in the ordinal affiliation of the largest carnivore, with each sequentially evolving to a maximum body mass of ~1000 kg (Fig. 2). After the initial size increase, the ratio of body masses of coexisting carnivorous and herbivorous mammals remained similar across the entire Cenozoic (Pearson correlation = 0.819, $P < 0.000$; fig. S4). This suggests at least an indirect relation in which the maximal sizes of carnivores followed the overall size distribution of mammals, but not necessarily a direct causal relation between the largest carnivores and herbivores. Indeed, the largest carnivores probably did not prey on the largest herbivores. The disparity in maximum size between carnivores and herbivores persists in contemporary mammals: Lions, tigers, and bears are about an order of magnitude smaller than elephants and rhinos. The asymptotic maximum size of carnivores of

~1000 kg is consistent with the recent prediction that this represents an upper limit for flesh-eating terrestrial mammals because of physiological and ecological constraints (18).

We compared the overall global trajectory of maximum body mass with time series of three major abiotic factors: global temperature (19), atmospheric oxygen levels (20), and terrestrial

land area (21) (Fig. 3 and table S1). Each of these variables has been hypothesized theoretically and sometimes shown empirically to affect body size evolution in mammals: temperature by affecting how mammals dissipate heat through Bergmann's rule (22–24); greater land area by allowing larger populations and reducing extinction probabilities for the largest mammals (25, 26); and higher atmospheric oxygen concentrations by allowing higher rates of metabolism and biomass production (27–29). We averaged the abiotic values, which were generally reported at a finer scale, using the durations for each geological sub-epoch so we could compare against the trajectory of global body mass over the Cenozoic (table S1). Binned values are superimposed over the finer-scale data shown in Fig. 3. Our analyses were not based on specific values and slopes of these curves at specific times. We varied bin widths and averaging techniques; results were robust with regard to the binning technique employed (9). These abiotic records are based on proxies (19, 21) or on modeling of carbon isotopic records (20); hence, they contain significant unresolved uncertainties, which complicate interpretations of the patterns.

Fig. 2. The trajectory of body mass evolution of selected trophic guilds over the Cenozoic. Green solid circles, herbivores; red open circles with dots, carnivores. Carnivore maximum body mass closely tracks that of herbivores (fig. S4). The ceiling for maximum size is different for herbivores and carnivores (~10 to 15 tons versus ~1 ton) but consistent over time within a trophic group, irrespective of taxonomic affiliation. The largest mammals before the K/Pg may have been omnivorous rather than strict herbivores; our interpretations are based solely on patterns for the Cenozoic.

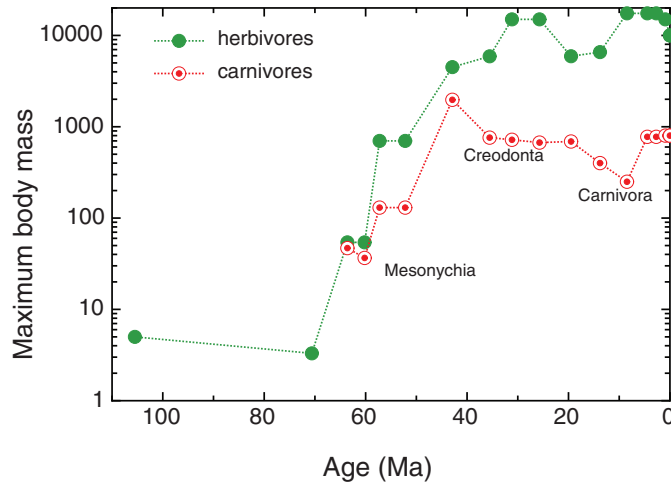
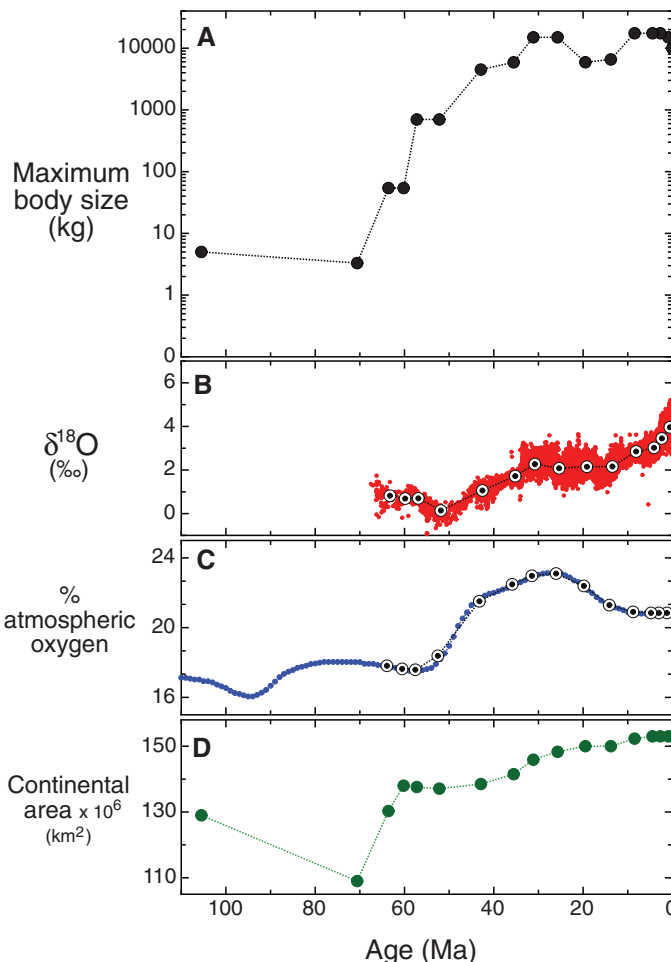


Fig. 3. Global fluctuations in maximum body size and various abiotic factors over the Cenozoic. (A) Maximum body mass (this study). (B) $\delta^{18}\text{O}$ adjusted in parts per thousand (19). (C) Percent atmospheric oxygen concentration (20). (D) Variations in terrestrial land area in square kilometers (21). Abiotic factors chosen were those that have been demonstrated or postulated to influence the body mass of mammals. Linear regression yields highly significant fits between all factors and global mammalian body mass over the Cenozoic; all but atmospheric oxygen concentration remain significant when data are restricted to the past 42.9 million years (table S3). Values for the Mesozoic were excluded because data are fairly scarce; data were truncated at the terminal Pleistocene to yield roughly comparable sub-epoch durations. Data vary in the fineness of their resolution; abiotic variables were binned to obtain values representing the averages of the temporal span represented by each body mass estimate. These binned values are superimposed on the finer-scale data.



All abiotic factors were significantly related to mammalian body mass over the Cenozoic (Fig. 3 and table S3). To determine whether significance was driven by the initial exponential phase, we also ran analyses using the temporal interval from the late Eocene through the Pleistocene (42.9 to 0.9 Ma; results were similar when early or middle Eocene values were chosen). Both global temperature and terrestrial land area remained highly significant: The largest mammals evolved when Earth was cooler and terrestrial land area was greater (table S3), but atmospheric oxygen level dropped out (table S3). However, as might be expected, temperature and land area were significantly related (Pearson correlation = 0.904, $P < 0.001$, $df = 13$): Lower global environmental temperatures (indexed by ^{18}O) corresponded to more water stored in ice caps, lower sea levels, and increased land areas, and probably to changes in vegetation cover and primary productivity.

That temperature and/or land area may have influenced the evolution of body mass in mammals is consistent with several well-established biogeographic principles. The influence of temperature is consistent with Bergmann's rule, a well-known ecogeographic trend of larger body mass in cooler habitats across space (24), and in a few instances, across time (30). Bergmann's rule probably reflects physiological adaptations to prevent heat loss, because larger animals have a reduced surface-to-volume ratio; or alternatively, to promote heat dissipation at smaller body masses (24). Our results are also consistent with the hypothesis that available land area constrains the upper body mass limit of mammals by limiting population through the size or number of home ranges that can be "packed in" or by reducing energy acquisition (25, 26). Among contempo-

rary mammals, maximum body mass is strongly influenced by terrestrial land area, with larger-bodied mammals being found in larger insular or continental "islands" (fig. S5). Thus, constraints on maximum body size potentially imposed by both abiotic factors ultimately may be traced to physiological processes related to endothermy.

However, some caution should be used in the interpretation of our results. Quantitative analyses of these abiotic variables were complicated by a lack of resolution, potential collinearities, and a lack of statistical power that precluded the use of more-rigorous tests to fully explore the relationships between the predictor variables. Moreover, for some of these abiotic factors the uncertainties are not well characterized, and we currently have no way of knowing how these may interact to influence our results. For example, the oxygen isotope curve is confounded by changes in the terrestrial ice volume, atmospheric oxygen concentration is related to temperature through fluctuations in carbon dioxide and carbon sequestration (19) and potentially to global land area through changes in primary productivity, and global land area is clearly related to temperature and sea level. Moreover, other factors such as changes in seasonality and precipitation were not explicitly incorporated; the late Cenozoic saw a global trend toward cooler, drier, and more seasonal climates (19, 31). Nevertheless, the potential role of abiotic factors in the overall trajectory of mammalian evolution cannot be ignored, and the available data suggest interesting and important trends, which should be explored further.

Our analysis implies that the increase in the maximum mass of mammals over the Cenozoic was neither a statistical inevitability driven by increasing species richness nor a random evolutionary walk from a small initial size, but rather

reflected processes operating consistently across trophic and taxonomic groups, and independent of the physiographic history of each continent. We find no support for other hypotheses for the evolution of maximum body mass (9), including the expected increase in variance due to random divergence from a common ancestor or to increasing species richness (13–15); nor do terrestrial mammals ever approach sizes that might invoke biomechanical constraints (32). The K/Pg extinction provided the ecological opportunity for mammals to become larger. Terrestrial mammals did so in an exponentially decreasing fashion, reaching a more or less maximal size by 40 Ma as evolutionary possibilities for increasing body size were progressively exhausted and abiotic factors began constraining the upper limit.

References and Notes

- Z.-X. Luo, *Nature* **450**, 1011 (2007).
- A. W. Crompton, in *Comparative Physiology: Primitive Mammals*, K. Schmidt-Nielsen, L. Bolis, C. R. Taylor, Eds. (Cambridge Univ. Press, Cambridge, 1980), pp. 1–12.
- Y. Hu, J. Meng, Y. Wang, C. Li, *Nature* **433**, 149 (2005).
- Q. Ji, Z. X. Luo, C. X. Yuan, A. R. Tabrum, *Science* **311**, 1123 (2006).
- J. Alroy, *Syst. Biol.* **48**, 107 (1999).
- P. J. Van Soest, *The Nutritional Ecology of the Ruminant* (Cornell Univ. Press, ed. 2, Ithaca, NY, 1994).
- J. A. Wolfe, G. R. Upchurch, *Nature* **324**, 148 (1986).
- J. Alroy, *Science* **280**, 731 (1998).
- Supporting material is available on Science Online.
- A. W. Trites, D. Pauly, *Can. J. Zool.* **76**, 886 (1998).
- F. M. Gradstein, J. G. Ogg, A. G. Smith, *A Geologic Time Scale 2004* (Cambridge Univ. Press, Cambridge, 2004).
- www.paleodb.org/cgi-bin/bridge.pl
- S. J. Gould, *J. Paleontol.* **62**, 319 (1988).
- D. W. MacShea, *Evolution* **48**, 1747 (1994).
- J. Trammer, *Evolution* **59**, 941 (2005).

- M. Clauss *et al.*, *Oecologia* **136**, 14 (2003).
- G. F. Gunnell, M. E. Morgan, M. C. Maas, P. D. Gingerich, *Paleogeog. Paleoclim. Paleocool.* **115**, 265 (1995).
- C. Carbone, A. Teacher, J. M. Rowcliffe, *PLoS Biol.* **5**, e22 (2007).
- J. C. Zachos, G. R. Dickens, R. E. Zeebe, *Nature* **451**, 279 (2008).
- P. G. Falkowski *et al.*, *Science* **309**, 2202 (2005).
- A. G. Smith, D. G. Smith, B. M. Funnell, *Atlas of Mesozoic and Cenozoic Coastlines* (Cambridge Univ. Press, Cambridge, 1994).
- C. M. Janis, *Annu. Rev. Ecol. Syst.* **24**, 467 (1993).
- G. F. Gunnell, in *Evolution of Tertiary Mammals of North America*, C. M. Janis, K. M. Scott, L. L. Jacobs, Eds. (Cambridge Univ. Press, Cambridge, 1998), pp. 91–109.
- B. Rensch, *Evolution above the Species Level* (Columbia Univ. Press, New York, 1959).
- P. A. Marquet, M. L. Taper, *Evol. Ecol.* **12**, 117 (1998).
- G. P. Burness, J. Diamond, T. Flannery, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14518 (2001).
- H. Tappan, in *Molecular Oxygen in Biology: Topics in Molecular Oxygen Research*, O. Hayaishi, Ed. (North-Holland, Amsterdam, 1974), pp. 81–135.
- J. L. Payne *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 24 (2009).
- R. H. Peters, *The Ecological Implications of Body Size* (Cambridge Univ. Press, Cambridge, 1983).
- F. A. Smith, J. L. Betancourt, J. H. Brown, *Science* **270**, 2012 (1995).
- G. Retallack, *J. Geol.* **109**, 407 (2001).
- R. M. Alexander, *Palaentology* **4**, 1231 (1998).
- This project was supported by the Integrating Macroecological Pattern and Process across Scales (IMPSS) NSF Research Coordination Network (NSF grant DEB-0541625 to F.A.S., S.K.M.E., and S.K.L.). This is IMPSS publication no. 11.

Supporting Online Material

www.sciencemag.org/cgi/content/full/330/6008/1216/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 to S3
References

8 July 2010; accepted 18 October 2010
10.1126/science.1194830

Modular Organic Structure-Directing Agents for the Synthesis of Zeolites

Raquel Simancas,¹ Djamel Dari,^{1,2} Noemí Velamazán,¹ María T. Navarro,¹ Angel Cantín,¹ José L. Jordá,¹ Germán Sastre,¹ Avelino Corma,^{1*} Fernando Rey¹

Organic structure-directing agents (OSDAs) are used to guide the formation of particular types of pores and channels during the synthesis of zeolites. We report that the use of highly versatile OSDAs based on phosphazenes has been successfully introduced for the synthesis of zeolites. This approach has made possible the synthesis of the elusive boggsite zeolite, which is formed by 10- and 12-ring intersecting channels. This topology and these pore dimensions present interesting opportunities for catalysis in reactions of industrial relevance.

Zeolites are crystalline microporous and mesoporous materials (1–4) that offer a wide range of applications because of their well-defined structures, which are formed by channels with pore apertures of molecular dimensions. An important objective during the synthesis of zeolites is to achieve control of the pore dimensions and their connectivity through the use of organic structure-directing agents (OSDAs) that, at the

limit, could act as template molecules. A large variety of quaternary organic ammonium salts have been successfully used as OSDAs (2, 4–6) as well as analogous molecules, such as phosphonium-derived organic cations (7–10). However, rather than design new molecules for each zeolite target it could be more efficient to have a type of OSDA that could be easily built by blocks similar to Legos, with a large variety of substituents. Poten-

tial new structures could be simulated with molecular modeling techniques, and an OSDA that directs its synthesis by minimizing the energy of the zeolite-OSDA system could be predicted or at least can be selected from a limited number of candidates.

The described procedure requires having a tool box of OSDA molecules that are easy to prepare and adapt while having the adequate polarity and basicity. We present a type of OSDA molecule with a nearly unlimited synthesis flexibility that is based on building-block units. These molecules are based on phosphazenes that can mobilize silica, have the adequate polarity and stability, and offer more structural possibilities than quaternary ammonium or phosphonium cations. We used these OSDAs for the synthesis of new zeolite structures,

¹Instituto de Tecnología Química, Consejo Superior de Investigaciones Científicas–Universidad Politécnica de Valencia (CSIC-UPV), Avenida de los Naranjos s/n, Valencia 46022, Spain.
²Laboratoire de Chimie Des Matériaux, Université Oran Es-Senia, Elmenaaour, Oran 1524, Algeria.

*To whom correspondence should be addressed. E-mail: acorma@itq.upv.es



Macroecology meets macroeconomics: Resource scarcity and global sustainability



James H. Brown^{a,b,*}, Joseph R. Burger^a, William R. Burnside^c, Michael Chang^a, Ana D. Davidson^d, Trevor S. Fristoe^a, Marcus J. Hamilton^b, Sean T. Hammond^a, Astrid Kodric-Brown^a, Norman Mercado-Silva^{a,e}, Jeffrey C. Nekola^a, Jordan G. Okie^f

^a Department of Biology, University of New Mexico, Albuquerque, NM 87131, United States

^b Santa Fe Institute, Santa Fe, NM, United States

^c National Socio-Environmental Synthesis Center, Annapolis, MD, United States

^d Department of Ecology & Evolution, Stony Brook University, Stony Brook, NY, United States

^e School of Natural Resources and the Environment, University of Arizona, Tucson, AZ, United States

^f School of Earth and Space Exploration, Arizona State University, Tempe, AZ, United States

ARTICLE INFO

Article history:

Received 29 November 2012

Received in revised form 17 July 2013

Accepted 18 July 2013

Available online 27 August 2013

Keywords:

Ecological economics

Economic growth

Human ecology

Macroecology

Resource scarcity

Global sustainability

ABSTRACT

The current economic paradigm, which is based on increasing human population, economic development, and standard of living, is no longer compatible with the biophysical limits of the finite Earth. Failure to recover from the economic crash of 2008 is not due just to inadequate fiscal and monetary policies. The continuing global crisis is also due to scarcity of critical resources. Our macroecological studies highlight the role in the economy of energy and natural resources: oil, gas, water, arable land, metals, rare earths, fertilizers, fisheries, and wood. As the modern industrial-technological-informational economy expanded in recent decades, it grew by consuming the Earth's natural resources at unsustainable rates. Correlations between per capita GDP and per capita consumption of energy and other resources across nations and over time demonstrate how economic growth and development depend on "nature's capital". Decades-long trends of decreasing per capita consumption of multiple important commodities indicate that overexploitation has created an unsustainable bubble of population and economy.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

The greatest challenge of the 21st Century is to secure a sustainable future for humanity. Our informal Human Macroecology Group at the University of New Mexico is one of several collaborative groups investigating the biophysical capacity of the Earth to support human populations and economies. Our approach is "macroecological". By "macro" we mean that our research, based mostly on statistical analysis of large datasets, considers a wide range of spatial and temporal scales, from local to global and from years to millennia. By "ecological" we indicate that our focus is on human-environment relationships, especially the flows of energy, materials, and information which obey well-established physical laws and biological principles, but have uniquely human features.

Our guiding principle is that there is much to be learned by studying humans from an explicitly ecological perspective – a perspective that should be complementary to, but is largely missing from the social sciences and from socioeconomic policy (Burnside et al., 2011).

Much of our work has focused on dependence on resources for population growth and economic development (Brown et al., 2011; Burger et al., 2012; Nekola et al., 2013). The results of our analyses provide a sobering perspective on the current economic situation – and one that contrasts with that of most economists. The global recession of 2008 was the deepest and most long-lasting since the Great Depression. It is not over yet. To recover completely and prevent an even greater crash, most economists and policy-makers are calling for economic growth. The implication is that if we can just get the right monetary, fiscal, and social policies implemented, then unemployment and deficits will go down, housing and industry will rebound, and the economy will start growing again at a healthy pace. This perspective comes from considering only the internal workings of the economy. But why is the recession global? Why is it so severe and long-lasting? Why is the

* Corresponding author at: Department of Biology, University of New Mexico, Albuquerque, NM 87131, United States. Tel.: +1 505 277 9337; fax: +1 505 277 0304.

E-mail address: jhbrown@unm.edu (J.H. Brown).

prescribed economic growth so hard to achieve? These are not just matters of jobs and deficits. The fundamental underlying cause of the decades-long economic trends that culminated in the current recession is depletion of global natural resources. Economic growth and development depend on more than moving money, people, and information; on more than capital and labor, principal and interest, credit and debt, taxation and investment. They also depend on “nature’s capital” (e.g., Costanza et al., 1997; Daily, 1997). Economies extract energy and material resources from the Earth and transform them to produce goods and services. In the last few decades critical resources have been overexploited (Goodland, 1995; Wackernagel and Rees, 1998; Rockström et al., 2009; Bardi, 2011; Burger et al., 2012; The Royal Society, 2012; Wijkman and Rockström, 2013).

2. Background

The human population has grown near-exponentially for about 50,000 years. *Homo sapiens* has expanded out of Africa to colonize the entire world and become the most dominant species in the history of the Earth. Our species has transformed the land, water, atmosphere, and biodiversity of the planet. This growth is a consequence of what we call the Malthusian–Darwinian Dynamic (Nekola et al., 2013). It represents the uniquely human expression of the universal biological heritage that we share with all living things. It has two parts: the Malthusian part, after Thomas Malthus, is the tendency of a population to increase exponentially until checked by environmental limits; the Darwinian part, after Charles Darwin, is the tendency of a population to adapt to the environment in order to push back the limits and keep growing. A special feature of humans is the central role of cultural evolution, which has resulted in rapid changes in behavior, social organization, and resource use.

The expansion of the human population has been accompanied by economic growth and development, and facilitated by technological innovations. The human economy has expanded from the hunting-gathering-bartering economies of subsistence societies to the industrial-technological-informational economies of contemporary civilization. Advances in agriculture used water, fertilizers, new varieties of plants, and animal and mechanical labor to grow food and fiber. Innovations in fisheries supplied additional, protein-rich food. New technologies used wood, bricks, cement, metals, and glass to construct living and working places. Newly developed vaccines and drugs kept parasites and diseases at bay. Energy from burning wood and dung, and subsequently coal, oil, and gas, supplemented with nuclear, solar, wind, and other sources, fueled the development of increasingly complex societies, culminating in our current interconnected civilization with its enormous infrastructure and globalized economy.

How long can recent demographic population and economic trends continue? For more than 200 years, “Malthusians” (e.g., Malthus, 1798; Ehrlich, 1968; Meadows et al., 1972) have argued that the human population cannot continue its near-exponential growth because essential resources supplied by the finite Earth will ultimately become limiting. This perspective has been countered by “Cornucopians” who have argued that there is no hard limit to human population size and economic activity, because human ingenuity and technological innovation provide an effectively infinite capacity to increase resource supply (e.g., Simon, 1981; Barro and Sala-i-Martin, 2003; Mankiw, 2008). So far, both the Malthusians and Cornucopians can claim to be right. Earlier civilizations have grown, flourished, and crashed, but these were always local or regional events (Tainter, 1988; Diamond, 2006). Innovations in agriculture, industry, medicine, and information

technology allowed the global population and its economy to grow (Dilworth, 2010).

Now, however, there is increasing concern that modern humans have depleted the Earth’s energy and material resources to the point where continued population and economic growth cannot be sustained on a global scale (Arrow et al., 1995, 2004; Goodland, 1995; Wackernagel and Rees, 1998; Rockström et al., 2009; Burger et al., 2012; Hengeveld, 2012; Klare, 2012; Mace, 2012; Moyo, 2012; The Royal Society, 2012; Ehrlich and Ehrlich, 2013; Wijkman and Rockström, 2013).

3. Energy

The most critical resource is energy. The development of the modern global industrial-technological-informational economy has been fueled by ever-increasing rates of energy consumption, mostly from fossil fuels. The dependence of economic growth and development on energy is incontrovertible. Much evidence for this is given in papers in this Special Issue by Day et al. (2014) and Hall and Day (2014) (this issue), and in other publications by these and other authors (e.g., Odum, 1971; Smil, 2008; Day et al., 2009; Hall and Day, 2009; Nel and Van Zyl, 2010; Hall and Klitgaard, 2011; Murphy and Hall, 2011; Tverberg, 2012.).

Our Human Macroecology Group has documented how economic development depends on the rate of energy use (Brown et al., 2011; see also references above). As indexed by Gross Domestic Product (GDP), the level of economic development across modern nations varies by nearly three orders of magnitude, from less than \$250 per capita in the poorest countries, such as Somalia, Burundi, and Congo-Kinshasa to more than \$85,000 per capita in the wealthiest, such as Luxembourg, Bermuda, and Norway (The Economist, 2013). There is a strong correlation between per capita GDP and per capita energy use (Fig. 1a). Energy use varies by about two orders of magnitude. In the poorest countries it is barely more than the 100 watts of human biological metabolism. In the richest countries it is more than 10,000 watts, because human metabolism has been supplemented more than 100-fold from exogenous sources, mostly fossil fuels (Brown et al., 2011). Temporal trends over the last few decades show a similar relationship between economic development and energy use (Fig. 1b). From 1980 to 2005 most countries experienced economic growth, accompanied by commensurate increases in energy use. In the few countries where GDP declined, energy consumption usually decreased as well. During the last decade economic growth was especially pronounced in the BRIC countries (Brazil, Russia, India, and China). Fig. 2 contrasts consumption of energy and other resources between 2000 and 2010 for China, where GDP increased more than 15% per year, and the US, where GDP grew by less than 4%.

The causal link between energy use and economic development is easy to understand. Just as a growing human body needs increasing amounts of food, a growing economy needs increasing quantities of energy, water, metal ores, and other resources. Fig. 1a shows that per capita energy use scales with approximately the 3/4 power of per capita GDP across nations (i.e., the slope of the log–log plot in Fig. 1a is 0.76). This means that the rate of energy use scales with GDP on a per individual basis similarly to the 3/4 power scaling of metabolic rate with body mass in mammals, often referred to as Kleiber’s rule (Kleiber, 1961). This similarity may not be coincidental. Both mammalian bodies and modern economies are sustained by consumption of energy supplied through complex branching networks (West et al., 1997). Regardless of whether the approximately 3/4 power scaling is due to a deep causal relationship or an amazing coincidence, both relationships reflect similar

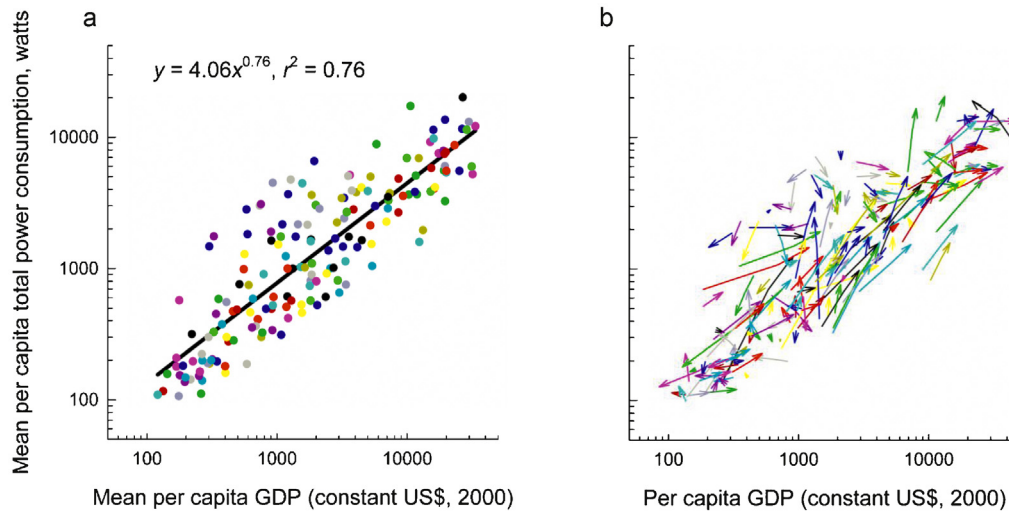


Fig. 1. Relationships between per capita energy use and per capita GDP: (a) Across countries, with each point representing the average energy consumption and the average GDP from 1980 to 2005; (b) over time, with each arrow showing the net change from 1980 to 2005. Note that per capita energy consumption scales as the 0.76 power of GDP (a), and the changes in energy consumption over the 25 years (b) parallels this scaling relationship. Replotted using data compiled by Brown et al. (2011).

underlying causes – the energy cost of maintaining the structure and function of a large, complex system.

The relationships in Fig. 1a can be used to develop future scenarios (Table 1; Brown et al., 2011). We emphasize that these are not predictions; they are simply extrapolations of current patterns of energy use and GDP. Nevertheless, the implications of these scenarios for “sustainable development” are sobering. As classically defined in the Report of the Brundtland Commission (1987, see also United Nations Development Programme, 2011), “Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs”. According to Table 1, to bring the current global population up to a US standard of living would require nearly a 5-fold increase in energy consumption, an obvious impossibility.

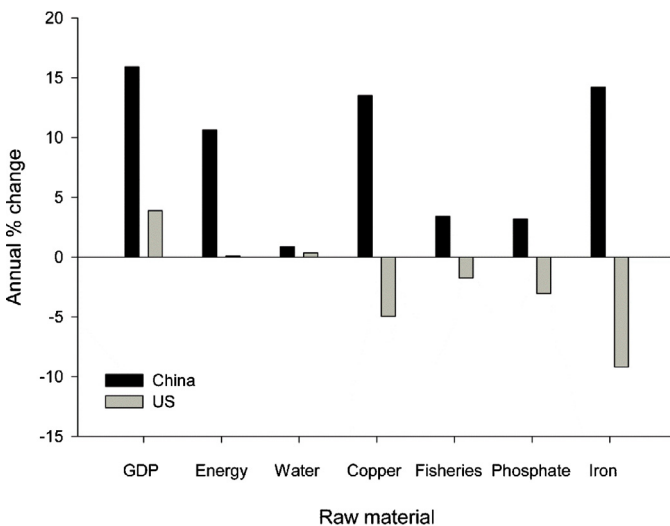


Fig. 2. Annual percent change in GDP and resource consumption for the US and China from 2000 to 2010. China’s economic growth of more than 15% per year was accompanied by commensurate increases in consumption of energy, water, metals, phosphate, and fisheries. Much slower growth of the US economy consumed much less of all these resources. Some of the changes in individual commodities also reflect trends due to globalization. For example, the shift in manufacture and export of electronics from the US to China is reflected in the decrease copper consumption in the US and the large increase in China.

Global energy use could potentially be reduced by 25% by offering everyone on Earth the current average Chinese standard of living, which could theoretically be accomplished by increasing the per capita GDP of poorer countries and decreasing it in richer countries (Brown et al., 2011). Note that China, far from being content with its current standard of living, is striving to grow its GDP as fast as possible (Klare, 2012; Moyo, 2012). More importantly, however, large increases in global energy consumption will be required to meet UN projected population and economic growth for 2025, just 12 years from now (Table 1).

What are the prospects for increasing energy production to meet the scenarios for future development? This is the subject of other papers in this collection and elsewhere. We simply point out that about 85% of current energy use comes from fossil fuels (37% from oil, 25% from gas, and 23% from coal; REN21, 2006). These are finite non-renewable resources. There is good evidence that global oil production has already peaked and will soon do so, and the reserves of gas and coal are being rapidly depleted. Recent increases in oil and gas extraction in North America using hydraulic fracturing technology (<http://www.>

Table 1

Current global energy use and projected energy requirements to meet alternative scenarios of population growth and economic development. These are based on extrapolating the relationship (correlation line) in Fig. 1a. The first column gives total global annual energy requirements in exajoules (EJ = 10¹⁸ J) and the second column gives the factor of increase relative to current consumption. So, for example, to bring the current world population up to a US standard of living would require an approximately 5-fold increase in global energy use, and to provide the entire world with a current Chinese lifestyle in 2025, incorporating UN projected population and economic growth, would require an approximately 2-fold increase.

Scenario	Energy requirement	
	EJ	Factor
World current	524	1.0
U.S. lifestyle	2440	4.7
Chinese lifestyle	392	0.75
Current trends to 2025 ^a	1142	2.2
U.S. lifestyle in 2025 ^a	5409	10.3
Chinese lifestyle in 2025 ^a	848	1.6

After Brown et al. (2011). For sources and calculations see www.jstor.org/stable/10.1525/bio.2011.61.1.7.

^a Assumes 2025 world population of 8 billion and 3.8% per year increase in global GDP.

iea.org/newsroomandevents/pressreleases/2012/november/name_33015.en.html) have simply increased the rate of depletion of the finite stocks. Oil is especially valuable, because it has the highest energy density of any fossil fuel and hence cannot be substituted for in many uses. The problem of “peak oil” is not that this and other finite geological resources (such as metals, phosphate, and rare earths; see below) have been completely used up, but that the rich, easily exploited stocks have been depleted. The remaining reserves are increasingly scarce, dispersed, difficult to extract, and far from human habitation, so the costs of maintaining even current rates of supply are increasing (e.g., [Murphy and Hall, 2011](#); [Tverberg, 2012](#)). Nuclear energy currently accounts for about 6% of global energy use and all renewable energy sources together account for only about 9%. Because large quantities of energy and material resources are required to develop these alternative energy sources (see below and [Hall and Klitgaard, 2011](#); <http://physics.ucsd.edu/do-the-math/2011/10/the-energy-trap>) prospects for increasing energy production sufficiently to meet projected demand are severely limited – and achieving them in the critical next few decades is highly unlikely.

4. Other resources

Energy is not the only essential resource that has been depleted to the point where it is becoming limited. To return to the biological analogy, just as a human being requires not only food energy but also water, protein, vitamins, minerals, clothing, and shelter to grow and survive, so the modern industrial-technological-informational economy requires not only energy but also water, cement, phosphate, metals, and rare earths. Rates of use of all these resources are also closely correlated with energy use and GDP ([Brown et al., 2011](#)). Many of these resources have been consumed to the extent that scarcity has resulted in reduced per capita consumption ([Burger et al., 2012](#); [Klare, 2012](#); [Moyo, 2012](#)). [Fig. 3](#) shows trajectories of global consumption since 1960. Per capita use of all these resources, except for iron, cement, and perhaps molybdenum have peaked, often decades ago. Some of these, such as fossil fuels, metal ores, and phosphate, are non-renewable, and humans have already extracted and burned or dispersed the richest reserves. Others, such as fresh water, fisheries, and wood, are potentially renewable but are being used at unsustainable rates ([Wackernagel and Rees, 1998](#); [Rockström et al., 2009](#); [Burger et al., 2012](#); [Hengeveld, 2012](#); [Klare, 2012](#); [The Royal Society, 2012](#); [Ehrlich and Ehrlich, 2013](#)). Experts in various commodities are beginning to warn not only about peak oil ([Hubbert, 1949](#); [Hirsch et al., 2006](#); [Sorrell et al., 2010](#)) but also about peak water and the over-harvesting of forests and fisheries ([Gleick and Palaniappan, 2010](#); [Foley et al., 2011](#)). It is clear that the [Bruntland Commission's \(1987\)](#) definition of sustainable development has already been violated, because resource use to meet “the needs of the present” has already compromised “the ability of future generations to meet their own needs”.

All of the natural resources in [Fig. 3](#) and many others are important for contemporary humans. Some are required just to keep the present population alive, whereas others are essential for the modern industrial-technological-informational economy. The finite amount of arable land and declining stocks of fresh water, fish (a major protein source), phosphate (an essential fertilizer), and wood (a source of fiber for fuel and housing) mean that major changes in food and shelter will be required to meet projected population growth. Some suggest that the “urban transition”, the trend for an increasing proportion of the population to reside in cities, will allow the Earth to accommodate continued population growth through more efficient use of space and resources (see [Ash](#)

[et al., 2008](#) and the following special issue of *Science*). However, the increased urban populations will need to be fed by a smaller proportion of farmers from a fixed amount of arable land. For rural food production to keep pace with increased urban consumption will require large investments of energy to power machines, and of water and fertilizers to increase yields ([Wackernagel and Rees, 1998](#); [Brown, 2012](#)). Futuristic scenarios in which cities produce a substantial proportion of their own food ([Ehrenberg, 2008](#)), need to be subjected to rigorous biophysical analysis. Even if this were theoretically possible, it may not be feasible, because the necessary changes in urban architecture and landscapes will require large energy and material subsidies.

Large quantities of fresh water and minerals, including copper, iron, molybdenum, nickel, cadmium, platinum, gold, silver, and rare earths are used in industry, including hi-tech electronics and optics. In addition to industrial uses, increased quantities of some minerals will be required to switch from fossil fuels to renewable energy sources. For example, increased deployment of solar energy will require increased use of silicon or cadmium for photovoltaic cells; copper, silver, or other non-magnetic metals for electrical transmission lines; and lead, zinc, nickel, cadmium, or lithium for storage batteries. The quantity of each of these elemental substances in the Earth's crust is fixed. Some of them, such as silicon, lead, and zinc, are relatively abundant, but others are much scarcer. The richest ores near populations have long since been mined, and their contents discarded in landfills and otherwise dispersed. Even though some recycling and substitution will often be possible, increasing quantities of energy and money will have to be expended to find, collect, and purify increasingly scarce minerals in order to maintain supply to meet ever-increasing demand. The result is a rapidly intensifying global race to corner the market ([Klare, 2012](#); [Moyo, 2012](#)). For example, China's rapid industrialization and economic growth in the first decade of the 2000s entailed large increases in consumption of copper and iron as well as energy from fossil fuels ([Fig. 2](#)).

5. Quality of life

Some suggest that level of economic development, often measured as per capita GDP, is a poor measure of what really matters. GDP quantifies the market value of all final goods and services produced in a country per unit time, usually one year. Economists and many others use it as the best available, but admittedly imperfect, index of economic growth and development. There is disagreement, however, on how well GDP measures standard of living (e.g., [Dasgupta and Weale, 1992](#); [United Nations Development Programme, 1990](#)). As an alternative to GDP, some social scientists have promoted the Human Development Index (HDI) or the Genuine Progress Indicator (GPI), which include factors such as life expectancy, education, income distribution, environmental costs, crime, and pollution ([Daly and Cobb, 1994](#); [Klugman, 2010](#); [Posner and Costanza, 2011](#); [United Nations Development Programme, 2011](#); [Kubiszewski et al., 2013](#)).

It has been suggested that the quality of life can be increased with minimal economic impact by eliminating inefficiencies in resource use and extravagant consumption by the wealthiest citizens of the wealthiest nations (e.g., [Diamandis and Kotler, 2012](#); [Jackson, 2012](#)). There is undoubtedly some room for economizing, by both increasing efficiency and eliminating unnecessary consumption. Energy efficiency can be increased by stricter fuel standards for automobiles, better insulation of buildings, improved mass transit, and so on. Substitution, such as renewable energy for fossil fuels and other conductors for copper wires, can reduce the depletion of some severely limited resources. Water can be saved

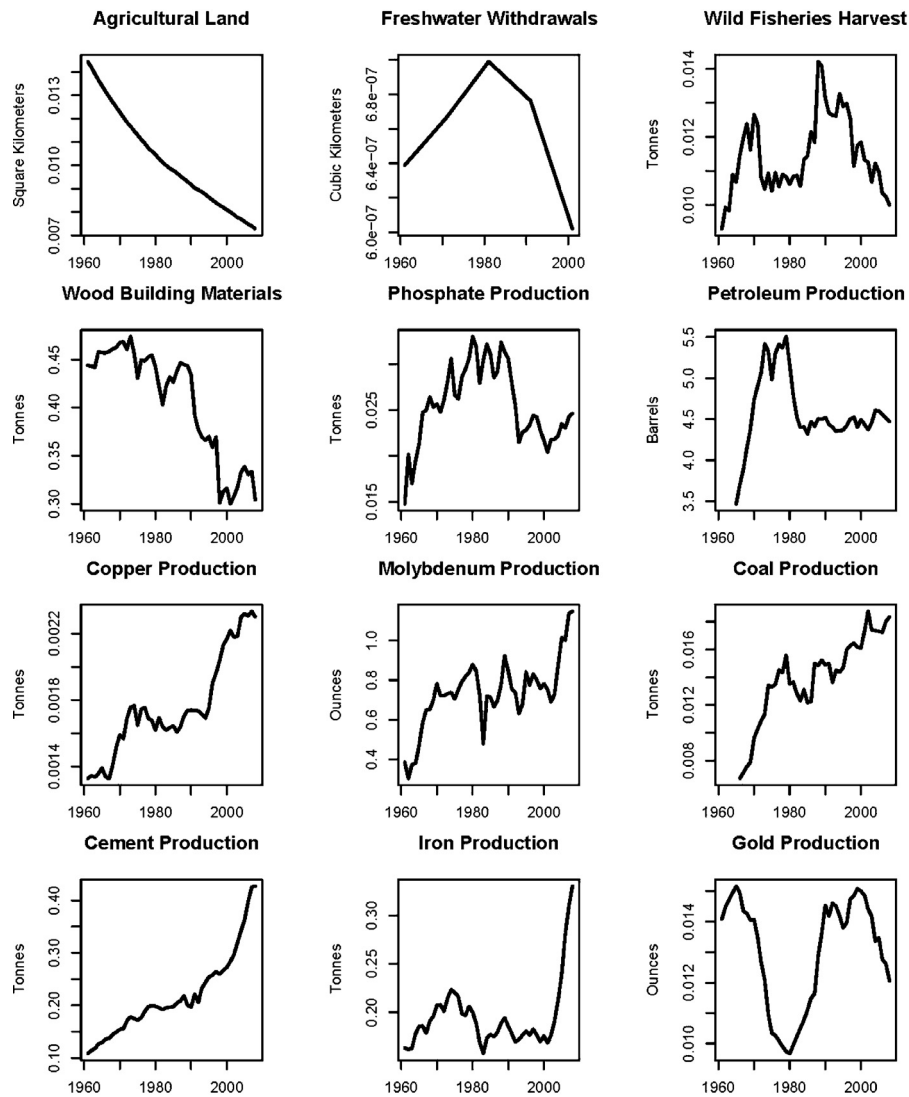


Fig. 3. Trajectory of per capita extraction and consumption of natural resources since the 1960s. Note that per capita supplies of all these resources, except for iron and possibly molybdenum and cement, have peaked, often decades ago, and are now declining. Data sources: per capita values represent the total values divided by global population size as reported by the World Resources Institute (<http://earthtrends.wri.org/>). Individual sources for global production/consumption values are as follows: Agricultural land in km² is from the World Development Indicators Database of the World Bank (<http://data.worldbank.org/data-catalog/world-development-indicators>) and represents the sum of arable, permanent crop, and permanent pasture lands. Freshwater withdrawal in km³ from 1960, 1970, 1980, and 1990 is from UNESCO (http://webworld.unesco.org/water/ihp/db/shiklomanov/part%273/HTML/Tb_14.html) and for 2000 from The Pacific Institute (<http://www.worldwater.org/data.html>). Wild fisheries harvest in tonnes is from the FAO Fishery Statistical Collection Global Capture Production Database (<http://www.fao.org/fishery/statistics/global-capture-production/en>) and is limited to diadromous and marine species. Wood building material production in tonnes is based on the FAO ForeSTAT database (<http://faostat.fao.org/site/626/default.aspx>), and represents the sum of compressed fiberboard, pulpwood+particles (conifer and non-conifer [C & NC]), chips and particles, hardboard, insulating board, medium density fiberboard, other industrial roundwood (C & NC), particle board, plywood, sawlogs+veneer logs (C & NC), sawn wood (C & NC), veneer sheets, and wood residues. Phosphate, copper, molybdenum, pig iron, gold, and combustible coal production data in tonnes is based on World Production values reported in the USGS Historical Statistics for Mineral and Material Commodities (<http://minerals.usgs.gov/ds/2005/140/>). Global coal production data is limited to 1966–2008. Petroleum production in barrels from 1965 to 2008 is based on The Statistical Review of World Energy (<http://www.bp.com/sectiongenericarticle800.do?categoryId=9037130&contentId=7068669>) and represents all crude oil, shale oil, and oil sands plus the liquid content of natural gas where this is separately recovered. These data are reported in 1000 barrels/day, and were transformed to barrels per capita per year. GDP in 1990 US dollars are from the World Resources Institute (<http://earthtrends.wri.org/>). All data were accessed May 2011 to October 2012. After Burger et al. (2012) with new graphs for iron, molybdenum, and gold added.

by behavioral and technological changes that reduce applications to industry and human landscapes and increase water use efficiency of agriculture. Recycling can add to the supply of both abiotic (metal ores, phosphate, water) and biotic (wood fiber) resources, reducing the depletion of the remaining natural stocks. Many kinds of conspicuous consumption, such as gas-guzzling automobiles, lavish climate-controlled houses and workplaces, giant home theater systems, smartphones, jet-set travel, and other extravagances, are obviously not essential to a happy, healthy lifestyle.

Nevertheless, there is little support for the proposition that large reductions in economic activity, and hence in resource consumption, can be achieved without sacrificing what really matters – quality of life (e.g., Costanza et al., 2009; Jackson, 2012; Wijkman and Rockström, 2013; but see Kubiszewski et al., 2013). The HDI and many variables that can be associated with quality of life are closely correlated with GDP (Fig. 4; see also Kelley, 1991). This is not surprising, because all of these variables tend to co-vary with each other, and also with rates of energy and

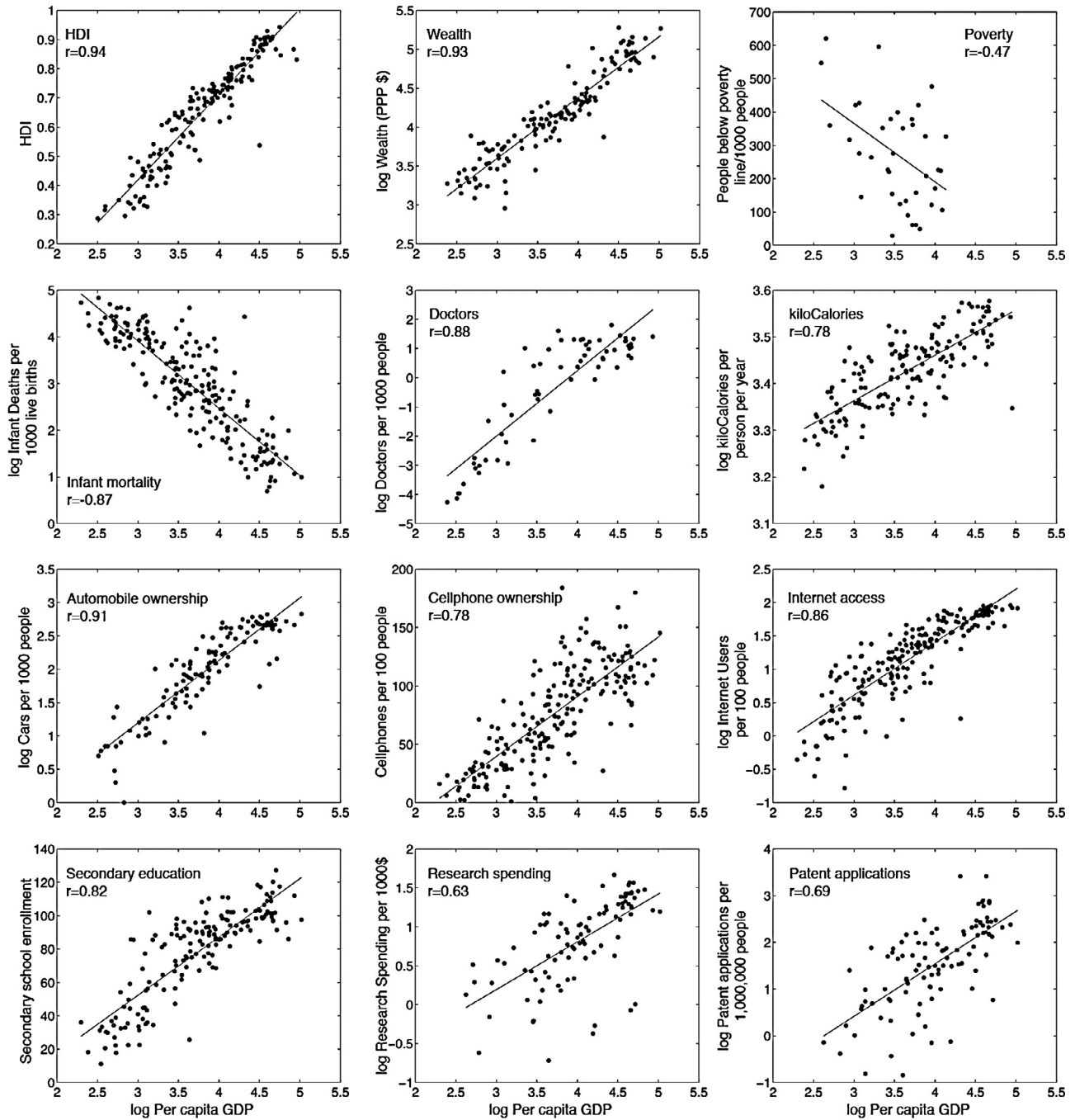


Fig. 4. Variation across countries in relationships between GDP and variables that reflect standard of living and quality of life. First row: overall standard of living: (A) Human Development Index (HDI), (B) per capita wealth, (C) poverty; second row: (D) health: infant mortality, (E) doctors, (F) calories in diet; third row: technology: (G) cars, (H) cell phones, (I) Internet users; fourth row: education and research: (J) secondary education, (K) research spending, (L) patents. In all cases each data point represents the value for a country, GDP is scaled logarithmically and plotted on the x-axis, the other variables are either log-transformed or not, depending on which gives better fit, and correlation coefficients are given. Variables are either per capita or per hundred or thousand population as in the original source. Note that all variables are well correlated with GDP per capita, although the goodness of fit and exact form of the relationships vary. Data from <http://databank.worldbank.org/data/home.aspx> accessed May 2011 to October 2012.

material resource use (Brown et al., 2011). The global per capita GPI peaked in 1978 (Kubiszewski et al., 2013), about the same time that per capita use of oil and several other resources peaked (Fig. 4; Burger et al., 2012) and the global Ecological Footprint exceeded global Biocapacity (http://www.footprintnetwork.org/en/index.php/GFN/blog/today_is_earth_overshoot_day1). There are statistical issues with the relationships shown in Fig. 4: problems of data quality and standardization of measurements across

countries, whether the variables on the Y-axis are scaled linearly or logarithmically, and how to account for the observed variation (i.e., the correlation coefficients). Nevertheless, these relationships go beyond mere correlations to indicate powerful mechanistic processes that require natural resources for economic growth and development. A developed economy with concomitant high rates of energy and other resource use is required to maintain infrastructure, eradicate poverty, and produce drugs, vaccines,

computers, and cell phones. Not only money, but also energy and materials are required to educate teachers, scientists, engineers, and physicians, to build and maintain the infrastructure of housing, workplaces, and transportation and communication facilities, and to train and employ all the people in the public and private service industries. Few people would voluntarily go back to the average lifestyle and standard of living in 1978 when the GPI peaked, even if it were possible to do so. The paper by Day et al. in this special feature (2013) shows how energy shortages will first and most severely reduce discretionary income, as people restrict expenditures to essential food and shelter. Discretionary income provides not only dispensable luxuries but also most things that we associate with quality of life: healthcare, education, science and the arts, travel and recreation. As the economist Milton Friedman is famous for saying, “There is no such thing as a free lunch.” Reductions in energy and material resource use will necessarily require sacrifices in quality of life.

6. Future prospects

So what does the future hold: an imminent end to population and economic growth because we have exceeded the biophysical limits of the finite Earth or a new period of growth and prosperity stimulated by technological innovation; a Malthusian reckoning or a Cornucopian rescue? Currently the global population comprises 7.1 billion people whose standards of living range from abject poverty to extravagant wealth but on average are comparable to typical average residents of China, Indonesia, and Algeria (HDI = 0.67–0.70: *The Economist*, 2013). Future projections of population and economic growth are widely variable and constantly being revised. Optimistic Cornucopian “sustainable development” scenarios for 2050 forecast a global population of 9–10 billion, 3–4% economic growth, and substantial reduction of poverty and disease in developing countries (e.g., *International Council for Science*, 2002; *Millennium Ecosystem Assessment*, 2005; *Sachs*, 2005; *United Nations World Population Prospects*, 2010; *Foley et al.*, 2011; *DeFries et al.*, 2012; *Diamandis and Kotler*, 2012). These are countered by pessimistic Malthusian scenarios (e.g., *Meadows et al.*, 2004; *Bardi*, 2011; *Brown et al.*, 2011; *Burger et al.*, 2012; *Hengeveld*, 2012; *Ehrlich and Ehrlich*, 2013), which suggest that a catastrophic crash is inevitable because the size of the present population and extent of current economic development already far exceed sustainable levels.

One thing is clear: ultimately Malthusian limitations must occur. It is mathematically, physically, and biologically impossible for continual exponential growth in population size and resource use in a finite environment. At some point, food shortages will limit population size or scarcity of other resources will halt economic growth and development. The only questions are when will this occur and what kind of adjustments will it entail?

The answers are uncertain, and we will not make predictions. Global civilization and its economy are complex dynamic systems (e.g., *Strumsky et al.*, 2010; *Tainter*, 2011; *Barnosky et al.*, 2012). Other such systems include hurricanes, forest fires, pandemic diseases, and the stock market. Such systems are composed of many components of many different kinds that interact with each other and with the extrinsic environment on multiple spatial and temporal scales. Their dynamics, driven by a combination of internal feedbacks and external forcings, are highly unpredictable.

We see several lines of evidence that the limits to growth and the concomitant declines in population and economy may be imminent. The first is the fact that per capita use of many resources has been declining for decades (*Burger et al.*, 2012; *Fig. 3*). Some may see the decrease in per capita consumption as encouraging

evidence of increased efficiency. But such “efficiency” is a response to demand increasing faster than supply, with corresponding increases in price. Abundant solar and wind energy have always been available, but they were not heavily used so long as there were abundant supplies of cheap fossil fuels with high energy density. Similarly, increased recycling of metals and wood fiber is an adaptive response to depletion of the richest natural stocks.

Second, contrary to conventional wisdom, most projections in *The Limits to Growth* have been accurate. Re-examination of the computer simulation model of *Meadows et al.* (1972) indicates that nearly all predictions, except for food production, remained on track at least through the early 2000s (*Meadows et al.*, 2004; *Bardi*, 2011, but see *Turner*, 2008). The widespread famines and resulting global population crash predicted by *Ehrlich* (1968) and *Meadows et al.* (1972) were averted primarily by the green revolution: applications of agricultural innovations that increased food production. But the critical technologies – genetic modification, use of supplemental fertilizers and water, and mechanization, implemented in the 1980s and 1990s, not only rely on fossil fuel inputs but also are facing diminishing returns in energy efficiency per unit yield (*Tilman et al.*, 2002). Now the world is again faced with a crisis of food scarcity, with frequent regional famines, thousands of deaths annually, and consequent social and political instability (*Ehrlich and Ehrlich*, 2013).

Third, despite the emphasis of economists, policymakers, and politicians on growth, the global economy has not recovered from the recession of 2008. The magnitude of the crash and the sluggish recovery suggest that, despite abundant unemployed labor, large amounts of corporate capital, and continuing technological innovation, factors outside conventional economic models are restricting growth. There is a surplus of human and monetary capital, but growth is limited by natural capital of energy and raw materials. The economic and political establishments have been slow to recognize and respond to the link between economy and resources. Implicitly, however, there is increasing recognition of the need for natural resources, especially energy, to fuel economic growth and development. There is also increasing recognition that the needed increases in resource production and consumption at the global scale have not occurred.

Finally, there has been far too little scientific, political, and media attention to the question, What is the carrying capacity of the earth for human beings? As *Cohen* (1995) has emphasized, the answer to the question “How many people can the Earth support?” depends on many things, but most importantly on standard of living and concomitant resource use. The present situation would probably not be so dire if meaningful action had been taken when the question of carrying capacity was raised by *Ehrlich* (1968), *Meadows et al.* (1972), and others decades ago. Now this has become the most important scientific and social issue of our time. It should be addressed by our greatest talents, including natural and social scientists, politicians and policymakers, and lay people. Unfortunately, many of the underlying issues, such as population control, equality of economic opportunity, and climate change, are politically charged. Both politicians and the public seem reluctant to confront the specter of a pessimistic future.

Our own assessment is that it is impossible for the Earth to continue to support the present number of people living their current lifestyles. The growth paradigm of traditional economics is no longer compatible with the biophysical carrying capacity of the finite Earth. The economic crash of 2008 and the lack of recovery are due, not to deficiencies in economic policy, but to increasing scarcity of natural resources; not to matters of traditional economics, but to fundamental biophysical constraints on human ecology. Substantial, sustained economic growth and development is no longer possible, because, for the first time in history, human

resource demands exceed global limits on resource supply. In the language of ecology, contemporary humans have exceeded the carrying capacity of the Earth. Unsustainable resource consumption has created a large bubble of population and economy. The bubble cannot keep on increasing: it must either deflate gradually or it will burst. This is not an optimistic assessment, but it must be taken seriously (Meadows et al., 2004; Bardi, 2011; Brown et al., 2011; Burger et al., 2012; Hengeveld, 2012; Ehrlich and Ehrlich, 2013; Wijkman and Rockström, 2013). Wishful thinking, denial, and neglect will not lead to a sustainable future for human civilization.

Acknowledgments

Financial support was provided by National Institutes of Health (grant T32EB009414 for JHB and JRB), National Science Foundation (grants OISE-0653296 and DEB-1136586 for ADD; grant DEB-0541625 for MJH; and grant DBI-1052875 to the National Socio-Environmental Synthesis Center for WRB), a NASA Astrobiology Institute Postdoctoral Program Fellowship (for JGO; administered by Oak Ridge Associated Universities), and Rockefeller Foundation (grant for MJH). We thank many colleagues for stimulating discussions and Paul Ehrlich and two anonymous reviewers for helpful comments on the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ecoleng.2013.07.071>.

References

- Ash, C., Jasny, B.R., Roberts, L., Stone, R., Sugden, A.M., 2008. *Reimagining cities*. *Science* 319, 739.
- Arrow, K., Bolin, B., Costanza, R., Dasgupta, P., Folke, C., Holling, C.S., Jansson, B.-O., Levin, S., Maler, K.-G., Perrings, C., Pimentel, D., 1995. *Economic growth, carrying capacity, and the environment*. *Science* 268, 520–521.
- Arrow, K., Dasgupta, P., Goulder, L., Daily, G., Ehrlich, P., Heal, G., Levin, S., Schneider, S., Starrett, D., Walker, B., 2004. *Are we consuming too much?* *J. Econ. Perspect.* 18, 147–172.
- Bardi, U., 2011. *The Limits to Growth Revisited*. Springer, London.
- Barnosky, A.D., Hadly, E.A., Bascompte, J., Berlow, E.L., Brown, J.H., Fortelius, M., Getz, W.M., Harte, J., Hastings, A., Marquet, P.A., Martinez, N.D., Mooers, A., Roopnarine, P., Vermeij, G., Williams, J.W., Gillespie, R., Kitzes, J., Marshall, C., Matzke, N., Mindell, D.P., Revilla, E., Smith, A.D., 2012. *Approaching a state shift in Earth's biosphere*. *Nature* 486, 52–58.
- Barro, R.J., Sala-i-Martin, X., 2003. *Economic Growth*, 2nd ed. MIT Press, Cambridge, MA.
- Brown, J.H., Burnside, W.R., Davidson, A.D., Delong, J.P., Dunn, W.C., Hamilton, M.J., Mercado-Silva, N., Nekola, J.C., Okie, J.G., Woodruff, W.H., Zuo, W., 2011. *Energetic limits to economic growth*. *Bioscience* 61, 19–26.
- Brown, L.R., 2012. *Full planet, empty plates: the new geopolitics of food scarcity*. WW Norton & Company.
- Brundtland, G.H., 1987. *World Commission on Environment and Development. Our Common Future*. Oxford University Press, New York.
- Burger, J.R., Allen, C.D., Brown, J.H., Burnside, W.R., Davidson, A.D., Fristoe, T.S., Hamilton, M.J., Mercado-Silva, N., Nekola, J.C., Okie, J.G., Woodruff, W.H., Zuo, W., 2012. *The macroecology of sustainability*. *PLoS Biol.* 10 (6), e1001345.
- Burnside, W.R., Brown, J.H., Burger, O., Hamilton, M.J., Moses, M., Bettencourt, L.M.A., 2011. *Human macroecology: linking pattern and process in big picture human ecology*. *Biol. Rev.* 87, 194–208.
- Cohen, J.E., 1995. *How Many People Can the Earth Support?* WW Norton and Company, New York, NY.
- Costanza, R., D'Arge, R., de Groot, R., Farber, S., Grasse, M., 1997. *The value of the world's ecosystem services and natural capital*. *Nature* 387, 253–260.
- Costanza, R., Hart, M., Posner, S., Talberth, J., 2009. *Beyond G.D.P.: the need for new measures of progress*. In: *Pardee Papers*, No. 4, p. 46.
- Daily, G.C. (Ed.), 1997. *Nature's Services: Societal Dependence on Natural Ecosystems*. Island Press, Washington, DC.
- Daly, H.E., Cobb Jr., J.B., 1994. In: Cobb, C.W. (Ed.), *For the Common Good: Redirecting the Economy Toward Community, the Environment, and a Sustainable Future* (No. 73). Beacon Press, Boston.
- Day, J.W., Hall, C.A.S., Yáñez-Arancibia, A., Pimentel, D., 2009. *Ecology in times of scarcity*. *Bioscience* 59, 321–331.
- Day, J.W., Hall, C.A.S., Moerschbaeche, M., Pimentel, D., Yáñez-Arancibia, A., Hall, C.A.S., 2014. *Sustainability and place: how emerging mega-trends of the 21st Century will affect humans and nature at the landscape level*. *Ecol. Eng.* 65, 33–48.
- Dasgupta, P., Weale, M., 1992. *On measuring the quality of life*. *World Dev.* 20, 119–131.
- DeFries, R.S., Ellis, E.C., Stuart Chapin III, F., Matson, P.A., Turner, B.L., Agrawal, A., Crutzen, P.J., Field, C., Gleick, P., Kareiva, P.M., Lambin, E., Liverman, D., Ostrom, E., Sanchez, P.A., Syvitski, J., 2012. *Planetary opportunities: a social contract for global change science to contribute to a sustainable future*. *BioScience* 62, 603–606.
- Diamandis, P.H., Kotler, S., 2012. *Abundance: The Future is Better than You Think*. Free Press, New York, NY.
- Diamond, J.M., 2006. *Collapse: How Societies Choose to Fail or Succeed*. Penguin Group, USA.
- Dilworth, C., 2010. *Too Smart for Our Own Good: The Ecological Predicament of Humankind*. Cambridge University Press, Cambridge, MA.
- Ehrenberg, R., 2008. *Let's get vertical: city buildings offer opportunities for farms to grow up instead of out*. *Science News* 174, 16–20, <http://dx.doi.org/10.1002/scin.2008.5591740818>.
- Ehrlich, P.R., 1968. *The Population Bomb*. Ballantine Books, New York, NY.
- Ehrlich, P.R., Ehrlich, A.H., 2013. *Can a collapse of global civilization be avoided?* *Proc. R Soc. B* 280, 1754.
- Foley, J.A., Ramankutty, N., Brauman, K.A., Cassidy, E.S., Gerber, J.S., Johnston, M., Mueller, N.D., O'Connell, C., Deepak, K., Ray, D.K., West, P.C., Balzer, C., Bennett, E.M., Carpenter, S.R., Hill, J., Monfreda, C., Polasky, S., Rockström, J., Sheehan, J., Siebert, S., Tilman, D., Zaks, D.P., 2011. *Solutions for a cultivated planet*. *Nature* 478, 337–342.
- Gleick, P.H., Palaniappan, M., 2010. *Peak water limits to freshwater withdrawal and use*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11155–11162.
- Goodland, R., 1995. *The concept of environmental sustainability*. *Ann. Rev. Ecol. Syst.* 26, 1–24.
- Hall, C.A.S., Day, J.W., 2009. *Revisiting the limits to growth after peak oil*. *Am. Sci.* 97, 230–237.
- Hall, C.A.S., Day, J.W., 2014. *Why aren't contemporary ecologists and economists addressing resource and energy scarcity: The major problems of the 21st century?* *Ecol. Eng.* 65, 49–53.
- Hall, C.A.S., Klitgaard, K.A., 2011. *Energy and the Wealth of Nations: Understanding the Biophysical Economy*. Springer Publishing, New York, NY.
- Hengeveld, R., 2012. *Wasted World: How our Consumption Challenges the Planet*. University of Chicago Press, Chicago, IL.
- Hirsch, R.L., Bezdek, R., Wendling, R., 2006. *Peaking of world oil production and its mitigation*. *AIChE J.* 52, 2–8.
- Hubbert, M.K., 1949. *Energy from fossil fuels*. *Science* 109, 103–109.
- International Council for Science, 2002. *Science and Technology for Sustainable Development*. International Council for Science, Paris.
- Jackson, T., 2012. *Prosperity Without Growth? Economics for a Finite Planet*. Routledge Press, London, UK.
- Klare, M.T., 2012. *The Race for What's Left: The Global Scramble for the World's Last Resources*. Metropolitan Books, New York, NY.
- Kelley, A.C., 1991. *The human development index: handle with care*. *Popul. Dev. Rev.* 17, 315–324.
- Kleiber, M., 1961. *The Fire of Life: An Introduction to Animal Energetics*. Wiley, New York, NY.
- Klugman, J., 2010. *Human Development Report 2010: the Real Wealth of Nations: Pathways to Human Development*. Macmillan, Palgrave.
- Kubiszewski, I., Costanza, R., Franco, C., Lawn, P., Talberth, J., Jackson, T., Aylmer, C., 2013. *Beyond GDP: measuring and achieving global genuine progress*. *Ecol. Econ.* 93, 57–68.
- Mace, G.M., 2012. *The limits to sustainability science: ecological constraints or endless innovation?* *PLoS Biol.* 10 (6), e1001343, <http://dx.doi.org/10.1371/journal.pbio.100>.
- Malthus, T.R., 1798. *An Essay on the Principle of Population*. Oxford University Press, Oxford, U.K.
- Mankiw, N.G., 2008. *Principles of Macroeconomics*, 5th ed. South-Western, Cengage Learning, Stamford, CT.
- Meadows, D.H., Randers, J., Meadows, D., 1972. *The Limits of Growth*. In: *A Report for The Club of Rome*.
- Meadows, D.H., Randers, J., Meadows, D., 2004. *Limits to Growth: The 30-Year Update*. Chelsea Green Publishing, White River Junction, VT.
- Millennium Ecosystem Assessment, 2005. *Ecosystems and Human Well-being: Synthesis*. Island. Press, Washington, DC.
- Moyo, D., 2012. *Winner Take All: China's Race for Resources and What it Means for the World*. Basic Books, New York, NY, pp. 257.
- Murphy, D.J., Hall, C.A.S., 2011. *Energy return on investment, peak oil, and the end of economic growth*. In: Costanza, R., Limburg, K., Kubiszewski, I. (Eds.), *Ecological Economics Reviews*. Ann. N.Y. Acad. Sci. 1219, 52–72.
- Nekola, J.C., Allen, C.D., Brown, J.H., Burger, J.R., Davidson, A.D., Fristoe, T.S., Hamilton, M.J., Hammond, S.T., Kodric-Brown, A., Mercado-Silva, N., Okie, J.G., 2013. *The Malthusian–Darwinian dynamic and the trajectory of civilization*. *Trends in Ecology and Evolution* 28, 127–130.
- Nel, W.P., Van Zyl, G., 2010. *Defining limits: energy constrained economic growth*. *Appl. Energy* 87, 168–177.

- Odum, H.T., 1971. *Environment, Power and Society*. John Wiley and Sons, New York, NY.
- Posner, S.M., Costanza, R., 2011. A summary of ISEW and GPI studies at multiple scales and new estimates for Baltimore City, Baltimore County, and the State of Maryland. *Ecological Economics* 70 (11), 1972–1980.
- Renewable Energy Policy Network for the 21st Century, 2006. *Renewables, Global Status Report: 2006 Update*. REN21 Steering Committee.
- Rockström, J., Steffen, W., Noone, K., Persson, A., 2009. A safe operating space for humanity. *Nature* 461, 472–475.
- Sachs, J.D., 2005. *The End of Poverty: Economic Possibilities for Our Time*. Penguin Group, USA.
- Simon, J.L., 1981. *The Ultimate Resource*. Oxford University Press, Oxford, UK.
- Smil, V., 2008. *Energy in Nature and Society: General Energetics of Complex Systems*. MIT Press, Cambridge, MA.
- Sorrell, S., Speirs, J., Bentley, R., Brandt, A., Miller, R., 2010. Global oil depletion: a review of the evidence. *Energy Policy* 38, 5290–5295.
- Strumsky, D., Lobo, J., Tainter, J.A., 2010. Complexity and the productivity of innovation. *Syst. Res. Behav. Sci.* 27, 496–509.
- Tainter, J.A., 1988. *The Collapse of Complex Societies*. Cambridge University Press, Cambridge, MA.
- Tainter, J.A., 2011. Energy, complexity, and sustainability: a historical perspective. *Environ. Innov. Soc. Trans.* 1 (2011), 89–95.
- Tilman, D., Cassman, K.G., Matson, P.A., Naylor, R., Polasky, S., 2002. Agricultural sustainability and intensive production practices. *Nature* 418, 671–677.
- Turner, G.M., 2008. A comparison of the limits to growth with 30 years of reality. *Global Environ. Chang.* 18 (3), 397–411.
- The Economist, 2013. *Pocket World in Figures*. Profile Books, London, UK, 2013 Edition.
- Tverberg, G.E., 2012. Oil supply limits and the continuing financial crisis. *Energy* 37, 27–34.
- The Royal Society, 2012. *People and the planet*. In: The Royal Society Science Policy Centre Report 01/12, Issued: April 2012, DES2470.
- United Nations Development Programme, 1990. *Human Development Report 1990*. University Press, Oxford/New York, N.Y.
- United Nations Development Programme, 2011. *Human Development Report 2011*. In: *Sustainability and Equity: A Better Future for All*. Palgrave Macmillan, New York.
- United Nations World Population Prospects: The 2010 Revision, 2010. <http://esa.un.org/wpp/Excel-Data/population.htm>
- Wackernagel, M., Rees, W., 1998. *Our Ecological Footprint: Reducing Human Impact on the Earth, Vol. 9*. New Society Publishers, Gabriola Island.
- West, G.B., Brown, J.H., Enquist, B.J., 1997. A general model for the origin of allometric scaling laws in biology. *Science* 276, 122–126.
- Wijkman, A., Rockström, J., 2013. *Bankrupting Nature: Denying Our Planetary Boundaries*. Routledge.