© 1995 Nature Publishing Group http://www.nature.com/naturemedicine

Genomic prospecting

The importance of preserving biodiversity extends beyond the discovery of new drugs to understanding how other species have dealt with medical problems we currently face.

STEPHEN J. O'BRIEN

There is no shortage of incurable diseases. Cancer, autoimmune disease, emerging viruses and parasites, degenerative pathologies like arthritis and Alzheimer's disease and over 2,000 hereditary diseases drive an US$11 billion National Institutes of Health budget to resolve the physiologic bases, to clone the genes and to move beyond symptomatic treatment for fatal and debilitating conditions that fill our hospitals. The Human Genome Project offers the prospect of identifying the genes affecting maladies from breast cancer and diabetes to werewolf-hair pattern and bipolar neuropathies. And another, largely untapped cache of genetic information has begun to emerge from an unexpected source: the genomes of free-ranging vertebrate populations that have acquired adaptations to abrogate or suppress the same diseases.

Over 99 per cent of the species that ever existed have gone extinct or failed some evolutionary challenge. Today's surviving species are descendants of the winners of ancient struggles who have evolved genetic defenses to infectious agents, cancers, and degenerative disease. Included in the genomes of all species are the genetic footprints of historic epidemics and of forced adaptation to neoplasia and to other life-threatening hazards. As these species have received no medical assistance, outcomes of the natural medical struggles are unforgiving, quite effective and largely undescribed.

Last April, a conference jointly sponsored by the Smithsonian Institution, the National Science Foundation and the National Institutes of Health met to explore the critical role of biological diversity on human health*. There was much discussion on the value of 'natural products,' the cornucopia of compounds discovered in the diverse biota with beneficial pharmaceutical properties. The National Cancer Institute (NCI) maintains a large program that tests thousands of extracts of rare plants, fungi, microbes and marine life for anticancer and anti-HIV properties every year[1]. This search for natural medicines, termed 'chemical prospecting' by Thomas Eisner[2] of Cornell University has led to identification of more than 500 active compounds — from aeroplysinin to zthyhydroxy benzene — that have medicinal activity. The economic and human rationale of conserving earth's biodiversity for medical potential alone was emphasized by Francesca Grifo (American Museum of Natural History), who noted that nine of the ten top-selling pharmaceuticals on the market today are natural products or their derivatives. Of the aesthetic, ethical and practical justifications for species conservation, the preservation of heretofore undiscovered natural product medicinals is to many the most convincing.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Fig. 1 Understanding the African cheetah's unusual genetic history has important implications for HLA variation and immune defenses. (Photograph by Karl Ammann.)

There are other potential medical benefits to species conservation in addition to pharmaceutical discovery. Using animals to interpret basic human biology can be important, as has been shown repeatedly. For centuries experiments in comparative anatomy, physiology, and metabolism dominated medical research. Human basic research is now so advanced that we use human biological data to interpret the status of endangered species[3,4]. How ironic that while advancing technology contributes to the creation of endangered species, it may be biomedical technology that actually provides the information required to reverse imminent extinctions. Another surprise is beginning to emerge from the genetic strategies uncovered in free-ranging and threatened species. Novel genetic adaptations relevant to man's unsolved diseases are being discovered and could lead to innovative therapies. I shall summarize some examples from our own experience that illustrate the medical value of 'genome prospecting' (mining the genomes of natural populations for medically relevant information), recognizing that far more await discovery.

The gene complex with the greatest extent of between-individual allelic variation in all mammals is the major histocompatibility complex (MHC), HLA in humans[5]. Stretched across two million base pairs on the short arm of human chromosome six, the HLA region consists of a collection of about seventy coding genes, of which fifteen mediate immune recognition and clearing of infectious disease agents. Three HLA class I, ten class II, and two TAP genes encode over 300 polymorphic alleles that provide the 'self' context or recognition signature for T cells to mount cellular and humoral (antibody)

Biodiversity and Human Health, April 3–4, 1995, Washington, DC, USA.

© 1995 Nature Publishing Group http://www.nature.com/naturemedicine

responses to invading pathogens. The purpose for this incredible genetic diversity had been a puzzle to immunologists and geneticists for many years. An important clue comes from a natural history study of the world's fastest land animal, the African cheetah (Fig. 1).

Difficulty in breeding cheetahs in zoos prompted a multidisciplinary study that showed cheetahs as a species to be remarkably inbred, having a reduction of 90 to 99 per cent of overall genomic diversity based on molecular genetic markers[6]. Perhaps most surprising was the demonstration that unrelated cheetahs would accept reciprocal skin grafts indicating immunological identity at the cheetah's MHC. The cheetah's genes were likely homogenized by a near extinction event followed by inbreeding in the species' ancestry. When cheetahs became exposed recently to feline infectious peritonitis virus, a coronavirus with low morbidity and mortality in domestic cats (about 1 per cent affected), 100 per cent of infected cheetahs developed symptoms (diarrhea, jaundice, depression, fever) and 60 per cent died[7]. The cheetah's homogeneous sensitivity to this agent in the face of genetic monomorphism at the MHC provided a rationale for the critical importance of MHC genetic variation in natural populations. When pathogens genetically circumnavigate the immune defence of an individual host, genetic diversity of immune response genes, particularly the MHC in the population, provide a 'moving target' that would protect the species from widespread affliction. By contrast, a genetically uniform population like the cheetah loses this advantage and may suffer far more casualties. The cheetah example is frequently quoted as a natural verification of the adaptive value of high MHC and immune gene polymorphism. Human demographic history has some dramatic parallels, particularly the loss of some 50 million native Americans (themselves genetically diminished) from infectious disease introduced by European contact[8].

A different moral for the consequences of a demographic contraction and inbreeding was demonstrated in the Florida panther (*F. concolor coryi*, see Fig. 2), a relict puma subspecies of 30 to 50 highly inbred individuals that teeters on the very edge of extinction in the Big Cypress Swamp Ecosystem in south Florida[9]. The consequences of inbreeding are dramatic: (1) an average of 95 per cent developmentally abnormal sperm per ejaculate, (2) an increase to 85 per cent incidence of cryptorchidism (heritable defect causing one or two testicles to remain undescended) over the past 10 years, (3) 80 per cent have heart murmurs, (4) an abrupt appearance of an atrial septal heart defect (patent foramen ovale, a common congenital human disorder) that has been fatal in at least three panthers and (5) an enormous microbial/parasite load. The Florida panther's genetic legacy proved so compelling that the U.S. Department of Interior decided to genetically augment the fragile population with a different puma subspecies, *F. c. stanleyi*. The Florida panther experience illustrates rather graphically the critical role that rare genetic variants can have on reproductive characteristics and on congenital heart abnormalities in wildlife species and in humans.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Fig. 2 The endangered Florida panther's tale emphasizes the role of rare genetic mutation on reproduction and congenital heart disease. (Photograph by Melody E. Roelke-Parker.)

Some fascinating insight about lethal infectious disease and natural genetic defences appeared from a wild mouse population in a squab farm near Lake Casitas, forty miles north of Los Angeles[10]. Murray Gardner (University of California at Davis) discovered that the mice were suffering a raging epidemic with a retrovirus that caused fatal hind-limb paralysis. But only about 20 per cent of the Lake Casitas mice became infected even though the virus caused 90 per cent mortality when inoculated into laboratory mouse strains. The reason that 80 per cent of the Lake Casitas mice were virus-free was the presence of a powerful trans-dominant restriction gene, *Fv-4*, that blocked infection of the retrovirus. *Fv-4* was subsequently shown to be a truncated version of the exogenous retrovirus. Integrated into chromosome 12 of resistant mice was a portion of retroviral *pol*, a full length *env* and 3'-LTR retroviral genes. In lymphoid cells of protected mice the *env* portion produced an Env glycoprotein that bound and blocked the murine retroviral receptor, thereby conferring resistance to infection. The LC mice had acquired a natural evolutionary solution to a fatal paralytic retroviral disease. In the presence of the retroviral pathogen, the *Fv-4* gene protected its carriers and their offspring, representing a remarkably effective genomic adaptation.

Chromosomally integrated endogenous retroviral genomes have been described in many mammal species including humans, and we have always suspected that they represent 'genomic fossils' of historic pathogenic outbreaks[11,12]. At least 0.1 per cent of the human genome is related to retroviral genomes. It is likely that all of these were acquired by infection into the germ line of a developing embryo during an outbreak that conferred endogenous immunity to the carriers. The wild Lake Casitas mouse population and functional retroviral integrations seen in other species are now considered to be the paradigm for acquisition of endogenous retroviral sequences. Expression of endogenous retroviral genes in humans is rare[12]. Any role they may play in disease resistance is only speculative until an interaction with pathogenic human

© 1995 Nature Publishing Group http://www.nature.com/naturemedicine

retroviruses is demonstrated.

If endogenous retroviral sequences originate as germline infections of naive populations during an outbreak of an exogenous counterpart, is the 'endogenization' process rapid or are there intermediate steps? A hint of an answer can be obtained by considering immunodeficiency viruses in free-ranging species. HIV-1, the aetiologic agent for AIDS, causes a slow drop in CD4-bearing T-lymphocytes and has led to collapse of the immune system, opportunistic infections and death in over 2 million people. The cosmopolitan HIV-1 and the less virulent west African HIV-2 strain erupted in human populations within the past few hundred years[13]. HIV origins are likely to have derived from simian immunodeficiency virus (SIV), a virus that is endemic in free-ranging sooty mangabey, green monkey and other *Cercopithecus* African monkey species. SIV-infected African monkeys do not develop immunodeficiency disease; yet SIV is quite virulent because it does induce AIDS when inoculated into Asian macaques, species that have never seen SIV in their native range[14]. The simplest explanation for the balance between SIV and African monkeys would be that they are descendants of an historic disease-inducing outbreak that effectively selected genetically resistant survivors to repopulate subsequent generations.

A nearly identical scenario has been uncovered with feline immunodeficiency virus (FIV) and wild cat species. FIV was discovered in an immunocompromised domestic cat and, like HIV, induces immune deficiency and death among domestic cats[15]. FIV has been found in sixteen non-domestic cat species (including puma, leopards, lions, and cheetahs) sampled for antibody and genomic sequences[16,17]. However, the FIV in wild cats, like the African monkeys, shows little evidence for immune suppression, associated disease or mortality. Since viral genomic sequences are highly divergent within and between different cat species, we suspect that the virus is relatively old, perhaps coexisting with wild cat species for thousands of years. Because retroviral sequence evolution is even less clocklike (in a molecular evolutionary sense) than gene sequences of host species, it is difficult to be precise about the timing. Nevertheless, the pattern of FIV divergence is generally monophyletic within each species (as if virus in each species entered that species only once and subsequently evolved), supporting the notion that FIV persistence in the felids is rather ancient. Like the monkeys, the cat species may have achieved an evolutionary balance whereby disease is postponed at least until after the reproductive years. Domestic cats, which do become ill, may simply represent a recent emergence of FIV in a heretofore naive species.

Is there an evolutionary explanation for the maintenance of a chronic exogenous FIV infection in wild cats? Is there an adaptive advantage to carrying the virus once historic adaptation has muted its pathogenesis? The answer could involve a temporary acquisition of immune protection against pathogenic FIV which may be present in low frequency or be produced spontaneously by rapid lentiviral mutation. Wild cats infected by an attenuated FIV strain would be immunized naturally against exposure to a virulent strain. An extended period of exogenous virus/host commensalism or symbiosis may provide the time period required for a genomic 'endogenization' event to occur. Although primates and cats have abundant endogenous retroviral sequences, none of these are lentivirus-related. With passenger lentiviruses so common among mammals, it may be simply a matter of time before the birth of endogenous lentivirus sequences in the free-living species. Finally, with the recent development of high resolution comparative gene maps of primate, cat, and other mammalian species, the genetic basis of lentivirus resistance can now be approached[18].

The prospect of genetic engineering to treat human pathologies will be a principal emphasis of twenty-first century medicine. The burgeoning biotechnology industry was born to accomplish this hope. Harold Varmus, Director of NIH, has moved to review gene therapy research and development as a high priority for medical research. In 1988, David Baltimore of the Massachusetts Institute of Technology coined the term 'intracellular immunization' to describe the delivery of virus-restricting genes to cells of sensitive or infected patients as a genetic therapy to virus disease[19]. Attempts to apply this to AIDS are beginning, but to be honest we really do not know which genes to use. The answer may be lurking in wild monkeys, cats, gazelles, woodchucks or armadillos. Molecular biologists and disease gene hunters need only to befriend a field biologist, a naturalist or a wildlife veterinarian. There is little question that time-tested and selectively improved natural solutions to our own medical ills await discovery, and with the rapid advances in medical biotechnology the tools to recognize them are now available. The potential benefits of chemical and genomic prospecting are dazzling, because they promise both to reveal naturally tested solutions for difficult health conditions and to contribute to conservation of the earth's biodiversity.

1. Roberts, L. Chemical prospecting: Hope for vanishing ecosystems. *Science* **256**, 1142–1143 (1992).
2. Flam, F. Chemical prospectors scour the seas for promising drugs. *Science* **266**, 1324–1325 (1994).
3. O'Brien, S.J. A role for molecular genetics in biological conservation. *Proc. natn. Acad. Sci. U.S.A.* **91**, 5748–5755 (1994).
4. O'Brien, S.J. Genetic and phylogenetic analyses of endangered species. *Annu. Rev. Genet.* **28**, 467–489 (1994).
5. Trowsdale, J. Genomic structure and function in the MHC. *Trends Genet.* **9**, 117–122 (1993).
6. O'Brien, S.J. *et al.* Genetic basis for species vulnerability in the cheetah. *Science* **227**, 1428–1434 (1985).
7. Heeney J. L. *et al.* Prevalence and implications of feline coronavirus infections of captive and free-ranging cheetahs (*Acinonyx jubatus*). *J. Virol.* **64**, 1964–1972 (1990).
8. Black, F.L. Why did they die? *Science* **258**, 1739–1740 (1992).
9. Roelke, M.E., Martenson, J.S. & O'Brien, S.J. The consequences of demographic reduction and genetic depletion in the endangered Florida panther. *Curr. Biol.* **3**, 340–350 (1993).
10. Gardner, M.B., Kozak, C. & O'Brien, S.J. The Lake Casitas wild mouse: Evolving genetic resistance to retroviral disease. *Trends Genet.* **7**, 22–27 (1991).
11. Benveniste, R. The contributions of retroviruses to the study of mammalian evolution. In *Molecular Evolutionary Genetics* (ed. MacIntyre, R.J.) 359–417 (Plenum, New York, 1985).
12. Wilkinson, D.A., Mager, D.L. & Leony J.C. Endogenous human retroviruses. in *The Retroviridae* (ed. Levy, J.A.) 465–535 (Plenum, New York, 1994).
13. Desrosiers, R. C. The simian immunodeficiency viruses. *Annu. Rev. Immunol.* **8**, 557–578, 1990.
14. Hirsch, V. M. & Johnson, P.R. Pathogenic diversity of simian immunodeficiency viruses. *Virus Res.* **32**, 183–203, 1994.
15. Pedersen, N. C., Ho, E.W., Brown, M.L. & Yamamoto, J.K. Isolation of a T-lymphotropic virus from domestic cats with an immunodeficiency-like syndrome. *Science* **235**, 790–793, 1987.
16. Olmsted, R.A. *et al.* Worldwide prevalence of lentivirus infection in wild feline species: Epidemiologic and phylogenetic aspects. *J. Virol.* **66**, 6008–6018 (1992).
17. Brown, E.W., Yuhki, N., Packer, C. & O'Brien, S. J. A lion lentivirus related to feline immunodeficiency virus: Epidemiologic and phylogenetic aspects. *J. Virol.* **68**, 5953–5968 (1995).
18. O'Brien, S. J. *et al.* Anchored reference loci for comparative genome mapping in mammals. *Nature Genet.* **3**, 103–112 (1993).
19. Baltimore, D. Intracellular immunization. *Nature* **335**, 395–396 (1988).

Laboratory of Viral Carcinogenesis
National Cancer Institute
Frederick, Maryland 21702, USA