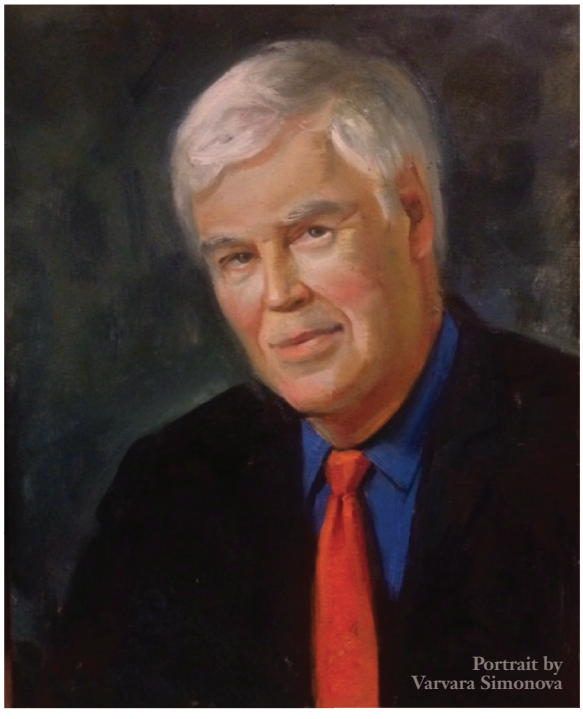

Stephen J. O'Brien

ANNUAL REVIEWS

Annual Review of Animal Biosciences

A Beautiful Life: High Risk–High Payoff in Genetic Science

Stephen J. O'Brien[1,2]

[1]Theodosius Dobzhansky Center for Genome Bioinformatics, St. Petersburg State University, St. Petersburg, Russia 199004; email: lgdchief@gmail.com

[2]Guy Harvey Oceanographic Center, Halmos College of Natural Sciences and Oceanography, Nova Southeastern University, Fort Lauderdale, Florida 33004, USA

Annu. Rev. Anim. Biosci. 2020. 8:1–24

First published as a Review in Advance on November 19, 2019

The *Annual Review of Animal Biosciences* is online at animal.annualreviews.org

https://doi.org/10.1146/annurev-animal-021419-083944

Keywords

genomics, conservation, AIDS, autobiography, evolution

Abstract

This narrative is a personal view of adventures in genetic science and society that have blessed my life and career across five decades. The advances I enjoyed and the lessons I learned derive from educational training, substantial collaboration, and growing up in the genomics age. I parse the stories into six research disciplines my students, fellows, and colleagues have entered and, in some cases, made an important difference. The first is comparative genetics, where evolutionary inference is applied to genome organization, from building gene maps in the 1970s to building whole genome sequences today. The second area tracks the progression of molecular evolutionary advances and applications to resolve the hierarchical relationship among living species in the silence of prehistory. The third endeavor outlines the birth and maturation of genetic studies and application to species conservation. The fourth theme discusses how emerging viruses studied in a genomic sense opened our eyes to host–pathogen interaction and interdependence. The fifth research emphasis outlines the population genetic–based search and discovery of human restriction genes that influence the epidemiological outcome of abrupt outbreaks, notably HIV–AIDS and several cancers. Finally, the last arena explored illustrates how genetic individualization in human and animals has improved forensic evidence in capital crimes. Each discipline has intuitive and technological overlaps, and each has benefitted from the contribution of genetic and genomic principles I learned so long ago from *Drosophila*. The journey continues.

> It's a very ancient saying but a true and honest thought, that if you become a teacher, by your pupils you'll be taught.
>
> —Oscar Hammerstein II, *The King and I* (1951)

INTRODUCTION

Penning an autobiography is a singular challenge, mostly because the task is hugely subjective. Two anecdotes around this exercise help me explain my conundrum.

NBC TV journalist and 16-year host of *Meet the Press* Tim Russert recalled a walk with his dad—folks called him Big Russ—in Buffalo when Tim was 6 years old (1). They passed by a neighborhood church where parishioners were queuing up for a holy service. Big Russ told his son, "Tim, I want us to wait here at the church door for a bit and then, when the service is completed and the worshipers emerge, I have a lesson to tell you."

As the service concluded and the people made their way home, Russ said,

> That was some poor fellow's funeral. Did you notice how long the service was? It took approximately 35 minutes—35 minutes to honor his memory, to say nice things, and to highlight his lifetime accomplishments. When you and I pass on in the future, that's the time we will get, so be sure the worshipers have something important, relevant, or meaningful to say.

Your whole life in 35 minutes. Wow!

A second recollection, which raises the prospect that most biographies are apocryphal, is from a classic book written in 1982 entitled *In Search of Excellence: Lessons from America's Best-Run Companies* (2). The authors offered profiles of a handful of Fortune 500 company CEOs asking each for the secrets to their success in leading the world's most influential commercial institutions. The CEOs had detailed explanations and insightful reasons for their achievements.

But then the authors interviewed deputies and managers of the companies' executives and got their view of the same question: "What were the most definitive characteristics of the CEO's success?" The opinions from these subordinates were equally descriptive and credible, but in nearly all cases very distinct from the reasons the CEOs themselves had mentioned. Why? Because the staff had the perspective of standing back over a decades-long career as observers and offered what they believed really made the difference. The CEOs had more cryptic, personal, and recent perspectives, influenced by their latest actions, challenges, and movements. Who was correct? Maybe both? Can I be so objective myself? I can only try.

My own career was stimulating, complex, curious, and fascinating. As a geneticist, I am supposing that "nature" awarded to me some good useful aptitudes, but perhaps as important was the "nurture" component: the teachers, colleagues, students, extraordinary support, empirical opportunities, unanticipated discoveries, peer recognition, science culture, and a bit of good luck. To some, my history sounds simple, even routine: a PhD in *Drosophila* genetics; a postdoc at the National Cancer Institute of the National Institutes of Health (NCI-NIH); 40+ years of staying put; then capped off by a genome bioinformatics adventure in St. Petersburg, Russia. The editors of the *Annual Review of Animal Biosciences* have invited me to expand on my view of how it all came together. Because I was anything but focused on my research endeavors, my steps look chaotic and episodic. To help tell this narrative more clearly, I emphasize discrete genetic research disciplines that I became enamored with, for better or for worse, and tell the stories through a genomic lens.

ORIGINS

I was born in 1944 to Bernard and Kathryn O'Brien in Rochester, New York. Kathryn was a college-educated secretary who raised four children (I was number three), and Bernard was a

self-trained radio engineer. He began by building crystal radio sets in the late 1920s, and with but a high school diploma, he went on to build the CBS radio station and then the TV station in Rochester, where he became chief engineer. In my early youth, I paid little attention to science, being more interested in baseball, dogs, and rock and roll. My family relocated to Bethesda, Maryland, outside Washington, DC, in 1958, where I had a rather undistinguished high school performance that, to be frank, would limit my college admission opportunities. One thing I did enjoy in that period was singing and dancing in amateur high school musicals. Looking back, I suspect that these on-stage appearances provided terrific preparation in building confidence for public speaking later in the science arena.

After hitchhiking 12 miles each day to Good Counsel High School in Wheaton, Maryland, I was excited to enter St. Francis College, a small liberal arts Catholic school run by Franciscan friars in the hills of western Pennsylvania. St. Francis was good for me. It protected me from my adolescent instincts, prejudices, and lack of career focus. The teachers were patient but demanding. I majored in biology and chemistry in a premedical track. My original plan was veterinary medicine, as I loved dogs and animals. But my dad advised me that the veterinary course was both rigorous and demanding, perhaps as much as medicine, so maybe I should consider medicine, because the payoff both culturally and financially could be more rewarding. The premed track I entered was not a perfect fit, and it showed from my transcripts. And then my life took a big turn.

A young new professor at St. Francis, Dr. James Edwards, arrived to teach the emerging fields of genetics, evolutionary theory, and microbiology. Edwards lectured to small classes with extraordinary depth, insight, knowledge, and enthusiasm. His talks were sagas of the great discoveries in the beginnings of genetics and molecular biology. The characters were familiar: George Beadle, Barbara McClintock, Thomas Hunt Morgan, Calvin Bridges, Watson and Crick, Marshall Nirenberg, Jacob and Monod—the list goes on. I was hooked. This was not about memorizing endless anatomical details or Latin names of countless plants or insects. Biology came alive in the details of transcription, translation, development, operons, gene maps, transfer RNAs, messenger RNAs, and empirical science design. A few students like me embraced it, devoured the excitement, and immersed ourselves into the field.

I would remember a lesson from my father, when I asked him why he chose radio engineering as his profession that he did so well in. His response was, "Well I was pretty good at it, but also the great advances and discoveries that were driving things forward in radio electronics were all accomplished by young, innovative, and aggressive engineering pioneers. I really wanted to meet them all."

To me at that time, such a field was molecular biology and genetics. My academic performance skyrocketed. Edwards embraced and navigated my dream. As a determined mentor, he encouraged me to take summer courses in calculus and differential equations, then to elect a physical chemistry course in my senior year, top out in all of these, and apply to the best universities for graduate training in genetics. I did this and was accepted at half of the high-end schools to which I had applied. I chose a graduate fellowship at Cornell to work with Bruce Wallace, a well-known evolutionary geneticist and himself a student of the population genetics pioneer Theodosius Dobzhansky. I drove to Ithaca the day after my graduation from St. Francis in 1966.

COMPARATIVE GENOMICS: GENETIC MAPS TO WHOLE GENOME SEQUENCE

Bruce Wallace was a creative and kind teacher who penned textbooks in population genetics, designed provocative *Drosophila* experiments, and led social science education. More direct oversight

and training were provided by Ross MacIntyre, a zealous young molecular geneticist, who as a Cornell Assistant Professor set a high bar for groundbreaking and important molecular genetic research. My first project, an empirical attempt to quantify overall genomic diversity in *Drosophila* populations with allozyme (allelic isozyme) variants, was successful, but my findings came in the wake of a report by Dick Lewontin and Jack Hubby (University of Chicago), who estimated that something like 30% of tested allozyme genes had electrophoretic variation (3, 4). After a few more population genetic reports (5), I developed a strong interest in functional genetics for genes of the α-glycerophosphate cycle, a component of energy metabolism in *Drosophila* flight muscle. I learned how to induce knockout mutations that caused weakened flight and early senescence phenotypes in flies (6–9). This work would stimulate a postdoctoral appointment at the National Institute on Aging in Baltimore with Bertram Sacktor, a pioneer in insect metabolism and energetics.

All the while, Ross and I were continually mapping new gene markers until we had enough enzyme genes to put together a biochemical map for *Drosophila* (10). To me, gene mapping was a keystone to functional genetics; we desperately needed a catalog for the genes and mutational alterations that affected all things biological. Calvin Bridges and Alfred Sturtevant had initiated *Drosophila* genetic maps 50 years before using observable variant phenotypes (eye color, bristle patterns, and wing morphology). By the late 1960s, it was time to add molecular enzymes and proteins to these maps to connect genetics to the fast-moving and exciting new discipline of molecular biology.

The NIH was a happening place in the early 1970s. The War on Cancer had been declared by Nixon and embraced by NCI. The discoveries of mammalian retroviruses were blossoming—scores of oncogenes were being uncovered by rescues from retroviral isolates from leukemias and sarcomas in chickens, mice, primates, and cats. Because I wanted to be part of this excitement, I would shift my postdoc from Baltimore to NCI in Bethesda to join a red-hot program headed by George Todaro, a truly brilliant biomedical researcher who among other things had formulated the term oncogene (11). I was hoping that my genetics and gene mapping background would offer a useful perspective.

This hunch led to my developing somatic cell hybrid panels for mapping human genes. Frank Ruddle (Yale) and Tom Shows (Roswell Park Comprehensive Cancer Center in Buffalo, New York) had streamlined human mapping to an efficient process for mapping the first human molecular isozyme genes on human chromosomes. I learned from these two how to construct somatic cell hybrids and to definitively map human genes. Later ZooFISH (chromosome painting) techniques augmented the hybrid approach, along with the radiation hybrid (RH) mapping methods developed by David Cox (UCSF) and Peter Goodfellow (Cambridge). Applying these state-of-the-art mapping technologies, the small group I was building at NCI was invited into more than 30 exciting collaborative studies to map the newest oncogenes, retroviral integration sites, tumor suppressor genes, and other gene determinants of cancer. Bill Nash, Bill Modi, Joan Menninger, and I participated in scores of gene assignments for new genes in the armament of cancer research. We had the technology, NCI had the money, and the cancer molecular biologists were handing us new gene sequences of interest at a steady pace (12–17).

Early on in this period, my students and I decided to emphasize a research area that could become our own principal emphasis. I was intrigued by the retroviral advances in cancer research but felt an area that was lacking was uncovering variants in immune response genes, which regulate virus transmission and disease progression. I wanted to choose a nonhuman species with retroviral pathology to begin host genetic analyses. Four obvious models were considered: mice, cats, monkeys, and chickens. I excluded mice (too many great researchers there already), monkeys (too expensive and difficult to get numbers), and chickens (too evolutionarily distant from humans to be so relevant). This left domestic cats, which had advancing virology for feline leukemia

virus (FeLV); feline sarcoma virus (FeSV); feline herpesvirus; feline coronavirus; and later feline immunodeficiency virus (FIV), first cousin to HIV.

We constructed several panels of cat–rodent somatic cell hybrids and began mapping the cat genome (18). We built a linkage map with a breeding pedigree between house cats and a related species, leopard cats (*Prionailurus bengalensis*), developed at the NIH Animal Breeding facility in Poolesville, Maryland (19). This was exciting to me and sparked a career-long fascination (indeed, obsession) with the biology, evolutionary genetics, and medicine of cats (20–23). Our choice had one unexpected advantage: The amazingly prescient cancer and oncogene researchers (Todaro, Huebner, Aronson, Skolnick, Lowy, Klausner, Sherr, Varmus, Gallo, Vogelstein, and others) became my friends and were very willing to help train my students in their latest technologies. With a determined, perhaps misguided, interest in cats, we were not really considered as competitive to these pioneers.

Our publication and subsequent expansion of the cat gene map made an important point in the emerging field of comparative genomics. Cat genes were organized into syntenic groups that were remarkably similar to the order of homologous genes on human chromosomes. Cats have 19 chromosome pairs, and humans have 23 pairs. Six of the chromosomes bore the same gene clusters in both species, and the other chromosomes had very few interchromosomal exchanges (called translocations). By 1990, ZooFISH and hybrid maps showed that one could rearrange the human gene syntenic segments into the cat gene order with as few as 13 scissor cuts or exchanges (22–23). The same comparison of syntenic segments required ~275 scissor cuts to rearrange the human map to the mouse map. Today, whole genome sequence analyses reveal additional intrachromosomal inversions and segment transpositions, but the highly conserved ancestral genome organization demonstrated by cat and human genome mapping in the 1980s remains true.

I was invigorated by the comparative genetics discoveries in the 1980s and 1990s. We took on the task of comparing homologous gene orders and rearrangements among distantly related mammal species, all the while interpreting them in terms of the evolutionary and genomic determinants. This concept blossomed to become the Comparative Genomics discipline, which details the similarities in gene order and the fixed evolutionary rearrangements that preceded the origins of modern species (23).

To promulgate this area, we needed precise and detailed gene maps of representative species of mammalian orders and families. As the human genome began to take shape, the importance of comparative data became crystal clear. In the 1980s, human gene mapping workshops kept up with the rapidly developing maps of each human chromosome. Indeed, the workshops had committees dedicated to human chromosome 1, to human chromosome 2, to human chromosome 3, to human chromosome 4…all the way to chromosome 23. Almost as an afterthought, our Comparative Committee was added to the group. The lead workers on gene mapping for mouse (Peter Lalley and Muriel Davisson), cow (Jim Womack and Harris Lewin), horse (Ernie Bailey), pig (Larry Shook and Alan Archibald), kangaroo (Jenny Graves), and cat (*moi*) agreed to build dense gene maps for each species and to compare the gene orders to impute translocation, inversion, and transposition rearrangements that punctuated genome evolution in mammals. The comparative community mapped genes, aligned the gene maps to different species, and included better gene mapping technologies [restriction fragment length polymorphism, polymerase chain reaction (PCR), DNA markers of various sorts].

In 1997, a group of us got together in Cape Hatteras, North Carolina, and designed a series of highly conserved PCR primers across the extensive human, mouse, and rat gene sequences, hoping to successfully amplify the gene homologs of cats, dogs, horses, pigs, and other nontraditional mammals. We called the markers comparative animal tagged sequence (or CATS) primers (24, 25). The original idea, which others would follow, was to provide a common platform for gene mapping

in any placental mammal using the same primers tagged to human–mouse gene homologs. It was an audacious step, but as Chinese philosopher Lao Tzu once mused, "A journey of a thousand miles begins with a single step." The comparative gene mapping committees working in parallel to map homologous genes in a dozen species were our first step.

Species gene maps were compiled and published annually by Cold Spring Harbor Press in what became six editions of *Genetic Maps: Locus Maps of Complex Genomes* from 1980 to 1993 (26). The first edition in 1980 was 28 pages; the last was a massive tome of 1,617 pages. When the data grew too large for anyone to carry, *Genetic Maps* was displaced by GeneBank, the NIH National Center for Biotechnology Information online database that went on to archive published, unabridged DNA sequences deposited from all biological species.

Around that time, whole genome sequences of model organisms were appearing (*Escherichia coli*, *Saccharomyces*, *Caenorhabditis*, *Drosophila*, and others), thanks to first-generation Sanger sequencing (developed by PerkinElmer–Applied BioSystems), 454 pyro-sequencing, and others, and subsequently the cheaper, faster, and robust second-generation gene sequencing offered by, e.g., Illumina, Pacific BioSystems, BGI Genomics, and Oxford Nanopore. An initial full human genome sequence appeared in 2001 at an estimated cumulative cost of $2.7 billion, but questions about its meaning, organization, and interpretation were numerous and seemed to need an evolutionary perspective provided by comparative genome sequencing of related species. The NIH National Human Genome Research Institute staffed a committee (including myself) to select and fund light Sanger sequence of 32 diverse species of mammals for a lesser price of $2–3 million per species (27). Cat and dog were among those selected for sequencing, leading to a major thrust in the area of veterinary medical models of these two companion animals (28, 29).

As the cost of genome sequencing was rapidly dropping, David Haussler (UC Santa Cruz), Oliver Ryder (San Diego Zoo), and I joined forces in 2009 to establish the Genome 10K Project (G10K), a consortium of geneticists, bioinformaticians, and organismal zoologists committed to facilitating whole genome sequence of approximately 10,000 species of vertebrates (30–32). We envisioned a truly transformative gift to the coming generation of biological researchers: a public whole genome sequence for each of the 63,000 named vertebrate species for which we could actually gather specimens. G10K is thriving as I write these words today. In the decade since its inception, its vision, energy, and commitment have spawned other consortia for genome sequence development for insects, plants, marine invertebrates, all vertebrates, fungi, and others (32–34). Coordination of all these consortia was attempted last year by the Earth BioGenome Project (EBP) led by Harris Lewin and colleagues (35). EBP proposes to coordinate whole genome sequence of all eukaryote species alive today, some 2 million species, at a projected cost of $4.7 billion. For the comparative genomics community, it is an ambitious but plausible goal. Soon, I expect we will really see genome sequence empowerment for countless species of biological interest and enquiry.

In 2015, conservation genomics and functional genomics made a major leap forward when the avian consortium of G10K, led by Erich Jarvis, Tom Gilbert, and Guojie Zhang, coordinated the simultaneous production of 50 related papers, many in a dedicated issue of *Science* (36, 37), that together analyzed the functional and evolutionary inferences derived from annotated whole genome sequences of 50 bird species. These papers dissected the genetic components of bird biology determining feather development, wing aerodynamics, bone density, olfactory reception, visual opsins, tooth loss, plumage coloration, and many other features. Today whole genome sequences have been accomplished for more than 300 vertebrate species, and my own group at the Theodosius Dobzhansky Center for Genome Bioinformatics in St. Petersburg has contributed to the annotation and analyses of many of these (32, 36–47). Comparative genomics approaches to whole genome sequences are now being used to impute the genome rearrangements that punctuated the mammalian radiations (48). I saw the genome revolution come of age.

EVOLUTION AND SPECIES NATURAL HISTORY

In my early days, I was strongly influenced by the writing, teachings, and personal counseling of Theodosius Dobzhansky, founder of the Modern Synthesis, a period in the 1940s and 1950s during which genetic advances and evolutional perspectives were joined to enrich both previously warring fields. Dobzhansky (49), considered by many the father of empirical population genetics, is well known for his quotation, "Nothing in biology makes sense except in the light of evolution."

No longer was evolution about a creationist debate. Rather, it became a keystone for biological inference and interpretation of functional genetics and development. In 2012, I named the center I established in St. Petersburg the Theodosius Dobzhansky Center for Genome Bioinformatics in his honor.

In the early 1960s, the powerful discipline of molecular evolution was jump-started by the molecular clock hypothesis first formulated by Linus Pauling and Emile Zuckerkandl (50). Since then, conceptual advances in molecular evolution have produced amazing insights and windows into the ancient evolutionary secrets of living species. Developed analytical algorithms of evolutionary theory and automated computer scripts allowed the construction of statistically robust phylogenies of living species. Many workers, including those in my own group, rapidly learned and applied four diverse but complementary philosophies to reconstruct the hierarchical history of species divergence and emergence. The methods had names: (*a*) numerical taxonomy-phenetic or distance-based methods; (*b*) cladistics or maximum parsimony methods; (*c*) maximum likelihood statistical approaches; and (*d*) Bayesian statistical methods. Today, evolutionary relationships of nearly all zoological groups have been solved with these tools and approaches (51).

In the comparative genomics area, we could not ignore the evolutionary lessons and more important opportunities that were arising. During the 1990s, Senior Editor Barbara Jasny at *Science* produced an annual Genome Issue featuring the best genome advances of model organisms and species. Intrigued by the depth of specific genome reorganization that the comparative mammal community was revealing, she asked me to help organize a *Science* issue on Comparative Genomics to illustrate genome exchanges that had been documented among approximately 20 mammal species with gene maps and ZooFISH (23). My comparative colleagues and I put together a large wall chart that displayed the chromosomes of 30 mammals colored according to their homologous syntenic segments with the human genome (52). At the top of the poster was a consensus phylogenetic diagram to show the generally accepted phylogenetic or evolutionary hierarchy of the 18 recognized orders of placental mammals, compiled by a consensus of mammal molecular researchers and paleontological systematists. The colorful genome poster, summarizing the comparative inference to date, adorned the lab refrigerator doors of gene researchers across the world.

But there was a catch. Our assessment of the consensus phylogenetic tree of mammals was a major stretch, made impossible because the morphologists and the molecular people could not agree at all on the hierarchy or the timing. At this stage, the research team that I directed, NCI-LGD (NCI's Laboratory of Genomic Diversity), spearheaded by Bill Murphy, Eduardo Eizirik, and Al Roca, reasoned that if we could use CATS primers to map genes, maybe we could design specific CATS primers to amplify genomic sequences of available species across mammals to resolve the mammal phylogeny using the molecular clock. Of the more than 100 CATS primers they designed, ~15 worked in 90% of the test species. Sadly, the rest failed owing to DNA drift across more than 100 million years of placental mammal evolution, but still we had enough to produce a first phylogenetic sampling and analysis across placental mammals (53). The dating of the mammal tree divergence nodes was calibrated with 15 time-calibrated ancient mammal species fossil remains that experts believe were precursors, or missing links, for most of the modern species

groups. The fossil date–calibrated mammal phylogeny gave a precise timing and hierarchy for the origins and family-level splits that precede the living mammals of today (53).

Nearly every really good idea in science occurs to multiple researchers simultaneously. Mark Springer's group (UC Riverside) had the same thought for mammal evolutionary history, and in 2001 our two groups published independent verification of the origins and divergence of mammal species in *Nature* (53, 54). Then the Springer and O'Brien teams quickly combined the data sets from each group and published collaborative affirmation of the major divergence node across the 100-million-year-old radiation of placental mammals (55). That relationship has been improved and tweaked but stands today as pretty much the same natural history discerned in 2001.

Molecular phylogenetic scenarios by my students and fellows would follow the mammal tree CATS primer strategy with some notable successes. The cat family Felidae is composed of 37 living species that range throughout the world, excluding the Arctic, the Antarctic, and Australia. Although catlike species occurred as early as 40 Mya in the fossil record, it was consistently agreed that modern living cat species all descended from an ancestor, *Pseudaelurus*, that lived in Asia ~10.8 Mya. The molecular phylogenetic analyses of modern Felidae, led by Warren Johnson, using time calibrated by 12 well-dated cat fossil precursors, affirmed the origins of 8 principal groups (we proposed most as genera) (56). The combined analyses of phylogenetic hierarchy, the present distribution of today's cat species, the geological record of sea level rise and fall, and the fossil record allowed us to propose not only a history of species divergence but also a plausible hypothesis for historic intercontinental migrations that preceded the present distribution of the 37 Felidae species (57).

Similar applications of phylogenetic reconstruction and interpretation are now widespread and have been summarized recently in the *Timetree of Life* series (51) and on websites based upon molecular evolutionary phylogeny reconstruction by a generation of molecular evolutionists (**http://www.timetree.org**). I had first become fascinated with the power of molecular evolutionary analyses when we embarked on solving the century-old mystery of the phylogenetic position of the giant panda relative to bears and the red panda, which looks a bit like a raccoon (58, 59). This was a fascinating study with a colorful history that I summarized in some detail in my semipopular book, *Tears of the Cheetah* (60). Our research groups would also contribute to phylogenies of the bat family Chiroptera (led by Emma Teeling and leading to the Bat1K genome sequencing project), primates (led by Polina Perelman and Jill Pecon-Slattery), pangolins (led by Shu-Jin Luo and Agostinho Antunes), and Carnivora (led by Robert Wayne) (61–65).

An important outcome of the phylogenetic and population studies of nontraditional species was the serendipitous discovery of several new species. In hindsight, the group I led would describe novel, previously unrecognized species of orangutan, clouded leopard, African wolf, South American tigrina, and elephant (66–69). Each new species was validated by multiple molecular genetic criteria, and nearly all have been accepted by the international conventions for species recognition. I was struck by a news headline when our group (led by Al Roca) described a second definitive species of African elephant, *Loxodonta cyclotis*, living in the African Congo. The headline read, "New species of elephant discovered in Africa—you think it would be hard to miss!" My joy continues.

GENETICS OF ENDANGERED SPECIES

Perhaps one of the more notable areas of my personal science adventure was the realization that genetic analyses of free-ranging wildlife species can reveal secrets of their past and projections for their future (20, 60). This first occurred when David Wildt and Mitch Bush (Smithsonian Institution) invited me into a study of African cheetahs in the early 1980s. Cheetahs were well

known as the world's fastest land animal, a beauty to observe sprinting across the African savannah. Unlike other cat species propagated by world zoos, cheetahs were very slow to breed and, when they did, produced congenital abnormalities at a relatively high frequency. Bush, a clinician, and Wildt, a reproduction specialist, wanted to explore the biological bases for the cheetahs' poor reproductive performance. Wildt found that cheetahs had a high incidence of morphologically malformed spermatozoa (~70% abnormal compared with <30% abnormal in cats, human, and dogs). In a collection trip to a cheetah breeding center in Pretoria, South Africa, they retrieved blood samples from 60 cheetahs for allozyme diversity quantification. We screened 52 markers and found zero variation in this initial sample, the first evidence of genetic impoverishment for cheetahs (70). Jan Martenson, a talented technical assistant who ran the allozyme tests, quipped, "These cheetahs are so boring and monotonous. For a moment, I wondered if Wildt and Bush really collected 50 different cheetahs?" The results were that hard to believe. Every cheetah was genetically identical to all the others.

The clincher to the stunning result came when we surgically exchanged postage stamp–sized pieces of skin graft between 14 unrelated cheetahs. The grafts were all accepted as if they were identical twins (71). Subsequently, we added additional measures that all pointed to the conclusion that the cheetah as a species displays 90–99% less overall variation than other cats or mammals in general (72). The latest affirmation of the cheetah's genetic impoverishment came from its whole genome sequence (38). The cheetah would become the poster-child species for the perils of severe population reduction, mating with close relatives, and shriveling genetic diversity, a prelude for extinction (60, 72).

Studies of other species soon revealed that they too had reduced variation because of a near-extinction event in their past. The Florida panther, a southeastern US subspecies of cougar or mountain lion, also showed depleted genic variation and consequent congenital pathologies (95% spermatozoa pleiomorphism, cryptorchidism, cardiac defects, and compromised immune system) (73). The evidence for genetic diminution, combined with clinical details assembled by Melody Roelke, DVM, led to a Florida panther conservation management initiative to augment the >30 Florida panther survivors with 8 healthy wild-caught female Texas cougars, a subspecies that had gene flow with the Florida subspecies 150 years ago. The restoration program was a dramatic success, resulting in rapid recovery of population numbers, tripling of population size and density, and a measurable increase in fitness for the subspecies intercross offspring (74). Similar genetic reduction and restoration proposals have been offered, but never implemented, for Asian lions, Amur leopards, Amur tigers from the Russian Far East, and California Channel Island foxes (75–78).

Genetic diversity is not the only mammal conservation issue genetics has weighed in on (20). New species discoveries, as outlined above, plus genetic-based categorizations of subspecies also informed wildlife conservation units. For example, the definition of explicit tiger subspecies has confirmed the historically separated units of tiger conservation, as well as the postulated founder effect for today's tigers caused by the Toba volcanic explosion in Southeast Asia ~73,000 years ago (78, 79). Subspecies verification also happened with leopards, pumas, tigers, and lions (77–81). Population genetic and coalescent dating approaches allowed us to pinpoint the timing of historic demographic events like the cheetah and Florida panther bottlenecks (12,000 years ago around the end of the last ice age for cheetahs and a few hundred years ago for the Florida panther owing to deliberate decimation) (60, 71–74, 80).

Ecologists have embraced genetic parentage and kinship estimators to resolve field behaviors in lions and other species. For lions, the development of the pride social unit had an evolutionary explanation that involved conscious mate choices. Oxford zoologist William Hamilton had coined the term kin selection to describe a component of natural selection impacted by close

relatives other than parents (82). Put simply, if evolution is all about transmitting one's genes successfully, would there not be an advantage to helping your brothers and sisters transmit theirs? Dennis Gilbert's comprehensive genetic assessment of Serengeti lion pride organization solved this conundrum where for lions everything sexual is a family event, according to lion ecologist Craig Packer (University of Minnesota) (83, 84).

Whole genome sequences are further pinpointing the genes that mediated historic events of evolutionary adaptation. Pavel Dobrynin's PhD genomics analysis showed that cheetah sperm problems were linked to five damaging knockout mutations in their *AKAP4* gene, which plays a major role in human spermatogenesis, azospermia, oligospermia, gonadal dysfunction, and oogenesis (38). Another fascinating example involves the well-known artificial selection experiment for tame behavior involving silver foxes, designed in 1959 by Dmitry Belyaev in Novosibirsk, Siberia. Anna Kukekova (University of Illinois) recently reported a genome analysis indicating that the docile behavior in the foxes was probably a result of mutations in the very large gene *SorCS1* (85). Today few management studies or planning workshops about threatened species convene without consideration of the genetic and genomic history of that species.

PLAGUES AMONG US

It occurred to me very early in my career that gene interaction and environment were critically important considerations in understanding the demography and survival of every living species. One of the strongest and most influential of environmental components had to be the role of pathogens, notably viruses, bacteria, helminths, and parasites, that afflicted individuals, populations, and species. In nature, the delicate balance between hosts and their pathogens is akin to a deadly arms race, waged fiercely almost daily and multiplied over individuals, within populations, across geography, and among the more than 5,500 mammal species that survive on earth today. The struggle is frequently lost, never won, but temporarily circled with the survivors rising once again to compete another day.

In the 1960s, there were public voices, encouraged by the success of antibiotics and vaccines against smallpox, measles, mumps, and polio, who predicted a swift shift in biomedical emphasis from infections to chronic diseases (86). But then came AIDS, hepatitis B, hepatitis C, papillomavirus, hantavirus, Ebola, SARS, West Nile virus, Zika, Legionnaires' disease, mad cow disease, etc. Modern medicine had far from conquered infections. Not even close.

Transmissible FeLV and FeSV in house cats became my first exposure to the power of the agents to rapidly overcome a healthy population (87). FeLV was supposed to afflict only domestic cats, but that conception was shattered when we encountered a deadly FeLV outbreak in the fragile Florida panther population during its recovery (88). Today we believe that pathogen outbreaks often cause the last step in countless populations' and species' extinction events.

When the HIV–AIDS epidemic began in the early 1980s, we never anticipated that there would soon emerge a closely related lentivirus in domestic cats. FIV, first cousin of the deadly HIV, mimicked human AIDS pathogenesis in infected cats, with a devastating prevalence of 5–15% of the hundreds of millions of feral cats worldwide infected with it (89). Our worries that the deadly FIV may have moved into wild cats were confirmed by a huge epidemiological sero-survey of the 37 Felidae species that revealed up to 20 species endemic with FIV (90–92). But for some inexplicable reason, the FIV-infected wild cats did not seem to be getting sick.

Then, in 1994, a devastating contagion abruptly killed 33% of the FIV-infected lions in the Serengeti. At first, we were afraid that the endemic FIV had caused the death of these lions, but we were wrong (93). Some in-depth histological and molecular analyses of lion necropsy materials revealed that the agent in the Serengeti lions was a variant of canine distemper virus (CDV), a

morbillivirus (think measles) that was previously thought to cause disease in dogs but never in cats (93, 94). Not this time.

The indigenous Maasai tribesmen who herd cattle outside the Serengeti had pet dogs that were themselves endemic with CDV, a likely source of the lion agent. Dogs are prohibited from the Serengeti National Park, so how did the lions contact them? We now think the connection involved marauding Serengeti hyena packs that also developed CDV titers at around the same time. Hyenas, who rumble with both Maasai dogs and lions, became our best guess as the carriers of the CDV from Maasai dogs to Serengeti lions. In truth, Packer and Roelke's vigilant surveillance of the lions and dogs opened our eyes to a close-up view of a horrific fatal outbreak in a large predatory carnivore (93, 94). Quickly thereafter, a vaccine initiative targeting the Maasai dogs ensued, which seems to have prevented further outbreaks over the coming decades.

A lesson of host gene influence during viral outbreaks became crystal clear when the SARS (severe acute respiratory syndrome) epidemic emerged in November 2002. The flulike disease defied any and all treatments as it abruptly appeared in the emergency rooms of Chinese hospitals around Guangdong, Guangxi, and Hong Kong (95, 96). SARS was new to medicine, to science, and to people. In 9 months, SARS traveled to 29 countries, infected 8,098 people, and caused 774 deaths (near 10% mortality). The alarming speed of spread and virulence of transmission were dramatic. Consider that a single patient admission at the Prince of Wales Hospital in Hong Kong transmitted to 112 cases within a few hours. A nearby sprawling Hong Kong apartment complex, Amory Gardens, home to 10,000 residents, developed 329 SARS cases after an infected woman coughed into the ventilation system. The condo became a ghost town within a few days.

The SARS epidemic was quickly shown to be caused by a novel coronavirus related to that which causes one-third of common colds. Although SARS subsided in May 2003, likely as a consequence of draconian quarantine measures, the deadly epidemic scared us all. Sixteen years later, we still have only a vague understanding of the precise mode of transmission, no laboratory-based clinical diagnosis, and no vaccine or efficacious treatment for SARS. The abrupt and deadly disease had spread at lighting speed, leaving dire economic consequences in the billions of dollars, along with a dramatic Asian cultural shift toward avoiding handshakes, kisses, or any physical contact (95, 96).

Coronavirus pathologies at the time were actually more familiar to veterinary clinicians, who had seen coronavirus diseases in mice, turkeys, chickens, pigs, dogs, and cats (97). Although I had been aware of these, the SARS episode set off an alarm as I recalled a devastating coronavirus epidemic we encountered in the 1980s in cheetahs (71). At the same wildlife park in Oregon where we had exchanged skin grafts between unrelated cheetahs, a pair of imported cheetahs from Sacramento arrived and grew ill with fevers, jaundice, and acute tremors. With little but symptomatic treatment, the two cheetahs deteriorated and succumbed to an unknown contagion. Biopsies and viral screening revealed they were infected with a domestic cat coronavirus called feline infectious peritonitis virus (FIPV). In cats, FIPV induces an accumulation of proteinaceous immune complexes in the peritoneum that strangulates the organs, leading to death. The FIPV quickly spread through fecal transmission across the park, and within a few months all 60 cheetahs had seroconverted (71). By the time of the SARS outbreak (in 2003), newly available molecular tools allowed us to recover FIPV sequence from archival specimens to trace their origins. The cheetah virus was nearly indistinguishable from FIPV isolated from domestic cats (98). Why is that important?

The cheetahs' FIP morbidity (symptoms) was 100% (including chronic diarrhea, jaundice, and tremors), and 60% of the infected cheetahs died of the disease within three years. In domestic cats, FIPV morbidity is usually $<10\%$, and mortality $\sim 1\%$. In human SARS, the mortality was $<10\%$. A near-uniform response of cheetahs to the FIPV was no coincidence. The FIPV, once it had

overcome the defenses of its first cheetah victim, had overcome them all. Cheetahs were all genetically the same (remember the skin graft results). The cheetahs' smoking gun was their endemic genetic uniformity, including the immune defense genes (98). This homogeneous clinical response to FIP in cheetahs convinced me and my colleagues that the genetic context of immune defenses had an enormous influence on the outcome of a novel emerging virus in a naïve population. This lesson I would never forget.

AIDS-RESTRICTION GENES AND THE ERA OF GENOME-WIDE ASSOCIATION STUDIES

A clear potential of the comparative inferences unfolding in my early scientific career was the hope to translate animal insights to human medicine. Our research group at NCI began a remarkable journey in the early 1980s when we looked for a chronic infectious disease that likely had host genetic components contributing to epidemiologic outcomes. We chose HIV–AIDS as our target because the disease first emerged in 1981, clearly involved an infectious agent (HIV), and was deadly to 95% of its victims. We were hoping that the natural genetic variants, once unveiled, might lead to better diagnosis, to effective drugs, or even to prophylactic agents.

After first consulting with human geneticists and AIDS experts and securing NCI financial support, we mounted a search for what we called AIDS-restriction genes, host genetic variants that influence epidemiologic heterogeneity among HIV-exposed individuals. We entered into collaborations with diligent epidemiologists, who were studying principal AIDS risk groups (actively gay men, hemophiliacs who received contaminated clotting factor before the advent of HIV screening in 1984, and IV drug users who were sharing contaminated hypodermic needles in urban settings). We sought to connect the clinical data documented by the epidemiology communities with DNA variants across the human genome as they were described by human geneticists and the fledgling human genome projects.

To do this, we gathered blood samples from the at-risk cohort volunteers. Our strategy was to genotype large numbers of individuals in the AIDS cohort studies with DNA variants in candidate genes (those genes implicated as important in AIDS pathogenesis) for HIV infection, AIDS progression, AIDS-defining diseases (Kaposi's sarcoma, lymphoma, *Pneumocystis carinii* pneumonia), other clinical indicators of AIDS, and differential responses to available therapy (after 1996 termed HAART for highly active antiretroviral therapy). For detection, patients were parsed into groups with different disease outcomes so we could compare allele frequencies of variants between groups, as well as conformance to Hardy–Weinberg equilibria and other population genetic equilibria measures. When a certain gene allele showed higher frequency in one group versus another (e.g., HIV negative versus HIV positive), it was the first signal that the allele conferred resistance to HIV infection.

Between 1985 and 1996, we accumulated more and more patients from some 25 cohorts who agreed to be part of our ambitious study consortium. We recruited nearly 20,000 patients, a rich repository for the discipline of genetic epidemiology. At NCI-LGD, Cheryl Winkler and Mary Eichelberger built a cell transformation laboratory that produced immortal B-cell lines using Epstein–Barr virus transformation protocol from most of the blood samples. Raleigh Boaze extracted high molecular weight from each sample. I had hired Michael Dean, a superb young molecular biologist, to develop the human genotyping technology. Mike carried out the DNA genotyping for hundreds of gene candidates (Mary Carrington led the *HLA* and *KIR* typing effort), all in search of a statistically significant population genetic signal for AIDS influence.

For more than a decade, the clinic doctors and nurses, our epidemiologic collaborators, Cheryl Winkler, Mike Dean, Mary Carrington, and I continued to add more patients, more genes, more

genetic variants, and more sophisticated computer programs to search for AIDS-restriction genes. By the mid-1990s, we had screened thousands of patients for hundreds of gene variants distributed across the human chromosomes. Every so often, we spotted a genetic difference between groups, but they all evaporated under closer inspection. We would monitor new research advances that were reported in the AIDS literature, searching for new genes to test. Finally, 12 years after we had begun what was becoming a tedious, expensive, and thus far disappointing fishing expedition, there appeared a glimmer of hope (99–101).

In July 1996, four separate studies from independent groups (not us) reported in *Nature*, *Cell*, and *Science* a requirement for the CCR5 chemokine receptor molecule for HIV to enter lymphoid cells (102–105). HIV binds to two receptors on T-lymphoid cells, CCR5 and CD4, in a stepwise process to enter these cells and destroy them. In an infected victim, HIV then produces more than a billion viral particles each day. CD4-bearing T lymphocytes eventually become depleted, leading to collapse of cell-mediated immunity, thereby allowing normally innocuous infections and certain rare cancers to proliferate and kill the victim.

We designed PCR primers to amplify the *CD4* and *CCR5* human genes and search for gene variants the day after the announcements appeared. Our team quickly discovered a large (32–nucleotide pair long) deletion in the *CCR5* gene (*CCR5-Δ32*) in some but not all patients. When we genotyped the AIDS cohorts, an incredible result appeared. Although the *CCR5-+/+* and *CCR5-+/−* genotypes were found in both HIV-positive and HIV-negative patients, the HIV-negative patients were never homozygous for *CCR5-Δ32/Δ32*. It seemed the people with *CCR5-Δ32/Δ32* were completely resistant to HIV infection, no matter how many times they were exposed. The reason was that the requisite CCR5 receptor—the doorway by which HIV enters T lymphocytes—was slammed shut. Other groups would opine that *CCR5-Δ32* might confer resistance to HIV (106, 107), but our result (99), published in *Science* on September 27, 1996, proved it with the genotypes of 1,995 people from 6 separate AIDS cohorts. Later we would extend and expand these studies around the details, geographic distribution, and origins of *CCR5-Δ32* in scores of derivative reports (108–110).

CCR5-Δ32 remains today one of the more wide-reaching success stories of human gene association. Several anti-AIDS drugs that block CCR5–HIV binding (fuszon-enfuvirtide and maraviroc) were developed and approved as AIDS treatment by the Food and Drug Administration (110). We noticed that *CCR5-Δ32* mutation is found mainly in Europeans (but virtually absent in pure East Asian and African ethnicities), indicating it first occurred in Europe after the "Out of Africa" migration and was likely raised to a 10% allele frequency by the sixteenth-century Black Death and earlier bubonic plague episodes. Both the Berlin patient and London patient, the only two individuals to have had HIV cleared or cured (of the 37 million HIV carriers alive today), received a stem cell transplant using *HLA*-match donors who were homozygous for *CCR5-Δ32/Δ32* (111–113). There are provocative reports that CCR5 action is important in graft-versus-host disease (GVHD), which can kill up to one-third of bone marrow transplant recipients in cancer therapy (114, 115). Phase I and II trials using maraviroc (a CCR5 antagonist) treatment after bone marrow transplants showed remarkable ablation of GVHD incidence without side effects. *CCR5-Δ32* has been shown to mediate more rapid recovery from stroke (116). Alternatively, homozygous *CCR5-Δ32/Δ32* carriers are at a fivefold-increased risk for developing encephalitis after infection with West Nile virus (117). Finally, Chinese scientist Jiankui He's ethically challenged CRISPR-based germline editing claim this year targeted *CCR5* to protect two newborn babies from HIV infection (118, 119).

The experience and follow-up with CCR5 invigorated our team to use the AIDS cohorts to discover some 36 additional AIDS-restriction genes (110) (**Table 1**). Each was published, and most were replicated in independent cohort studies. However, because nearly all showed a

Table 1 AIDS-restriction genes described by the NCI's Laboratory of Genomic Diversity (1996–2012)[a]

No.	Year	Gene	Allele	Mode	Effect
1	1996	*CCR5*	*Δ32*	Recessive	Prevents infection
2	1996	*CCR5*	*Δ32*	Dominant	Delays AIDS
3	1997	*CCR2*	*64I*	Dominant	Delays AIDS
4	1998	*CCR5P*	*P1*	Recessive	Accelerates AIDS
5	1998	*SDF1*	*3'A*	Recessive	Delays AIDS
6	1999	*CCR5*	*Δ32*	Dominant	Prevents lymphoma
7	1999	*HLA*	*A,B,C,"Homozy"*	Codominant	Accelerates AIDS
8	2000	*IL10*	*5'A*	Dominant	Limits infection
		IL10	*5'A*	Dominant	Accelerates AIDS
9	2001	*HLA*	*B*35Px*	Codominant	Accelerates AIDS
10	2002	*RANTES*	*−403A*	Dominant	Accelerates AIDS
			In1.1C	Codominant	Accelerates AIDS
11	2002	*KIR*	*3DS1*	Epistatic (Bw4–801)	Delays AIDS
12	2003	*EOTAXIN-MCP1*	*Hap7*	Dominant	Enhances infection
13	2003	*HLA*	*B*57*	Codominant	Delays AIDS
14	2003	*IFNG*	*179T*	Dominant	Accelerates AIDS
15	2003	*CXCR6*	*E3K*	Dominant	Accelerates PCP
16	2004	*APOBEC3G*	*H186R*	Recessive	Accelerates AIDS
17	2004	*DCSIGN*	*−336T*	Dominant	Decreases infection
18	2006	*HLA*	*B27*	Codominant	Delays AIDS
19	2006	*TSG101*	*Hap2*	Dominant	Accelerates AIDS
20	2006	*TRIM5*	*Hap4*	Dominant	Increases infection
21	2007	*Cul5*	*HapI*	Codominant	Accelerates CD4 loss
22	2007	*PP1A* (cyclophilinA)	*SNP-4*	Dominant	Accelerates AIDS
23	2007	*HLA*	*Bw4*	Dominant	Reduces HIV transmission
24	2008	*MYH9*			End-stage renal disease
25	2008	*MYH9*			HIV FSGN
26	2008	*mtDNA*	*Hap-J, U5a*	Dominant	Accelerates AIDS
27	2008	*mtDNA*	*Hap-H*	Dominant	Increases lipatrophy post HAART
28	2008	*mtDNA*	*Hap-J*	Dominant	Delays CMV-NRD
29	2009	*HCP5*	*T>G; rs2395029*	Dominant	HIV set point
30	2009	*HLA*	*rs9264942*	Dominant	HIV set point
31	2009	*PROX1*	*Hap-CGT*	Recessive	Delays AIDS progression
32	2009	*APOBEC 3B*	*ΔV_*	Recessive	Increases infection
33	2010	*PECI*	*G*	Dominant	Accelerates AIDS
34	2010	*ACSM4*	*A*	Codominant	Delays AIDS
35	2010	*NCOR2*	*T*	Dominant	Increases infection
36	2010	*IDH1*	*a. C*	Dominant	Prevents infection
37	2011	*PARD-3b*	*b. C*	Codominant	Delays AIDS

[a]All these genes were discovered or validated using the NCI's Laboratory of Genomic Diversity. Eight AIDS cohorts including more than 10,000 study participants were used in these studies from 1985 to 2012 (see 110 for citations).

Abbreviations: AIDS, acquired immunodeficiency syndrome; CMV-NRD, cytomegalovirus neuroretinal disorder; FSGN, focal segmental glomerulosclerosis; HAART, highly active antiretroviral therapy; HIV, human immunodeficiency virus; NCI, National Cancer Institute.

quantitative allelic difference (unlike the gene-knockout effect of *CCR5-Δ32*), clinical follow-up and translation have been less aggressive than for *CCR5*. Finally, we have used similar gene association studies, including genome-wide association studies, to discover multiple restriction genes for HIV–AIDS in Africa (120), for hepatitis B and C in China (121, 122), and for nasopharyngeal carcinoma in China (123). These advances joined in complex gene discovery during the genomics era make me quite proud of the NCI-LGD team, who really were among the early formulators of a human "genetic epidemiology" discipline.

FORAYS IN FORENSIC SCIENCE

My earliest tiptoe into forensic genetics began with the work of Walter Nelson-Rees, a forensic pioneer who uncovered cell culture mix-ups or contaminations with HeLa cells in the 1970s. Walter would request samples of well-known cell lines studied by major cancer research groups and then resolve HeLa karyotypic markers, visible chromosome translocations, and rearrangements specific for the rapidly growing HeLa cells (made famous recently by the book and movie about the unwitting donor, Henrietta Lacks; 124). Nelson-Rees would catalog and publish all the cases in which he received a cell line with a certain name and origin label that really turned out to be HeLa cell contaminants (125, 126). Often his public announcement of cell-line misidentification ruffled the feathers of the world's cell biologists, who expected respectful confidentiality.

At the time, we were researching human allozymes and their population frequencies. Nelson-Rees urged me to develop a panel of allozyme markers that would individualize human cell identification and also a cell line's species origins. We did this quickly, contributing to several releases of cell-line screwups, as well as developing statistical rigor associated with individual genetic uniformity (127–129). This was a prelude to the better-known DNA fingerprinting and product rule calculations used today in courtrooms across the globe.

In the fall of 1994, a year after the televised O.J. Simpson murder trial taught laymen the power of DNA fingerprinting for individual identification in capital crimes, I was contacted by Constable Roger Savoie of the Royal Canadian Mounted Police to help solve the murder case of a young mother on Prince Edward Island (130, 131). The Mounties had uncovered a leather jacket in the woods near the crime scene that was spattered with the victim's blood, identified by DNA fingerprinting. Nested in the jacket lining were whorls of white hairs that were not human but feline. The prime suspect lived with his parents and a large white tomcat named Snowball. Because of our reputation as cat genetic experts, Savoie asked me to try to determine whether the jacket hairs were from Snowball, which would connect the suspect to the crime scene.

Marilyn Raymond and Victor David, crack geneticists on the NCI-LGD feline research team, took on the project and performed the DNA analyses that in the end allowed for a conviction and sentencing. The jacket hair DNA showed a perfect match to Snowball. We then had to build an ad hoc population genotype database of 19 stray cats from Prince Edward Island. The Island cats had abundant variation and allowed us to compute the likelihood of a chance match of Snowball to another local cat to a vanishingly tiny probability of 1 in 45,000,000. The suspect was convicted and sentenced, setting an international precedent for animal genetic individualization in capital crime cases. Constable Savoie, who recruited us and oversaw the case against the suspect, was awarded the high honor of Mountie of the Year in 1997 for his persistence. Details of the Snowball case are described in more depth in Reference 60 and in a derivative TV episode of *Forensic Files* (132).

Our connecting of DNA forensics to species conservation came with the saga of rhinoceros marauding in Africa. Two species, white and black rhinos, survive today as critically endangered species in southern and eastern Africa. Both species were objects of intensive conservation

protection and relocation efforts in the twentieth century. At the end of the nineteenth century, Southern white rhinoceros (*Ceratotherium simum simum*) numbers had declined to between 20 and 50 animals but were restored to more than 20,000, primarily in South Africa, in an effort led by dedicated South African conservationist Ian Player (133). Black rhinoceros populations dove from approximately 850,000 in the early twentieth century to 2,400 in 1995 and increased to 5,055 by 2013 (134). Both rhinoceros species have seen a hundredfold increase in illegal hunting since 2007. More than 7,000 rhinoceros have been killed through poaching across Africa, with South Africa suffering the highest losses. The slaughter is driven by an appetite for rhino horns, which are considered both of medicinal value and as a cultural status symbol in Asian countries, mainly Vietnam and China. Organized crime circuits with automatic rifles, helicopters, and sophisticated trafficking and smuggling systems earn enormous profits. The risks are relatively modest: a small fine, a few weeks in prison, or a slap-on-the-hand warning.

Enter Cindy Harper, DVM, an amazing veterinary geneticist at the University of Pretoria. At first, Cindy supported her lab and research by assessing the parentage and breeding structure of farm animals and some captive wildlife programs. She embraced the rhino poaching crisis with a personal passion. She optimized a group of 23 short tandem repeat DNA markers specific for rhinos to assess species, gender, and individual identification of confiscated specimens. Cindy's team developed an extensive database of rhinoceros DNA profiles and demographic information named RhODIS® (Rhino DNA Index System), modeled after CODIS, the US Federal Bureau of Investigation criminal DNA database (135, 136).

The RhODIS® forensic system involved training and certification of police and wildlife officials, a chain-of-custody compliant sampling methodology used for live and dead rhinoceros and rhinoceros horns, an eRhODIS™ field data collection app, and state-of-the-art DNA genetic individualization. Harper has trained more than 500 rangers, police officers, and conservationists. To date, more than 20,000 individual rhinoceros specimens and genotypes have been accessioned to the RhODIS® database. These data include more than 5,800 forensic case samples for which links were made between recovered horns, blood-stained evidence items, and specific rhinoceros carcasses in >120 cases. In the most recent cases, this forensic genetic individualization allowed heavy punishments upon conviction, establishing international legal precedents for prosecuting and convicting smugglers of rhino horns suitable for trafficking in other endangered species (136).

Cindy asked me to serve as her PhD supervisor and her counselor. I used our Snowball experience to fine-tune the forensic development of chain of custody, plus the population genetic statistics for confirming DNA matches based upon accepted DNA fingerprinting product rule calculations. We learned an enormous amount of legal detail from the stories of Cindy and her team. To date, Cindy's group has submitted hundreds of DNA forensic case reports. So far, more than 50 prosecutions of poachers have ensued, with most leading to swift convictions and life-altering sentences of 1–20 years of prison. Because of Cindy's dedication and determination, DNA fingerprint identification has set a new high standard for wildlife forensics and prosecution. Her vigilance has changed the world positively and I, for one, am humbled and proud to have played a minor supporting role.

COMMUNICATING SCIENCE ADVANCES

Publishing research findings in peer-reviewed journals, hopefully of high esteem, is the currency of good science, critical for performance reviews and securing of funding for future research opportunities. Our team of talented students, fellows, and colleagues released more than 850 articles together, including several more wide-ranging review articles (scattered across the citations below). I also learned that clear and more general syntheses beyond technical reports are of special

value. Since my early youth, I had admired the skillful prose and illustrations of *Scientific American*, knowing that these articles would inspire or inform the general science background of important discoveries featured in high school and college textbooks.

I dreamed that someday I might write a piece in *Scientific American* that would reach beyond my traditional peer audience to young students of science, teachers, and nonspecialists. Looking back, my 50-year career allowed for five *Scientific American* articles, each on a different topic [cheetahs (72), giant pandas (59), *CCR5-Δ32* (100), Felidae natural history (57) and the cat domestication process (137)]. In 2003, I published a semi-popular book, *Tears of the Cheetah and Other Tales from the Genetic Frontier* (60). *Tears* contained 14 science adventure stories stimulated by accounts of our work that had been featured in TV and newsprint media. Although these professional outlets made us look amazingly good, *Tears* allowed what I hoped was a more accurate and detailed narrative, because I lived the tales personally. I wrote *Tears* to explain our work to my mother; to my sister Carol, a folk singer; and to other interested nonscientists.

The 14-year series of *Genetic Maps*, *Vol. I–VI* was undertaken to be a telephone book–like compendium of genetic maps of all species (26). The notion was successful, but it simply got too big, with the genome revolution unfolding, and yielded to scores of important web-based gene and DNA databases. In 2006, Bill Nash, Joan Menninger, and I decided to collect all the G-banded mammalian karyotype figures that were published (or not) and present them in a single *Atlas of Mammalian Chromosomes* (138). In 2020, a dramatically updated version of the *Atlas*, edited by Alexander Graphodatsky, Polina Perelman, and myself, will appear with karyotypes of 1,020 species plus chromosome painting illustrations of comparative studies of evolutionary genome organization (139). Lastly, my wife, Diane, reminds me that my 20-year tenure as editor of *Journal of Heredity* (1988–2008; published by the American Genetic Association, of which Diane was Managing Editor) clearly opened multiple doors and contacts that led to many new collaborations and dear friends.

CONCLUSIONS

So I doubt this treatise can be abridged in the 35 minutes Big Russ had allotted for his son. And I am certain that there are countless alternative views to be offered someday by dear friends, colleagues, and chroniclers. But this is my story told in my words. I do apologize to so many colleagues who have important roles but are unmentioned here. To me the ride was exhilarating, satisfying, and remarkably satisfactory. Although I might have preferred major league baseball or rock music if I were blessed with the required talent (I was not), my adventures and experience in science have more than exceeded my expectations. It is not just the rush of an important flashpoint discovery but the satisfaction of hearing of personal discoveries without any appropriation (not so necessary anyway), and the smile of seeing my own scientific philosophy and even temperament displayed when a former student delivers a science lecture advance. As I traveled across the globe, tracking lions, pandas, whales, and sick people, I worked with impressive, dedicated people and tried to help them out through science.

My closing parable relates to why I enjoy the extraordinary science experts I encounter so frequently. Folk singer Harry Chapin once interviewed a hero of his, country singer Pete Seeger. Harry asked Pete how he felt after all these years being involved in social activism. Seeger paused, leaned back in his chair, and drawled slowly, "Harry, my involvement in these causes, benefits, marches, demonstrations—I'm not sure they made a difference…"

Pete was being modest, because for 40 years he had stood up for every major issue of his time, embracing human rights and fighting fascism, Nazis, and racism starting in the late 1930s and 1940s. He continued,

> But I can tell you one thing. Involvement with these issues means you are involved with the good people, people with live hearts, live eyes and live heads… Just think about it in terms of your lives. Who are the people who are your best friends? Who are the people you keep coming back to? Who are the people that make your life worthwhile?…Usually they are the people who are committed to something. So in the final analysis commitment, in and of itself—irrespective of whether you win or not, is something that truly makes your life worthwhile.

So many times across my life I have embraced this precise emotion in pursuit of important science queries. I have been graced to work and interact with the precise people "with live hearts, live eyes and live heads." Certainly scientists discuss, argue, clash, and compromise with their colleagues. We may not always agree or succeed, but we certainly have tried to hit the highest possible standard. My dad urged me to meet all the innovators in my field, as Pete Seeger wanted to befriend the prescient pathfinders. I have been privileged to do both, one day at a time across so many years. To me, that may be the essence of a wonderful life and an exciting science career.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Looking back, it seems clear that my career meanderings were blessed with special opportunities, gifted students, fellows, colleagues, and benevolent supporters. I cannot mention them all, but I am grateful to each. Nearly all discoveries derive from countless interactions with my terrific graduate students ($N \sim 20$), postdoctoral fellows (>30), dedicated lab and office staff, prescient collaborators, dear colleagues, and friends (**http://dobzhanskycenter.spbu.ru/images/pdf/SJO-CV-June-16-2019.pdf**). I employ the royal "we" in several spots to emphasize that the triumphs I enjoyed are theirs, with myself as a cheerful grateful add-on. The employment support and comfort I received from the US National Cancer Institute–National Institutes of Health (NIH) (1971–2011); the Theodosius Dobzhansky Center in St. Petersburg, Russia (2012–2019); and Nova Southeastern University, Fort Lauderdale, Florida (2013–2019), were the bedrock of my work across the years.

For my initial inspiration, my teachers James Edwards, Ross MacIntyre, Bruce Wallace, and Theodosius Dobzhansky deserve special mention. Support from largely government moneys was paramount in my success. I enjoyed an NIH training grant for my graduate education at Cornell. The NCI-NIH provided a hefty budget as several wonderful scientist-administrators believed in me and steered resources to my discretion. Dick Adamson, Richard Klaussner, George VandeWoude, Tony Fauci, George Todaro, and Robert Gallo deserve special thanks. I could not have done this well without their encouragement and cheerful support. So important in my reminiscence are the joys, laughter, support, and love I received from my family—Diane, the steady, cheerful, happy lifetime partner, and Kirsten and Meghan, my daughters, who taught me humility and caution and to set all my goals and standards so very, very high.

LITERATURE CITED

1. Russert T. 2007. *Wisdom of Our Fathers: Lessons and Letters from Daughters and Sons*. New York: Random House
2. Peters TJ, Waterman RH Jr. 1982. *In Search of Excellence: Lessons from America's Best-Run Companies*. New York: Harper & Row. Presumed 1st ed.

3. O'Brien SJ, MacIntyre RJ. 1969. An analysis of gene enzyme variability in natural populations of *Drosophila melanogaster* and *D. simulans*. *Am. Nat.* 103:97–113
4. Lewontin RC, Hubby JL. 1966. A molecular approach to the study of genic heterozygosity in natural populations. 11. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics* 54:595–609
5. O'Brien SJ, MacIntyre RJ. 1971. Empirical demonstration of a transient linkage disequilibrium in *Drosophila*. *Nature* 230:335–36
6. O'Brien SJ, MacIntyre RJ. 1972. The α-glycerophosphate cycle in *Drosophila melanogaster*. I. Biochemical and developmental aspects. *Biochem. Genet.* 7:141–61
7. O'Brien SJ, MacIntyre RJ. 1972. The α-glycerophosphate cycle in *Drosophila melanogaster*. II. Genetic aspects. *Genetics* 71:127–38
8. O'Brien SJ, MacIntyre RJ. 1972. The α-glycerophosphate cycle in *Drosophila melanogaster*. III. The effect of "null" mutations at α Gphd 1 locus on viability. *Am. Nat.* 106:967–71
9. O'Brien SJ, Shimada Y. 1974. The α-glycerophosphate cycle in *Drosophila melanogaster*. IV. Metabolic, ultrastructural, and adaptive consequences of α Gphd 1 "null" mutations. *J. Cell Biol.* 63:864–82
10. O'Brien SJ, MacIntyre RJ. 1971. A biochemical genetic map of the *Drosophila* genome. *Drosophila Inf. Serv.* 46:89–93
11. Huebner RJ, Todaro GJ. 1969. Oncogenes of RNA tumor viruses as determinants of cancer. *PNAS* 64:1087–94
12. O'Brien SJ, Nash WG, Goodwin JL, Lowy DR, Chang EH. 1983. Dispersion of the *ras* family of transforming genes to four different chromosomes in man. *Nature* 302:839–42
13. Kinzler KW, Bigner SH, Bigner DD, Trent JM, Law ML, et al. 1987. Identification of an amplified highly expressed gene in a human glioma. *Science* 236:70–73
14. Watson DK, McWilliams-Smith MJ, Kozak C, Reeves R, Gearhart J, et al. 1986. Conserved chromosomal positions of dual domains of the ets protooncogene in cats, mice, and humans. *PNAS* 83:1792–96
15. Park M, Dean M, Cooper CS, Schmidt M, O'Brien SJ, Blair DG. 1986. Mechanism of *met* oncogene activation. *Cell* 45:895–904
16. Nienhuis AW, Bunn HF, Turner PH, Gopal TV, Nash WG, et al. 1985. Expression of the human c-*fms* proto oncogene in hematopoietic cells and its deletion in the $5q^-$ syndrome. *Cell* 42:421–28
17. Bonner T, O'Brien SJ, Nash WG, Rapp UR, Morton CC, Leder P. 1984. The human homologs of the raf (mil) oncogene are located on human chromosomes 3 and 4. *Science* 223:71–74
18. O'Brien SJ, Nash WG. 1982. Genetic mapping in mammals: chromosome map of the domestic cat. *Science* 216:257–65
19. Li G, Hillier LW, Grahn RA, Zimin AV, David VA, et al. 2016. A high-resolution SNP array-based linkage map anchors a new domestic cat draft genome assembly and provides detailed patterns of recombination. *G3* 6:1607–16
20. O'Brien SJ, Johnson W. 2005. Big cat genomics. *Annu. Rev. Genom. Hum. Genet.* 6:407–29
21. O'Brien SJ, Menotti-Raymond M, Murphy WJ, Yuhki N. 2002. The Feline Genome Project. *Annu. Rev. Genet.* 36:657–86
22. O'Brien SJ, Wienberg J, Lyons LA. 1997. Comparative genomics: lessons from cats. *Trends Genet.* 13:393–99
23. O'Brien SJ, Menotti-Raymond M, Murphy WJ, Nash WG, Wienberg J, et al. 1999. The promise of comparative genomics in mammals. *Science* 286:458–81
24. O'Brien SJ, Womack JE, Lyons LA, Moore KJ, Jenkins NA, Copeland NG. 1993. Anchored reference loci for comparative genome mapping in mammals. *Nat. Genet.* 3:103–12
25. Lyons LA, Laughlin TF, Copeland NG, Jenkins NA, Womack JE, O'Brien SJ. 1997. Comparative anchor tagged sequences (CATS) for integrative mapping of mammal genomes. *Nat. Genet.* 15:47–56
26. O'Brien SJ. 1980–1993. *Genetic Maps: Locus Maps of Complex Genomes*. New York: Cold Spring Harbor Lab. Press. 6 vols.
27. Lindblad-Toh K, Garber M, Zuk O, Lin MF, Parker BJ, et al. 2011. A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* 478(7370):476–82

28. Pontius JU, Mullikin JC, Smith D, Lindblad-Toh K, Gnerre S, et al. 2007. Initial sequence and comparative analysis of the cat genome. *Genome Res.* 11:1675–89
29. Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, et al. 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 438(7069):803–19
30. Genome 10K Community Sci. 2009. Genome 10K: a proposal to obtain whole-genome sequence for 10,000 vertebrate species. *J. Hered.* 100(6):659–74
31. Pennisi E. 2009. No genome left behind. *Science* 326:794–95
32. Koepfli K-P, Paten B, Genome 10K Community Sci., O'Brien SJ. 2015. The Genome 10K Project: a way forward. *Annu. Rev. Anim. Biosci.* 3:57–211
33. Voolstra A, Christian R, GIGA Community Sci. (COS), Wörheide G, Lopez JV. 2017. Advancing genomics through the Global Invertebrate Genomics Alliance (GIGA). *Invertebr. Syst.* 31:1–7
34. Robinson GE, Hackett KJ, Purcell-Miramontes M, Brown SJ, Evans JD, Goldsmith MR. 2011. Creating a buzz about insect genomes. *Science* 331(6023):1386
35. Lewin HA, Robinson GE, Kress WJ, Baker WJ, Coddington J, et al. 2018. Earth BioGenome Project: sequencing life for the future of life. *PNAS* 115:4325–33
36. Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, et al. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* 346(6215):1320–31
37. Zhang G, Li C, Li Q, Li B, Larkin DM, et al. 2014. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* 346(6215):1311–20
38. Dobrynin P, Liu S, Tamazian G, Xiong Z, Yurchenko AA, et al. 2015. Genomic legacy of the African cheetah, *Acinonyx jubatus*. *Genome Biol.* 16:277. **https://doi.org/10.1186/s13059-015-0837-4**
39. Cho YS, Li H, Haolong H, Hang L, Jiaohui X, et al. 2013. The tiger genome and comparative analysis with other feline genomes. *Nat. Commun.* 4:2433
40. Yim HS, Cho YS, Guang X, Kang SG, Jeong JY, et al. 2014. Minke whale genome and aquatic adaptation in cetaceans. *Nat. Genet.* 46:88–92. **https://doi.org/10.1038/ng.2835**
41. Kim S, Cho YS, Kim HM, Chung O, Kim H, et al. 2016. Comparison of three dietary groups in mammals: carnivore, omnivore, and herbivore genome analyses with a new leopard assembly. *Genome Biol.* 17(1):211
42. Koepfli K-P, Tamazian G, Wildt D, Dobrynin P, Kim C, et al. 2019. Whole genome sequencing and re-sequencing of the sable antelope (*Hippotragus niger*): a resource for monitoring diversity in *ex situ* and in situ populations. *G3* 9(6):1785–93. **https://doi.org/10.1534/g3.119.400084**
43. Marra NJ, Stanhope MJ, Jue NK, Wang M, Sun Q, et al. 2019. White shark genome reveals ancient elasmobranch adaptations associated with wound healing and the maintenance of genome stability. *PNAS* 116(10):4446–55. **https://doi.org/10.1073/pnas.1819778116**
44. Choo SW, Rayko M, Tan TK, Hari R, Komissarov A, et al. 2016. Pangolin genomes and the evolution of mammalian scales and immunity. *Genome Res.* 26(10):1312–22
45. Grigorev K, Kliver S, Dobrynin P, Komissarov A, Wolfsberger W, et al. 2018. Innovative assembly strategy contributes to understanding the evolution and conservation genetics of the endangered *Solenodon paradoxus* from the island of Hispaniola. *Gigascience* 7(6):giy025. **https://doi.org/10.1093/gigascience/giy025**
46. O'Brien SJ. 2012. Genome empowerment for the Puerto Rican parrot – *Amazona vittata*. *GigaScience* 1:13–16
47. Tamazian G, Simonov S, Dobrynin P, Makunin A, Logachev A, et al. 2014. Annotated features of domestic cat – *Felis catus* genome. *GigaScience* 3:13
48. Farré M, Kim J, Proskuryakova AA, Zhang Y, Kulemzina AI, et al. 2019. Evolution of gene regulation in ruminants differs between evolutionary breakpoint regions and homologous synteny blocks. *Genome Res.* 29:576–89. **https://doi.org/10.1101/gr.239863.118**
49. Dobzhansky T. 1973. Nothing in biology makes sense except in the light of evolution. *Am. Biol. Teach.* 35(3):125–29
50. Morgan GJ. 1998. Emile Zuckerkandl, Linus Pauling, and the molecular evolutionary clock, 1959–1965. *J. Hist. Biol.* 31(2):155–78
51. Hedges SB, Kumar S. 2009. *The Timetree of Life*. Oxford, UK: Oxford Univ. Press

52. O'Brien SJ, Eisenberg JF, Miyamoto M, Hedges SB, Kumar S, et al. 1999. Genome maps 10. Comparative genomics. Mammalian radiations. Wall chart. *Science* 286:463–78
53. Murphy WJ, Eizirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ. 2001. Molecular phylogenetics and the origins of placental mammals. *Nature* 409:614–18
54. Madsen O, Scally M, Douady CJ, Kao DJ, DeBry RW, et al. 2001. Parallel adaptive radiations in two major clades of placental mammals. *Nature* 409(6820):610–14
55. Murphy WJ, Eizirik E, O'Brien SJ, Madsen O, Scally M, et al. 2001. Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science* 294:2348–51
56. Johnson WE, Eizirik E, Murphy WJ, Pecon-Slattery J, Antunes A, et al. 2006. The late Miocene radiation of modern Felidae: a genetic assessment. *Science* 311:73–77
57. O'Brien SJ, Johnson WE. 2007. The evolution of cats. Genomic paw prints in the DNA of the world's wild cats have clarified the cat family tree and uncovered several remarkable migrations in their past. *Sci. Am.* 297:68–75
58. O'Brien SJ, Nash WG, Wildt DE, Bush ME, Benveniste RE. 1985. A molecular solution to the riddle of the giant panda's phylogeny. *Nature* 317:140–44
59. O'Brien SJ. 1987. The ancestry of the giant panda. *Sci. Am.* 257:102–7
60. O'Brien SJ. 2003. *Tears of the Cheetah and Other Tales from the Genetic Frontier*. New York: St. Martin's
61. Teeling E, Springer MS, Madsen O, Bates P, O'Brien SJ, Murphy WJ. 2005. A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science* 307:580–84
62. Perelman P, Johnson WE, Roos C, Seuánez HN, Horvath JE, et al. 2011. A molecular phylogeny of living primates. *PLOS Genet*. 7(3):e1001342
63. Wayne RK, Benveniste RE, Janczewski DN, O'Brien SJ. 1989. Molecular and biochemical evolution of the Carnivora. In *Carnivore Behavior, Ecology, and Evolution*, ed. JL Gittleman, pp. 465–94. New York: Comstock Pub. Assoc., Cornell Univ. Press
64. Eizirik E, Murphy WJ, Koepfli K-P, Johnson WE, Dragoo JW, et al. 2010. Pattern and timing of diversification of the mammalian order Carnivora inferred from multiple nuclear gene sequences. *Mol. Phylogenet. Evol.* 56(1):49–63
65. Gaubert P, Antunes A, Mengd H, Lind M, Peignée S. 2017. The complete phylogeny of pangolins: scaling up resources for the molecular tracing of the world's most trafficked mammals. *J. Hered.* 109(4):347–35
66. Buckley-Beason VA, Johnson WE, Nash WG, Stanyon R, Menninger JC, et al. 2006. Molecular evidence for species-level distinctions in clouded leopards. *Curr. Biol.* 16:2371–76
67. Trigo TC, Freitas TRO, Kunzler G, Cardoso L, Silva JCR, et al. 2008. Inter-species hybridization among Neotropical cats of the genus *Leopardus*, and evidence for an introgressive hybrid zone between *L. geoffroyi* and *L. tigrinus* in southern Brazil. *Mol. Ecol.* 17(19):4317–33
68. Roca AL, Georgiadis N, Pecon-Slattery J, O'Brien SJ. 2001. Genetic evidence for two species of elephant in Africa. *Science* 293:1473–77
69. Lu Z, Karish WB, Janczewski DN, Frazier-Taylor H, Sajuthi D, et al. 1996. Genomic differentiation among natural populations of orangutan (*Pongo pygmaeus*). *Curr. Biol.* 6:1326–36
70. O'Brien SJ, Wildt DE, Goldman D, Merril CR, Bush M. 1983. The cheetah is depauperate in genetic variation. *Science* 221:459–62
71. O'Brien SJ, Roelke ME, Marker L, Newman A, Winkler CA, et al. 1985. Genetic basis for species vulnerability in the cheetah. *Science* 227:1428–34
72. O'Brien SJ, Wildt DE, Bush M. 1986. The cheetah in genetic peril. *Sci. Am.* 254:84–92
73. Roelke ME, Martenson JS, O'Brien SJ. 1993. The consequences of demographic reduction and genetic depletion in the endangered Florida panther. *Curr. Biol.* 3:340–50
74. Johnson WE, Onorato DP, Roelke ME, Land ED, Cunningham M, et al. 2010. Genetic restoration of the Florida panther. *Science* 329:1641–45
75. Wildt DE, Bush M, Goodrowe KL, Packer C, Pusey AE, et al. 1987. Reproductive and genetic consequences of founding isolated lion populations. *Nature* 329:328–31
76. Gilbert DA, Lehman N, O'Brien SJ, Wayne RK. 1990. Genetic fingerprinting reflects population differentiation in the California Channel Island fox. *Nature* 344:764–67

77. Uphyrkina O, O'Brien SJ. 2003. Applying molecular genetic tools to the conservation and action plan for the critically endangered Far Eastern leopard (*Panthera pardus orientalis*). *C. R. Biol.* 326:93–97
78. Liu YC, Sun X, Driscoll C, Miquelle DG, Xu X, et al. 2018. Genome-wide evolutionary analysis of natural history and adaptation in the world's tigers. *Curr. Biol.* 28(23):3840–49. **https://doi.org/10.1016/j.cub.2018.09.019**
79. Luo S-J, Kim J-H, Johnson WE, van der Walt J, Martenson J, et al. 2004. Phylogeography and genetic ancestory of tigers (*Panthera tigris*). *PLOS Biol.* 2:2277–93
80. Culver M, Johnson WE, Pecon-Slattery J, O'Brien SJ. 2000. Genomic ancestry of the American puma (*Puma concolor*). *J. Hered.* 91:186–97
81. Antunes A, Troyer JL, Roelke ME, Pecon-Slattery J, Packer C, et al. 2008. The evolutionary dynamics of the lion *Panthera leo* revealed by host and viral population genomics. *PLOS Genet*. 4(11):e1000251
82. Hamilton WD. 1964. The genetical evolution of social behaviour. *J. Theor. Biol.* 7:17–52
83. Gilbert DA, Packer C, Pusey AE, Stephens JC, O'Brien SJ. 1991. Analytical DNA fingerprinting in lions: parentage, genetic diversity, and kinship. *J. Hered.* 82:378–86
84. Packer C, Gilbert DA, Pusey AE, O'Brien SJ. 1991. Kinship, cooperation and inbreeding in African lions: a molecular genetic analysis. *Nature* 351:562–65
85. Kukekova AV, Johnson JL, Xiang X, Feng S, Liu S, et al. 2018. Red fox genome assembly identifies genomic regions associated with tame and aggressive behaviours. *Nat. Ecol. Evol.* 2:1479–91
86. Spellberg B, Taylor-Blake B. 2013. On the exoneration of Dr. William H. Stewart: debunking an urban legend. *Infect. Dis. Poverty* 2:3
87. Hardy WD, Essex M, McClelland AJ. 1980. *Feline Leukemia Virus*. New York: Elsevier/North Holland
88. Brown M, Cunningham MW, Roca AL, Troyer JL, Johnson WE, O'Brien SJ. 2008. Genetic characterization of emerging feline leukemia virus in the free-ranging Florida panther population. *Emerg. Infect. Dis.* 14:252–59
89. Pedersen NC, Ho EW, Brown ML, Yamamoto JK. 1987. Isolation of a T-lymphotropic virus from domestic cats with an immunodeficiency-like syndrome. *Science* 235(4790):790–93
90. Olmsted RA, Langley R, Roelke ME, Goeken RM, Adger-Johnson D, et al. 1992. Worldwide prevalence of lentivirus infection in wild feline species: epidemiologic and phylogenetic aspects. *J. Virol.* 66:6008–18
91. Troyer JL, Pecon-Slattery J, Roelke ME, Johnson W, VandeWoude S, Vazquez-Salat N. 2005. Seroprevalence and genomic diversity of circulating strains of feline immunodeficiency virus among *Felidae* and *Hyaenidae* species. *J. Virol.* 79:8282–94
92. Pecon-Slattery J, Troyer JL, Johnson WE, O'Brien SJ. 2008. Evolution of feline immunodeficiency virus in Felidae: implications for human health and wildlife ecology. *Vet. Immunol. Immunopathol.* 123:32–44
93. Roelke-Parker ME, Munson L, Packer C, Kock R, Cleaveland S, et al. 1996. A canine distemper virus epidemic in Serengeti lions (*Panthera leo*). *Nature* 379:441–45
94. Carpenter MA, Appel MJG, Roelke-Parker ME, Munson L, Hofer H, et al. 1998. Genetic characterization of canine distemper virus in Serengeti carnivores. *Vet. Immunol. Immunopathol.* 65:259–66
95. Cent. Dis. Control Prev. 2004. Revised U.S. surveillance case definition for severe acute respiratory syndrome (SARS) and update on SARS cases—United States and worldwide, December 2003. *JAMA* 291(2):173–74
96. Drazen JM. 2003. SARS—looking back over the first 100 days. *N. Engl. J. Med.* 349:319–20
97. Holmes KV. 2003. SARS-associated coronavirus. *N. Engl. J. Med.* 348:1948–51
98. Pearks-Wilkerson AJ, Teeling EC, Troyer JL, Bar-Gal GK, Roelke ME, et al. 2004. Coronavirus outbreak in cheetahs: lessons for SARS. *Curr. Biol.* 14:227–28
99. Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, et al. 1996. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the *CKR5* structural gene. *Science* 273:1856–62
100. O'Brien SJ, Dean M. 1997. In search of AIDS-resistance genes. *Sci. Am.* 277:44–51
101. O'Brien SJ, Nelson GW, Winkler CA, Smith MW. 2000. Polygenic and multifactorial disease gene association in man: lessons from AIDS. *Annu. Rev. Genet.* 34:563–91
102. Choe H, Farzan M, Sun Y, Sullivan N, Rollins B, et al. 1996. The β-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell* 85(7):1135–48

103. Doranz BJ, Rucker J, Yi Y, Smyth RJ, Samson M, et al. 1996. A dual-tropic primary HIV-1 isolate that uses fusin and the β-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. *Cell* 85(7):1149–58
104. Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, et al. 1996. Identification of a major co-receptor for primary isolates of HIV-1. *Nature* 381:661–66
105. Alkhatib G, Combadiere C, Broder CC, Feng Y, Kennedy PE, et al. 1996. CC CKR5: a RANTES, MIP-lα, MIP-lβ receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science* 272:1955–58
106. Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, et al. 1996. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 86(3):367–77
107. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, et al. 1996. Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 382(6593):722–25
108. Dean M, Carrington M, O'Brien SJ. 2002. Balanced polymorphism selected by genetic versus infectious human disease. *Annu. Rev. Genom. Hum. Genet.* 3:263–92
109. Carrington M, O'Brien SJ. 2003. The influence of HLA genotype on AIDS. *Annu. Rev. Med.* 54:535–51
110. O'Brien SJ, Hendrickson S. 2013. Host genomic influences on HIV/AIDS. *Genome Biol.* 14:201–14
111. Hutter G, Nowak D, Mossner M, Ganepola S, Mussig A, et al. 2009. Long-term control of HIV by CCR5 delta32/delta32 stem-cell transplantation. *N. Engl. J. Med.* 360:692–98
112. Allers K, Hütter G, Hofmann J, Loddenkemper C, Rieger K, et al. 2011. Evidence for the cure of HIV infection by CCR5Δ32/Δ32 stem cell transplantation. *Blood* 117:2791–99
113. Gupta RK, Abdul-Jawad S, McCoy LE, Mok HP, Peppa D, et al. 2019. HIV-1 remission following CCR5Δ32/Δ32 haematopoietic stem-cell transplantation. *Nature* 568:244–48
114. Reshef R, Luger SM, Hexner EO, Loren AW, Frey NV, et al. 2012. Blockade of lymphocyte chemotaxis in visceral graft-versus-host disease. *N. Engl. J. Med.* 367(2):135–45. **https://doi.org/10.1056/NEJMoa1201248**
115. Reshef R, Mangan JK, Luger SM, Loren AW, Hexner EO, et al. 2014. Extended CCR5 blockade in graft-versus-host disease prophylaxis—a phase II study. *Blood* 124(21):2491
116. Joy MT, Ben Assayag E, Shabashov-Stone D, Liraz-Zaltsman S, Mazzitelli J, et al. 2019. CCR5 is a therapeutic target for recovery after stroke and traumatic brain injury. *Cell* 176:1143–57
117. Glass WG, McDermott DH, Lim JK, Lekhong S, Yu SF, et al. 2006. CCR5 deficiency increases risk of symptomatic West Nile virus infection. *J. Exp. Med.* 203:35–40
118. Cohen J. 2019. Inside the circle of trust. *Science* 365:430–37
119. Cyranoski D. 2019. The CRISPR-baby scandal: what's next for human gene-editing. *Nature* 566:440–42. **https://www.nature.com/articles/d41586-019-00673-1**
120. Xie W, Agniel D, Shevchenko A, Malov S, Svitin A, et al. 2017. Genome-wide analyses reveal gene influence on HIV disease progression and HIV-1C acquisition in Southern Africa. *AIDS Res. Hum. Retrovir.* 33:597–609
121. Zeng Z, Guan L, An P, Sun S, O'Brien SJ, Winkler CA. 2008. A population-based study to investigate the host genetic factors associated with hepatitis B infection and pathogenesis in the Chinese population. *BMC Infect. Dis.* 8:1
122. Khakoo SI, Thio CL, Martin MP, Brooks CR, Gao X, et al. 2004. HLA and NK cell inhibitory receptor genes in resolving hepatitis C virus infection. *Science* 305:872–74
123. Tang M, Lautenberger JA, Gao X, Sezgin E, Hendrickson SL, et al. 2012. The principal genetic determinants for nasopharyngeal carcinoma in China involve the *HLA* class I antigen recognition groove. *PLOS Genet.* 11:e1003103
124. Skloot R. 2011. *The Immortal Life of Henrietta Lacks*. New York: Broadway Books
125. Lucey BP, Nelson-Rees WA, Hutchins GM. 2009. Henrietta Lacks, HeLa cells, and cell culture contamination. *Arch. Pathol. Lab. Med.* 133:1463–67
126. Nelson-Rees W, Flandermeyer RP. 1976. Inter- and intraspecies contamination of human breast tumor cell lines HBC and BrCa5 and other cell cultures. *Science* 191:96–97; 195:1343–44
127. Harris NL, Gang DL, Quay SC, Poppema S, Nelson Rees WA, O'Brien SJ. 1981. Contamination of Hodgkin's disease cell cultures. *Nature* 289:228–30

128. O'Brien SJ, Shannon JE, Gail MH. 1980. A molecular approach to the identification and individualization of human and animal cells in culture: isozyme and allozyme genetic signatures. *In Vitro* 16:119–35
129. O'Brien SJ, Kleiner G, Olson R, Shannon J. 1977. Enzyme polymorphisms as genetic signatures in human cell cultures. *Science* 195:1345–48
130. Menotti-Raymond MA, David VA, O'Brien SJ. 1997. Pet cat hair implicates murder suspect. *Nature* 386:774
131. Menotti-Raymond MA, David VA, Wachter LA, Butler JM, O'Brien SJ. 2005. An STR forensic typing system for genetic individualization of domestic cat (*Felis catus*) samples. *J. Forensic Sci.* 50:1061–70
132. Dowling P, exec. prod. 2002. Purr-fect match. *Forensic Files*, Season 7, Episode 7, aired Feb. 12. **https://www.youtube.com/watch?v=e8ip3HoiG8k**
133. Player I. 2013. *The White Rhino Saga*. Johannesburg: Jonathan Ball
134. Milliken T, Shaw J. 2012. *The South Africa–Viet Nam Rhino Horn Trade Nexus: a deadly combination of institutional lapses, corrupt wildlife industry professionals and Asian crime syndicates*. Rep., TRAFFIC, Johannesburg, S. Afr.
135. Harper CK, Vermeulen GJ, Clarke AB, De Wet JI, Guthrie AJ. 2013. Extraction of nuclear DNA from rhinoceros horn and characterization of DNA profiling systems for white (*Ceratotherium simum*) and black (*Diceros bicornis*) rhinoceros. *Forensic Sci. Int. Genet.* 7:428–33
136. Harper C, Ludwig A, Clarke A, Makgopela K, Yurchenko A, et al. 2018. Robust forensic matching of confiscated horns to individual poached African rhinoceros. *Curr. Biol.* 28(1):13–14
137. Driscoll CA, Clutton-Brock J, Kitchener AC, O'Brien SJ. 2009. The taming of the cat. *Sci. Am.* 300(6):68–75
138. O'Brien SJ, Menninger JC, Nash WG. 2006. *Atlas of Mammalian Chromosomes*. New York: John Wiley & Sons
139. O'Brien SJ, Graphodatsky A, Perelman P. 2020. *Atlas of Mammalian Chromosomes*. New York: John Wiley & Sons. In press. 2nd ed.