

# Taxonomy and Species Discovery



James Hanken

Museum of Comparative Zoology

Harvard University, USA

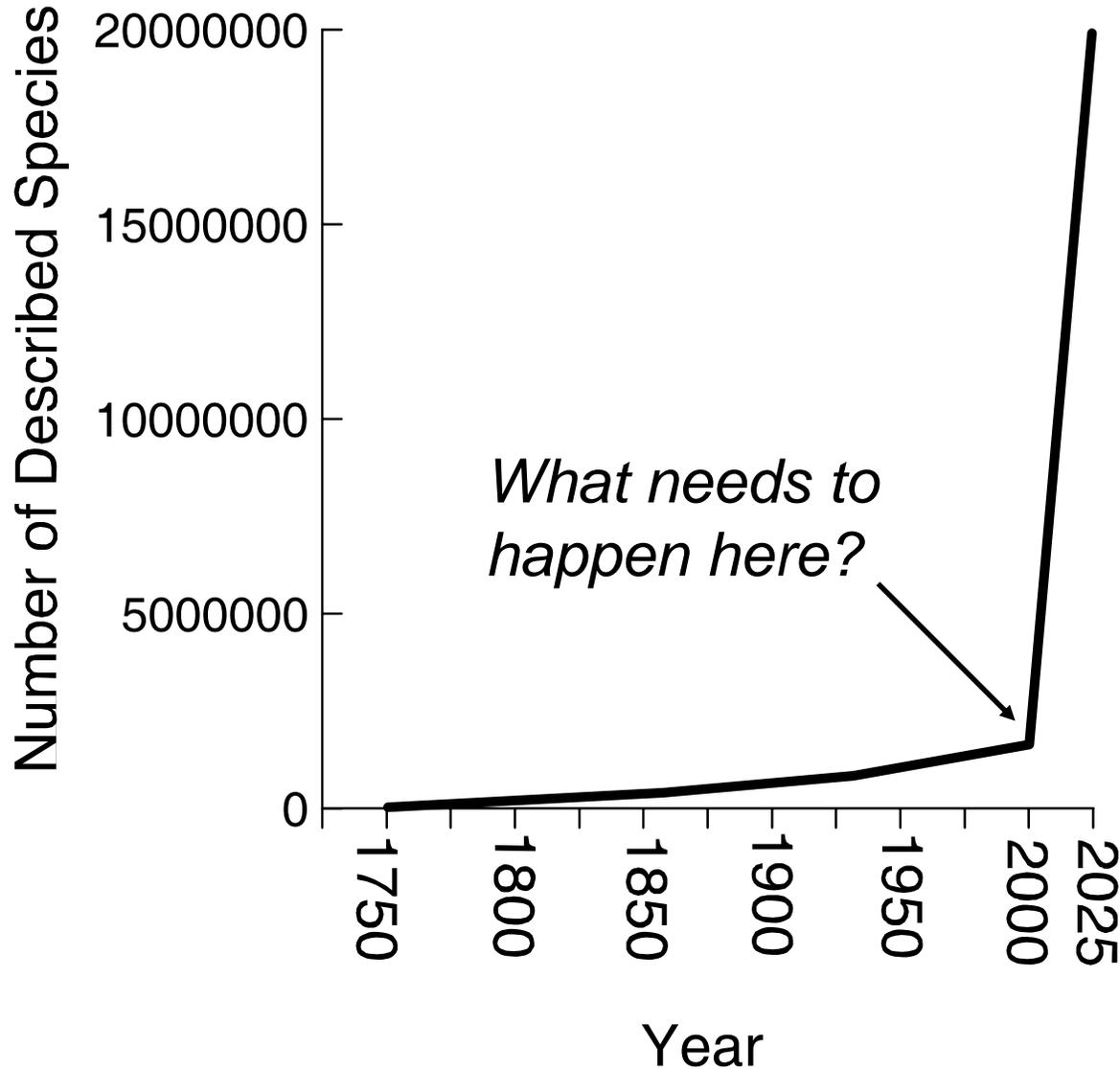
# Outline

- Current rate of species discovery.
- Practical (realistic) solutions to the “taxonomic impediment.”
- Deploying people and resources towards an “integrative taxonomy.”

# “Taxonomic impediment”

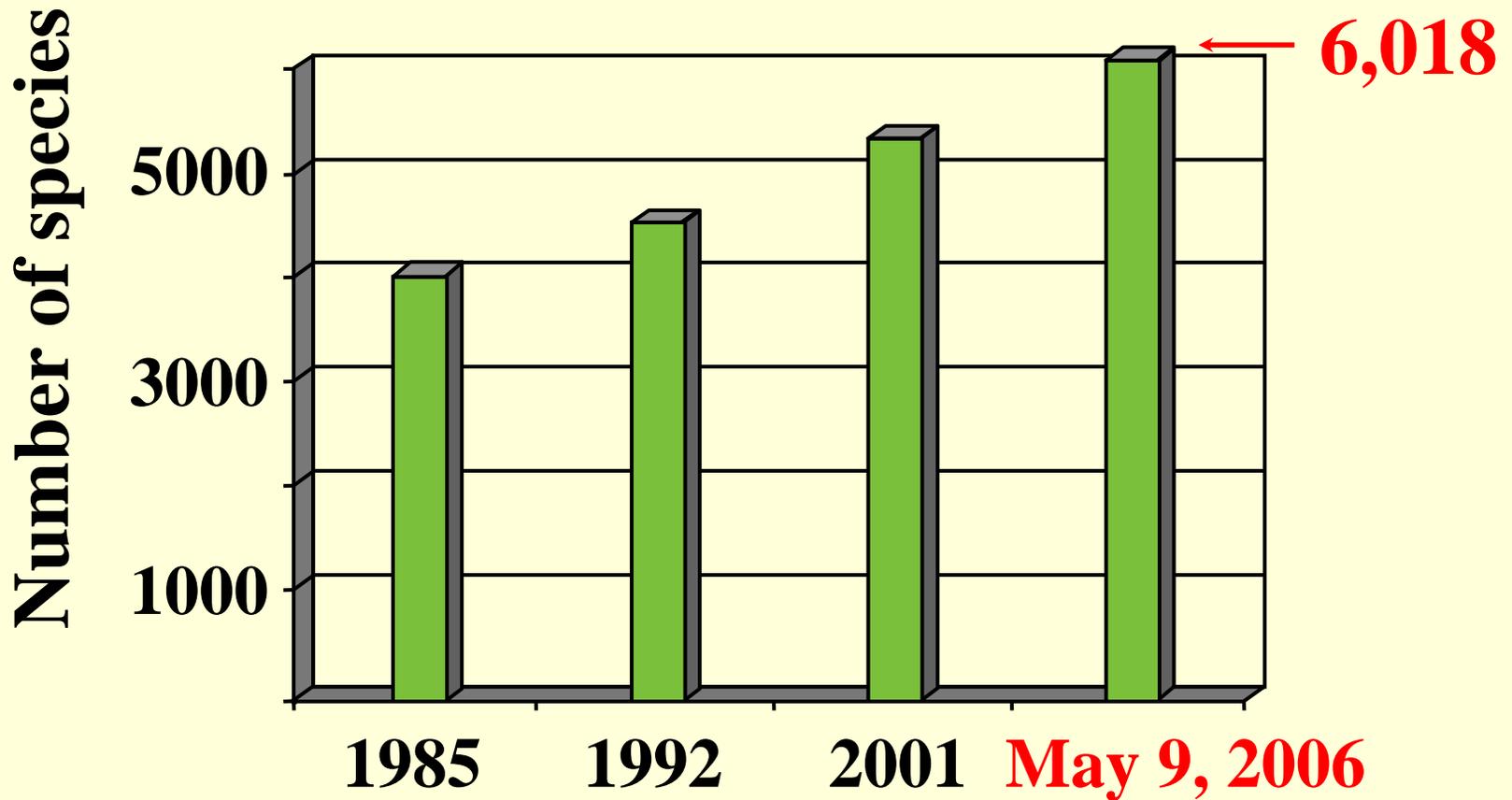
Rate of description of new species is suppressed by an insufficient number of trained professionals and inadequate support for their work.

# Rate of Species Discovery



# Increase in the Global Number of Amphibian Species: > 50% since 1985

---



**248 new species since Feb 2005!**



# ALL CATFISH SPECIES INVENTORY

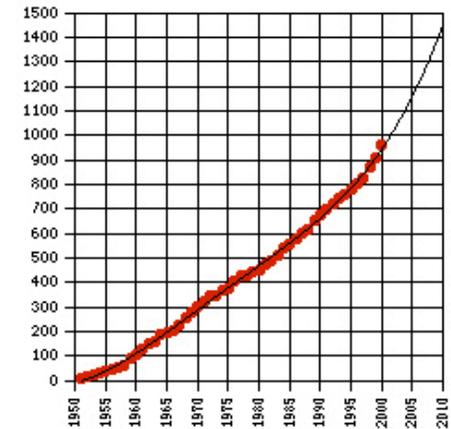
## All Catfish Species (Siluriformes) – Phase I of an Inventory of the Otophysi Proposal Stage

Asked over 170 taxonomists to estimate numbers of new species for taxa of expertise

APPENDIX 1:  
APPROXIMATE NUMBERS OF VALID GENERA AND SPECIES AND KNOWN/ESTIMATED  
UNDESCRIBED SPECIES OF CATFISHES BASE ON EXPERTISE OF ACSI PARTICIPANTS

No.	Family/Group	Approx. Valid Genera	Approx. Valid Species	Known/Est. New Species	ACSI Participating Taxonomists & Systematists (students named when possible)	No.	Family/Group	Approx. Valid Genera	Approx. Valid Species	Known/Est. New Species	ACSI Participating Taxonomists & Systematists (students named when possible)
1	Akysidae (includes <i>Parakysis</i> )	4	27	20	de Pinna, H.H. Ng, Roberts	22	Loricariidae	88	673	205	Akama, Aquino, Armbruster, Barrera, Baskin, Batten, Bichuette, Bockmann, Britski, Britto, Burr, Cardoso, Carvajal, Cezar Nolasco, Covain, Cristina Santana, de Pinna, Evans, Fisch-Muller, Garavello, Ghazzi, Hardman, Langeani, Le Bail, Lehmann, Lima, Lucinda, Lujan, Montoya-Burgos, Moreira, C. Oliveira & students, J.C. de Oliveira, Ortega-Lara, Oyakawa, Pavanelli, Pereira, Pimentel, Provenzano, Quaggio-Grassiotto, Quevedo, Rapp Py-Daniel, Retzer, Reis, Ribeiro, Royero, Salcedo, Sarmiento, Schaefer, Schmidt, Shibatta, D. Stewart, Takako, Taphorn, Thomas, Trajano, Vieira, Weber, Zawadzki, Zuanon
2	Amblycipitidae	3	25	20	H.H. Ng	23	Malapteruridae	2	19	0	Golubtsov, Roberts
3	Amphiliidae	9	58	35	Diogo, Golubtsov, H.H. Ng, Roberts, Skelton	24	Mochokidae	10	180	30	Bills, Bloomer, De Vos, Fermon, Friel, Golubtsov, H.H. Ng, Page, Skelton, van der Bank, Villet
4	Ariidae	18	138	16	Acero, Aguilera, Allen, Betancur, Kailola, Marceniuk, Menezes, Polanco, Vidthayanon, Wang, Zhang	25	Nematogenyidae	1	1	0	Arratia & students, de Pinna, C. Oliveira & students, Quaggio-Grassiotto, Shibatta, Vila
5	Aspredinidae	12	36	11	Britto, Figueiredo, Friel, Taphorn	26	Pangasiidae	3	26	?	Gustiano, Roberts, Teugels
6	Astroblepidae	1	54	10	Barrera, Buitrago, Carvajal, de Pinna, C. Oliveira & students, Ortega, Quaggio-Grassiotto, Sarmiento, Schaefer, Shibatta	27	Pimelodidae	30	67	45	Abreu, Akama, Azpelicueta, Buitrago, Burr, Dahdul, Ferraris, Garavello, Littmann, Lundberg, Ortega-Lara & students, Parisi, Shibatta & students, D. Stewart, Vari
7	Auchenipteridae (includes Ageneiosidae, Centromochlidae)	21	86	14	Akama, Baskin, Borges, Britski, Ferraris, Langeani, Lima, Lundberg, Pety, Reis, Royero, Soares-Porto, Sabaj, D. Stewart, Vari, Walsh	28	Plotosidae	10	35	?	Allen
8	Auchenoglanididae	5	28	? 0		29	Pseudopimelodidae	5	26	14	de Pinna, C. Oliveira & students, Ortega-Lara & students, Pety, Quaggio-Grassiotto, Shibatta & students
9	Austroglanididae	1	3	0	Bills	30	Schilbeidae	14	53	15	Ferraris, Golubtsov, H.H. Ng, Page, Roberts, Vari
10	Bagridae (includes <i>Obyra</i> , <i>Horabagrus</i> )	17	131	60	Ferraris, Linder, Manimekalan, Marquardt, H.H. Ng, Watanabe	31	Scoloplacidae	1	4	?	de Pinna, C. Oliveira, Quaggio-Grassiotto, Shibatta
11	Callichthyidae	8	177	45	Britto, Burgess, de Pinna, DoNascimento, Lima, C. Oliveira & students, Quaggio-Grassiotto, Reis, Shibatta	32	Siluridae	12	89	40	Ferraris, He & colleagues/students, H.H. Ng, Krudpan, Manimekalan, Roberts, Vari
12	Cetopsidae	7	23	20	de Pinna, Ferraris, Keith, Lundberg, Milani, Ortega, Provenzano, Rapp Py-Daniel, Vari	33	Sisoridae	16	112	70	Dimmick, Edds, He & colleagues/students, Manimekalan, H.H. Ng, Vidthayanon
13	Chacidae	1	3	1	Ferraris, H.H. Ng	34	Trichomycteridae	42	176	55	Arratia & students, Azpelicueta, Barrera, Barthem, Baskin, Bichuette, Bockmann, Carvajal, de Pinna, DoNascimento, L. Fernandez, Keith, Linder, C. Oliveira & students, Provenzano, Quaggio-Grassiotto, Sarmiento, Schaefer, Schmidt, Shibatta, Trajano, Vari, Vila, Villarreal, Wosiacki & students, Zuanon
14	Clariidae (includes <i>Heteropneustes</i> )	15	93	25	Adriaens, Devaere, Manimekalan, Musschoot, H.H. Ng, Skelton, Sudarto, Teugels	35	<i>Conorhynchos</i>	1	1	0	Bockmann, de Pinna, Ferraris, Lundberg
15	Claroteidae	8	58	3	Bills, Golubtsov, Lundberg	36	<i>Ancharius</i>	1	2	5	Friel, H.H. Ng
16	Cranoglanididae	1	3	0	—recently revised—	37	<i>Chiapas catfish</i> *	0	0	1	Hendrickson, Lundberg, Rodiles
17	Diplomystidae	2	6	1	Arratia & students, Lundberg, Vila	<b>Totals</b>		<b>436</b>	<b>2734</b>	<b>873</b>	
18	Doradidae	30	74	25	Agostinho, Akama, Burr, Dahdul, Keith, Langeani, Lundberg, Rapp Py-Daniel, Sabaj, Sousa, Weitzell						
19	Erethistidae	6	13	20	Ferraris, H. H. Ng						
20	Heptapteridae	24	190	52	Aguilera, Akama, Alejandra Diaz, Azpelicueta, Bichuette, Bockmann, DoNascimento, Ferraris, Garavello, Guazzelli, Lasso, Lundberg, Ortega-Lara & students, Oyakawa, Pety, Provenzano, Reis, Royero, Salcedo, Taphorn, Trajano, Zuanon						
21	Ictaluridae	7	44	15	Burr, Eisenhour, Garcia de León, Grady, Hardman, Hendrickson, Littmann, Page, Thomas						

### Cumulative Valid Species Curve

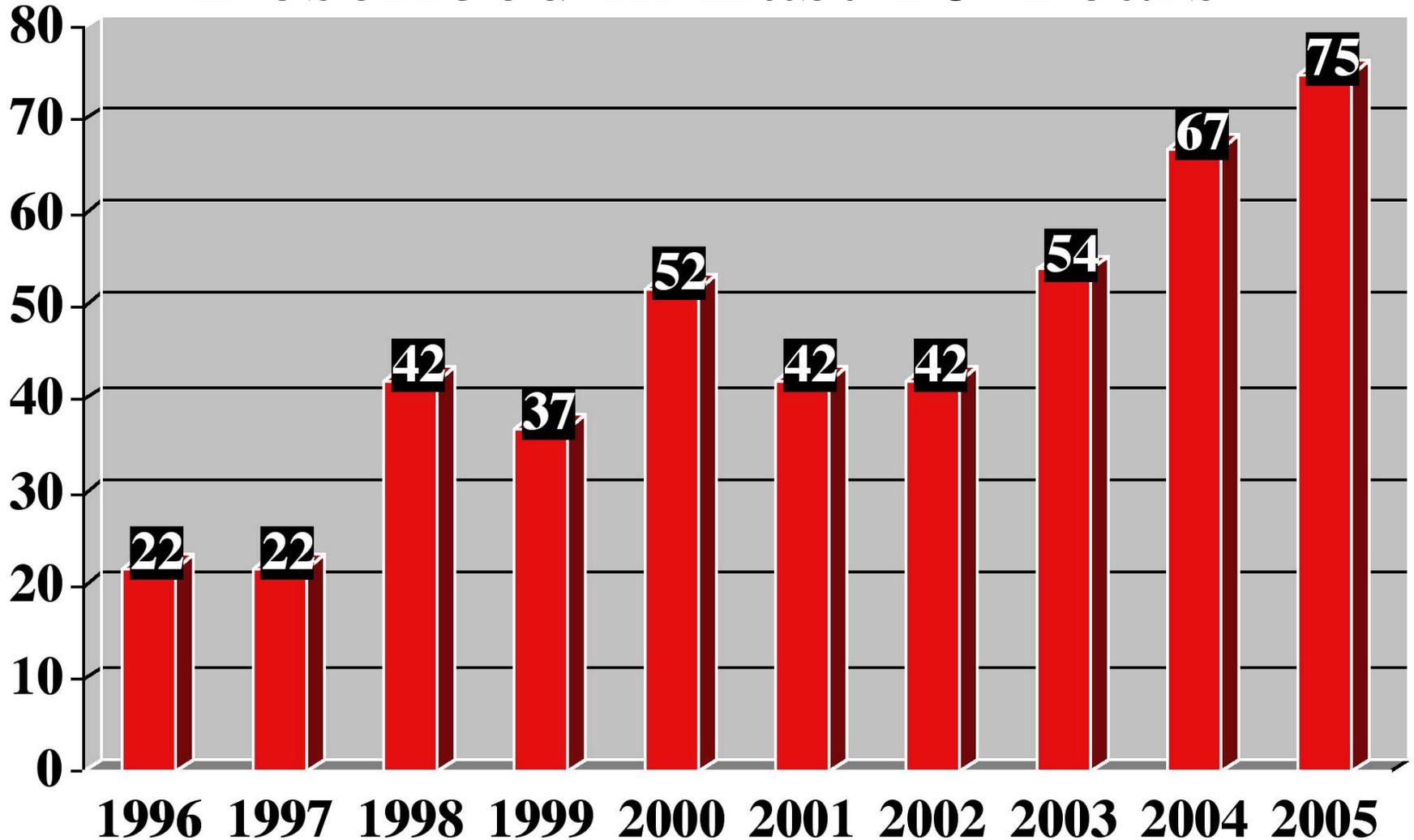


Siluriformes ~2,734 to 2,855 valid spp.  
~873 to 1,750 undescribed



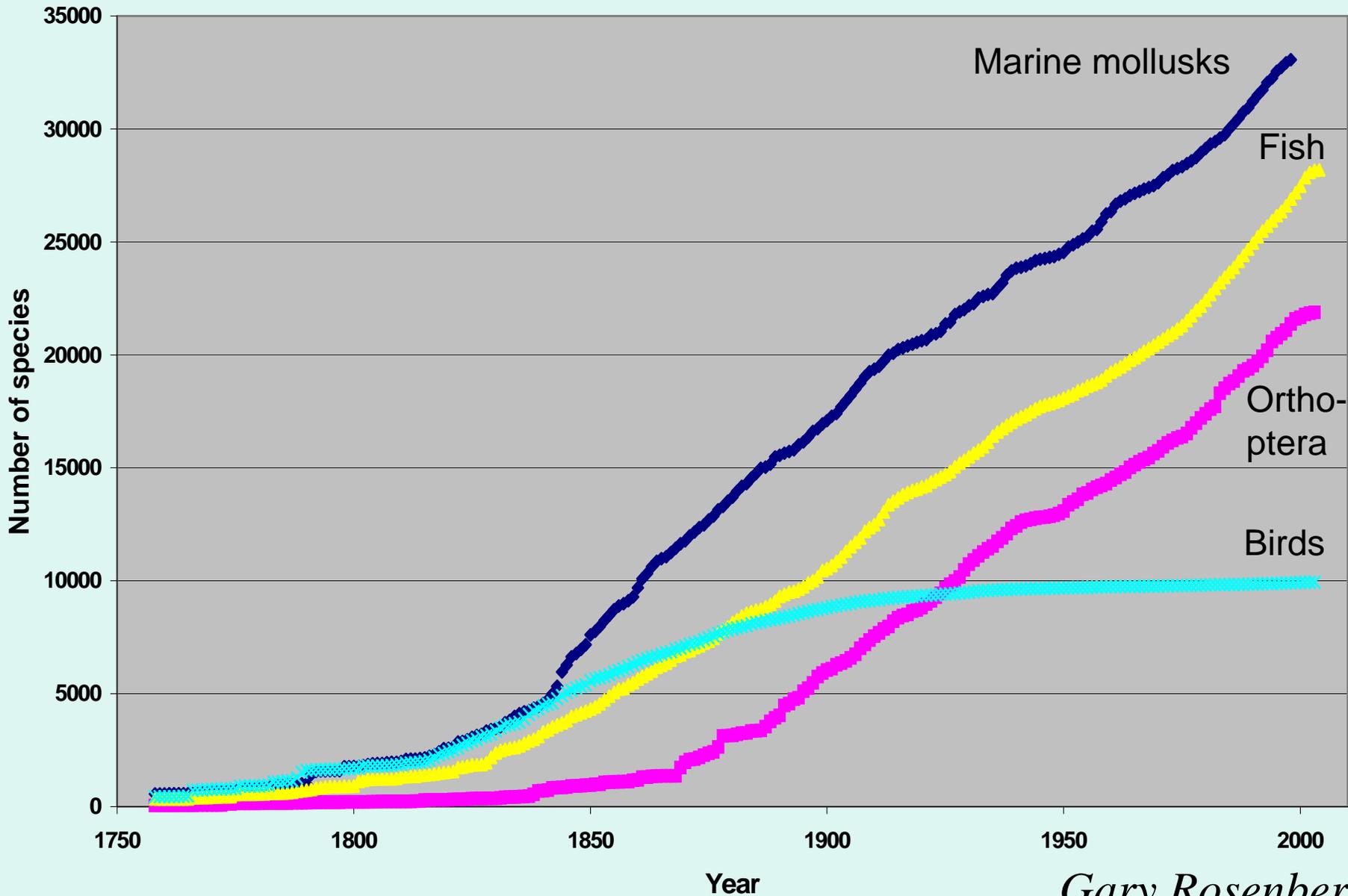
John Lundberg

# Number of New Species of Catfishes Described in Last 10 Years



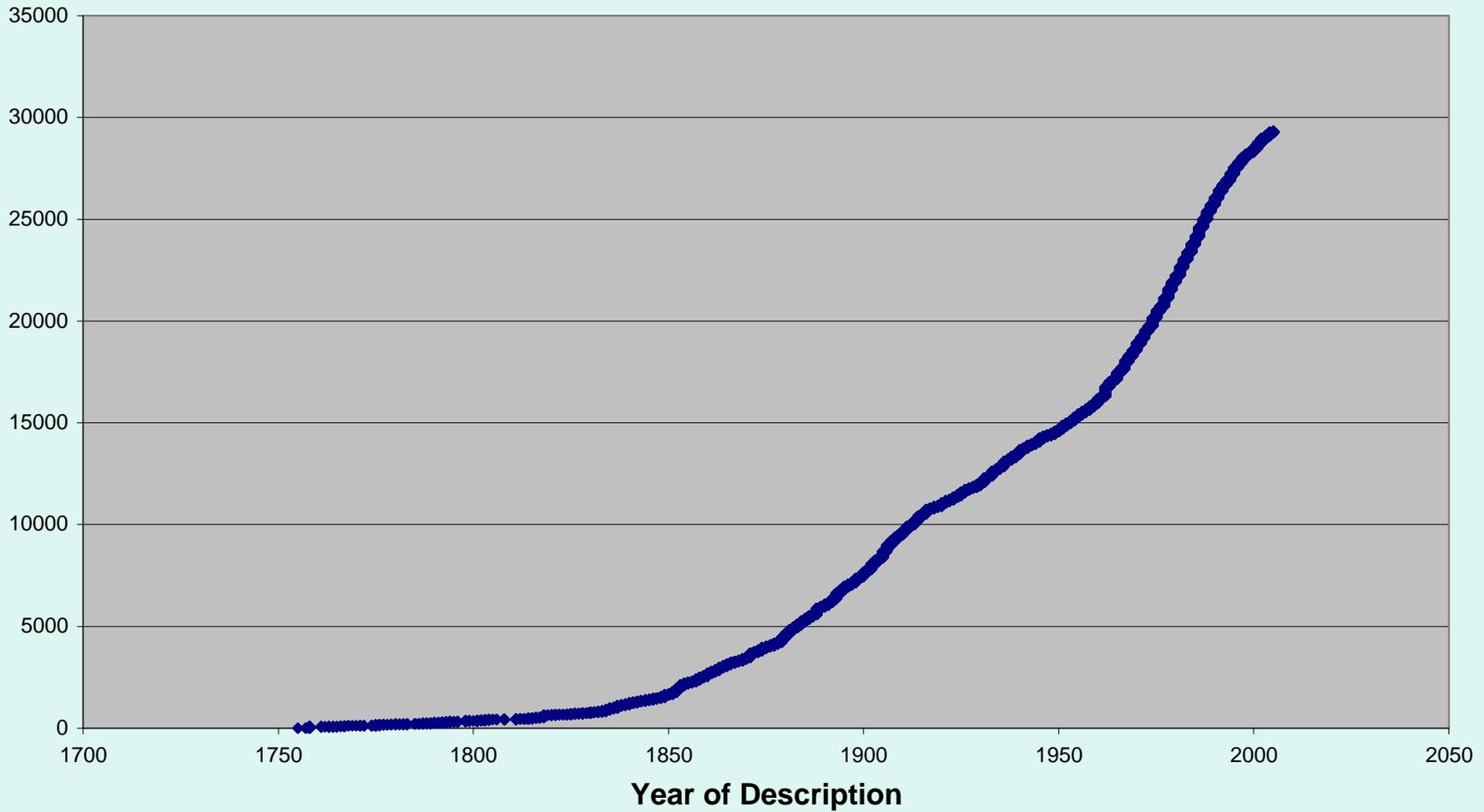
↑  
ACS

# Species naming curves for various taxa



*Gary Rosenberg*

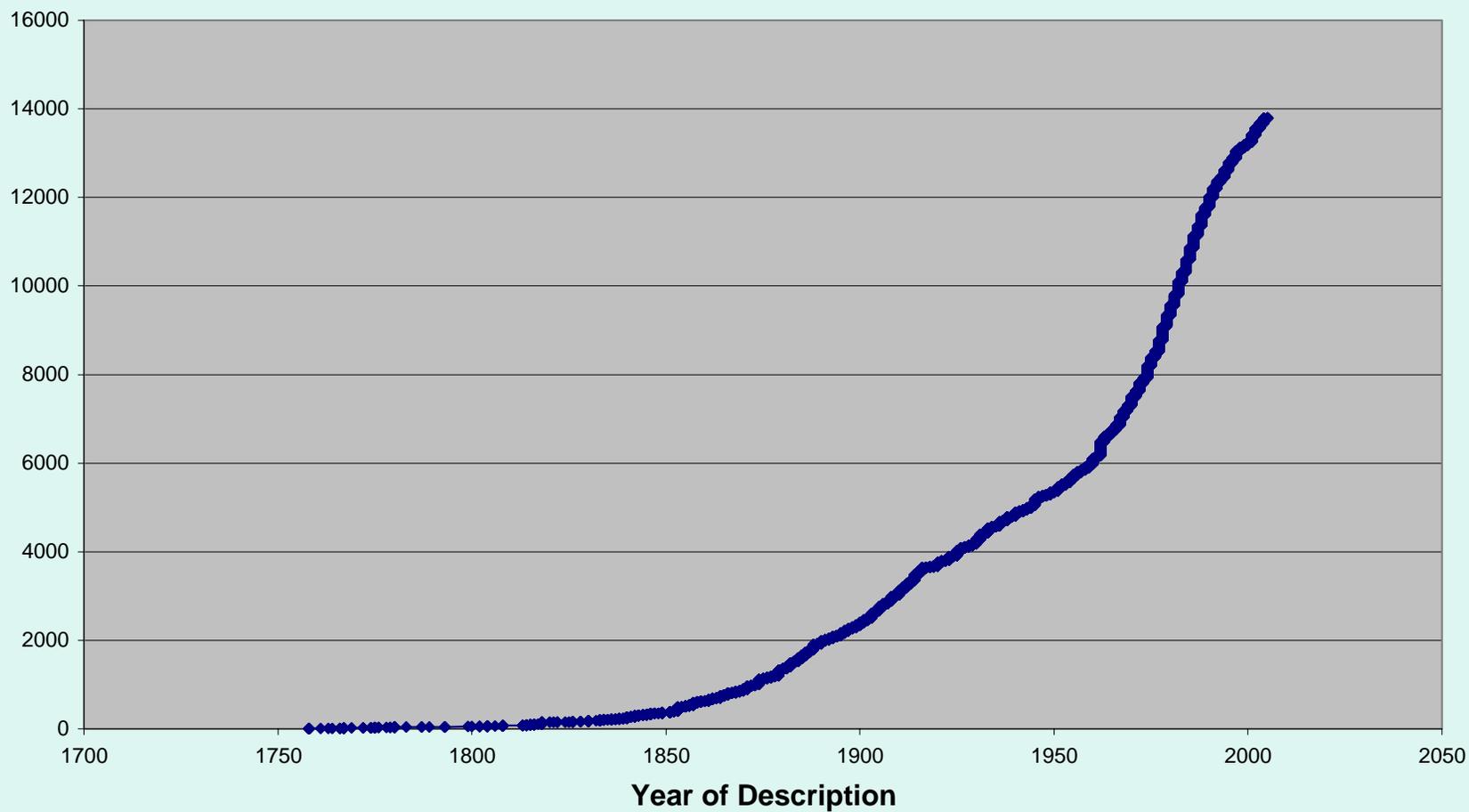
# All Crustaceans (N = 29,312)



*Jody Martin*



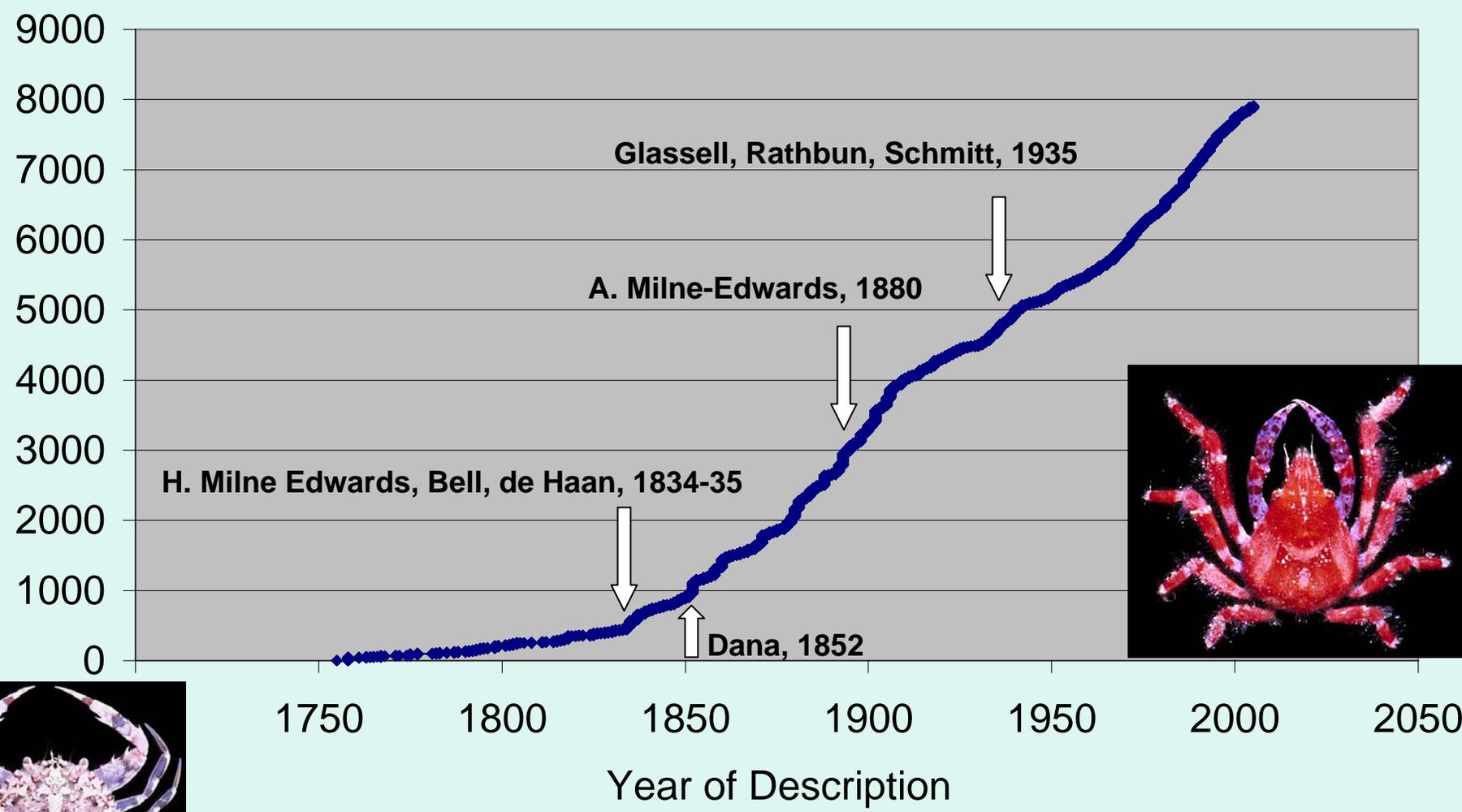
## Peracarida (N = 13,797)



*Jody Martin*



# Decapoda (N = 7,905)



*Jody Martin*

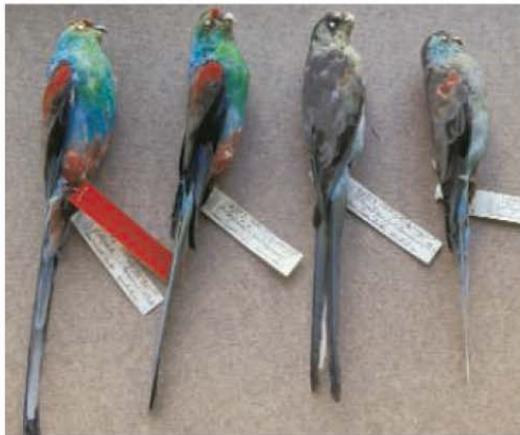
# Underlying Premise/Assumption

- The number of formally trained, practicing taxonomists is not likely to increase much above what it is now.
- Therefore, taxonomy must look to technical innovation as the principal means of increasing the rate of discovery and naming of new species.

## SYSTEMATICS

## Philadelphia Institution Forced to Cut Curators

A chronic budget shortfall has forced the oldest natural history institution in the United States to lay off 5% of its staff. Outside scientists are especially concerned that the Academy of Natural Sciences in Philadelphia is losing three of its 10 curators, including the overseer of a prized, nearly 200-year-old ornithology collection. The move is part of a trend of cutbacks at natural history museums. “We’re losing positions. It’s of national con-



**Scientific treasure.** The ornithology collection at the financially troubled Philadelphia academy includes specimens of the extinct Australian paradise parrot (*Psephotus pulcherrimus*).

cern,” says Smithsonian Institution ornithologist Helen F. James.

The academy, founded in 1812, runs a museum and research programs and houses 17 million biological specimens. Its \$12 million annual budget has faced deficits of \$500,000 to \$1 million for a decade, explains president and CEO D. James Baker, former head of the National Oceanic and Atmospheric Administration. As a result, Baker says leaders made the “painful decision” last month to lay off 13 of 250 employees across all divisions. The layoffs go into effect over the next 6 months. Thomas Lovejoy, head of the Heinz Center, an environmental think tank in Washington, D.C., and an academy board member, says that the cuts were inevitable. “They just had to address” the deficit, he notes.

The three curators losing their jobs are Leo Joseph, assistant curator and chair of ornithology; Richard McCourt, an associate botany curator; and Dominique Didier-Dagit, an associate curator of ichthyology. Some outside scientists who asked not to be identified suggest that these junior scientists weren’t pulling in enough grant money. Baker doesn’t deny the charge, saying that the acad-

emy tried to keep staff in “areas where we think there is research support from outside agencies.” (Joseph and McCourt referred calls to an academy spokesperson.)

The academy’s ornithology collection, which now has no curator, is a paramount concern. The holdings include many of the earliest specimens collected by North American ornithologists as well as the Australia collection of British ornithologist John Gould. Baker says the academy “has made an absolute commitment to preserve” this resource, which will still have a manager to make it available to scientists. But experts worry that the absence of a curator to add specimens and conduct his or her own research could undermine it. “A collection should be part of a living and breathing community,” says A. Townsend Peterson, ornithology curator of the Natural History Museum at the University of Kansas, Lawrence.

Baker is mum on future staffing plans, saying only that “we can grow our number of curators” if the budget outlook improves. But he predicts that a focus on certain areas, such as watershed management and molecular systematics, will create “a stronger institution.”

—JOCELYN KAISER

CREDITS (TOP TO BOTTOM): ESA; THE ACADEMY OF NATURAL SCIENCES

# New Technologies and Methodologies

- Digital conversion of the taxonomic infrastructure, e.g., natural history collections (specimens and ancillary data), taxonomic literature (species accounts), etc.
- Networking collections and other kinds of cyberinfrastructure.
- DNA barcoding and other genomics approaches.

QuickTime™ and a  
TIFF (LZW) decompressor  
are needed to see this picture.



*Mojave rattlesnake*

QuickTime™ and a  
TIFF (LZW) decompressor  
are needed to see this picture.

QuickTime™ and a  
TIFF (LZW) decompressor  
are needed to see this picture.

QuickTime™ and a  
TIFF (LZW) decompressor  
are needed to see this picture.

QuickTime™ and a  
TIFF (LZW) decompressor  
are needed to see this picture.

QuickTime™ and a  
TIFF (LZW) decompressor  
are needed to see this picture.

QuickTime™ and a  
TIFF (LZW) decompressor  
are needed to see this picture.

QuickTime™ and a  
TIFF (LZW) decompressor  
are needed to see this picture.

QuickTime™ and a  
TIFF (LZW) decompressor  
are needed to see this picture.

# High-resolution X-ray CT (Computed Tomography)

---



*Thorius aureus*

[http://www.digimorph.org/specimens/Thorius\\_aureus/](http://www.digimorph.org/specimens/Thorius_aureus/)

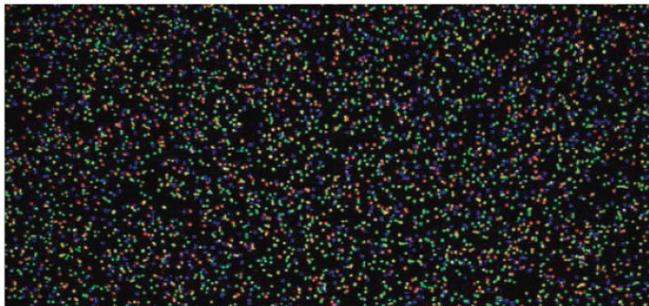


*Astylosternus*



QuickTime™ and a  
TIFF (LZW) decompressor  
are needed to see this picture.

QuickTime™ and a  
TIFF (LZW) decompressor  
are needed to see this picture.



**Charting islands.** Glowing dots on a glass slide mark cloned DNA being sequenced.

## GENE SEQUENCING

# The Race for the \$1000 Genome

Fast, cheap genetic analyses will soon become a reality, and the consequences—good and bad—will affect everybody

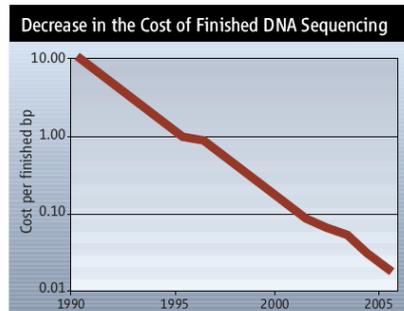
**MARCO ISLAND, FLORIDA**—Computers aren't the only things getting better and cheaper every time you turn around. Genome-sequencing prices are in free fall, too. The initial draft of the first human genome sequence, finished just 5 years ago, cost an estimated \$300 million. (The final draft and all the technology that made it possible came in near \$3 billion.) Last month, genome scientists completed a draft of the genome sequence of the second nonhuman primate—the rhesus macaque—for \$22 million. And by the end of the year, at least one company expects to turn out a full mammalian genome sequence for about \$100,000, a 3000-fold cost reduction in just 6 years.

It's not likely to stop there. Researchers are closing in on a new generation of technology that they hope will slash the cost of a genome sequence to \$1000. "Advances in this field are happening fast," says Kevin McKernan, co-chief scientist at Agencourt Bioscience in Beverly, Massachusetts. "And they are coming more quickly than I think anyone was anticipating." Jeffrey Schloss, who heads the sequencing-technologies grant program at the National Human Genome Research Institute (NHGRI) in Bethesda, Maryland, agrees. "People are roundly encouraged and nervous," Schloss says—encouraged because their own technologies are working, and nervous because their competitors' are too.

A host of these novel sequencing technologies were on display last month at a meeting here. Although no one at the meeting claimed to have cracked the \$1000 genome sequence yet, researchers are getting more confident that it's a real possibility. "From what I've listened to the last few days, there is no physical principle that says we shouldn't be able to do a \$1000 genome,"

says Harvard University sequencing pioneer George Church.

Even today, the declining cost of genome sequencing is triggering a flowering of basic research, looking at broad-ranging topics such as how the activation of genes is regulated and understanding genetic links to cancer. And as prices continue to drop, sequencing will revolutionize both the way biologists hunt for disease genes and the way medical professionals diagnose and treat diseases. In fact, some researchers say cheap sequencing technology could finally usher in personalized medicine in a major way. "The promise of cheap sequencing is in the understanding of disease and biology, such as cancer, where the genome changes over time," says Dennis Gilbert, chief scientist of Applied Biosystems, the leading gene-sequencing-technology company based in Foster City, California. "It will enable different kinds of science to be done." Of course, as with other forms



**Free fall.** As with computer technology, the plunging cost of DNA sequencing has opened new applications in science and medicine.

of high technology, that promise brings new risks as well. Researchers expect cheap sequencing to raise concerns about the proliferation of bioterrorism agents as well as patient privacy.

### The race is on

The first group to produce a technology capable of sequencing a human genome sequence for \$1000 will get instant gratification, as well as potential future profits: In September 2003, the J. Craig Venter Science Foundation promised \$500,000 for the achievement. That challenge has since been picked up by the Santa Monica, California-based X Prize Foundation, which is expected to up the ante to between \$5 million and \$20 million. But the competition really began in earnest in 2004, when the National Institutes of Health launched a \$70 million grant program to support researchers working to sequence a complete mammal-sized genome initially for \$100,000 and ultimately for \$1000. That program has had an "amazing" effect on the field, encouraging researchers to pursue a wide variety of new ideas, says Church. That boost in turn has led to a miniexplosion of start-up companies, each pursuing its own angle on the technology (see table, p. 1546).

All are racing to improve or replace a technology first developed by Fred Sanger of the U.K. Medical Research Council in the mid-1970s that is the basis of today's sequencing machines. The technique involves making multiple copies of the DNA to be sequenced, chopping it up into small pieces, and using those pieces as templates to synthesize short strands of DNA that will be exact complements of stretches of the original sequence. The synthesis essentially mimics the cell's processes for copying DNA.

The technology relies on the use of modified versions of the four bases that make up DNA, each of which is tagged with a different fluorescent marker. A short DNA snippet called a primer initiates the synthesis at a specific point on the template DNA, and the altered bases—which are vastly outnumbered by normal bases in the mix of reagents used to perform the synthesis—stop the process when one of them is tacked onto the end of the growing DNA strand. The result is a soup of newly synthesized DNA fragments, each of which started at the same point but ends at a different base along the chain.

Today's sequencers separate these fragments by passing the soup through tiny capillaries containing a gel; the shorter the fragment, the faster it moves through the gel. The process, known as capillary electrophoresis, is so effective that

\* Advances in Genome Biology and Technology Conference, Marco Island, Florida, 8–11 February 2006.

# People and Resources are Limited

- Taxonomic initiatives, including DNA barcoding, must coordinate with related activities in systematic and comparative biology to minimize needless redundancy and to maximize the scientific return from investment of “resources” (labor and expense).
- DNA barcoding is consistent with an “integrative taxonomy.”

# TREE OF LIFE web project

## Explore the Tree of Life

### Browse the Site

- [Root of the Tree](#)
- [Popular Pages](#)
- [Sample Pages](#)
- [Recent Additions](#)
- [Random Page](#)
- [Treehouses](#)
- [Biographies](#)
- [Picture Sampler](#)

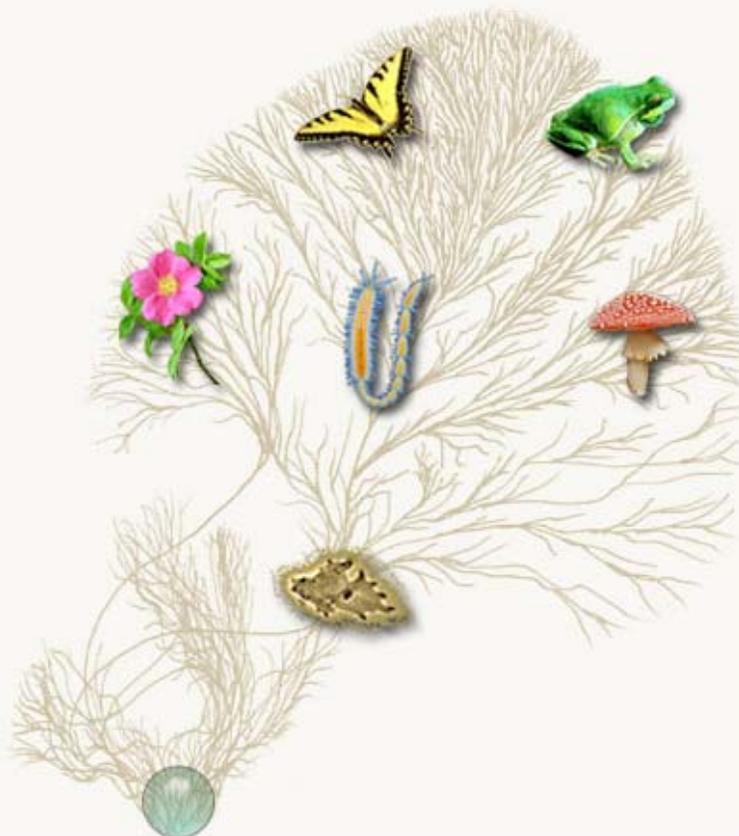
Search

[advanced](#)

### News

Download the structure of the tree used on the ToL web site ...

[read more](#)



### Learn about ...

**Caudata**  
(Salamanders)



"Salamanders are amphibians of the order Caudata (also called Urodela). This taxon includes the newts, a group of rough-skinned species in the salamander family Salamandridae..."

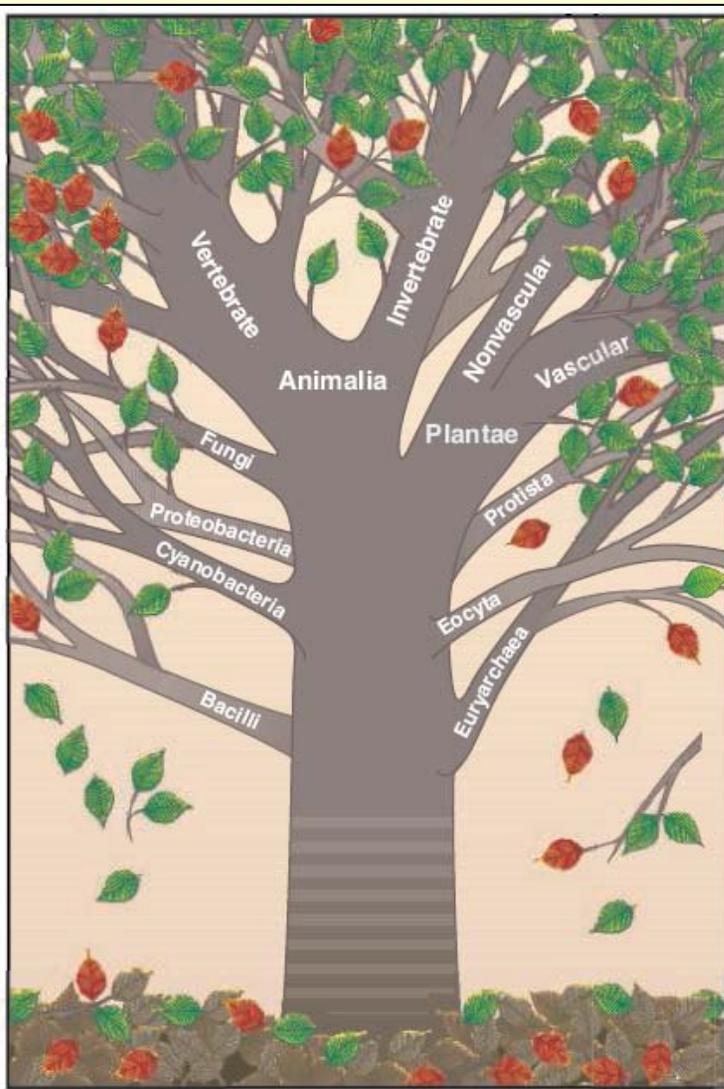
[read more](#)

[previously featured pages](#)

The Tree of Life Web Project (ToL) is a collaborative effort of [biologists from around the world](#). On more than 3000 World Wide Web pages, the project provides information about the diversity of organisms on Earth, their evolutionary history ([phylogeny](#)), and characteristics.

Each page contains information about a particular group of organisms (e.g., [echinoderms](#), [tyrannosaurs](#), [phlox flowers](#), [cephalopods](#), [club fungi](#), or the [salamanderfish of Western Australia](#)). ToL pages are linked one to another hierarchically, in the form of the evolutionary tree of life. Starting with the [root of all Life on Earth](#) and moving out along diverging branches to individual species, the [structure of the ToL project](#) thus illustrates the genetic connections between all living things.

[read more about the Tree of Life Web Project...](#)



**Building the Tree of Life.** A current view of the Tree of Life (7). Information is biased toward vertebrate animals and vascular plants (the thick branches); lesser-known groups such as bacteria, fungi, and protists are largely underrepresented. Also shown are species known to science (green leaves), extinct species (leaf litter, brown), endangered species (falling leaves), and species for which "barcode" information is available (red leaves).

“If the supertree approach establishes the trunk and thick branches of the Tree of Life, then perhaps the barcoding approach is more appropriate for discerning the twigs and leaves of the tree.

It may turn out that two different databases are needed for these two distinct purposes—establishing relationships versus diagnosing species.”

Crandall & Buhay (2004)

