Special Feature JAMIA

Presentation of the Morris F. Collen Award to Joshua Lederberg, PhD

Edward H. Shortliffe, MD, PhD, Thomas C. Rindfleisch, MS

The American College of Medical Informatics is an honorary society established to recognize those who have made sustained contributions to the field. Its highest award, for lifetime achievement and contributions to the discipline of medical informatics, is the Morris F. Collen Award. Dr. Collen's own efforts as a pioneer in the field stand as the embodiment of creativity, intellectual rigor, perseverance, and personal integrity.

At most once a year, the College gives its highest recognition to those whose attainments have, throughout their careers, substantially advanced the science and art of medical informatics. In 1999, the College was proud to present the Collen Award to Joshua Lederberg, PhD. Biologist, geneticist, computer scientist, and scholar, Dr. Lederberg's lifelong contributions to biomedicine and science, as well as to medical informatics, make him highly deserving of the recognition embodied in the Collen Award.

The Early Years

Joshua Lederberg was born in Montclair, New Jersey, on May 23, 1925, the son of a rabbi. His interest in a scientific career was mentioned in an essay that he wrote at the age of seven. His family moved to New York City when Lederberg was a child, and Lederberg was able to attend Stuyvesant High School, a school that allowed him to concentrate in the sciences. In New York, Lederberg was also able to take advantage of facilities such as the American Institute, which made laboratory space and equipment available to talented high school science students.

On graduating from high school at 16, Lederberg took advantage of a local scholarship to attend Columbia University and, after completing his undergraduate degree, he began medical school there. He did not see combat service during World War II, but from July 1943 he was enrolled in the U.S. Navy's V-12 training program, which combined an accelerated premedical and medical curriculum with active service as a hospital corpsman in a U.S. naval hospital.

As a medical student, Dr. Lederberg undertook a series of experiments with Edward L. Tatum, demonstrating sexual recombination in bacteria. With his growing interest in a career in laboratory science. Lederberg decided to leave medical school to pursue a PhD, which he received from Yale in 1948 at the age of 23.

After graduate school, Dr. Lederberg joined the Genetics Department at the University of Wisconsin, which at the time was part of the University's School of Agriculture. He eventually helped form the Department of Medical Genetics and served as its first chair.

Affiliation of the authors: Stanford University, Stanford, California.

Correspondence and reprints: Edward H. Shortliffe, MD, PhD, Professor of Medicine, Stanford Medical Informatics, Stanford University School of Medicine, MSOB X-201, 251 Campus Drive, Stanford, CA 94305-5479; e-mail: ⟨ehs@smi.stanford.edu⟩.

Received for publication: 1/12/00; accepted for publication: 1/14/00.

The Nobel Prize

In 1958, when he was but 33 years old, Joshua Lederberg received a telegram of the sort that only a handful of scientists ever receive. He was awarded the Nobel Prize for "his discoveries concerning genetic recombination and the organization of the genetic material of bacteria." He shared the prize with George W. Beadle and Edward L. Tatum, who won "for their discovery that genes act by regulating definite chemical events."

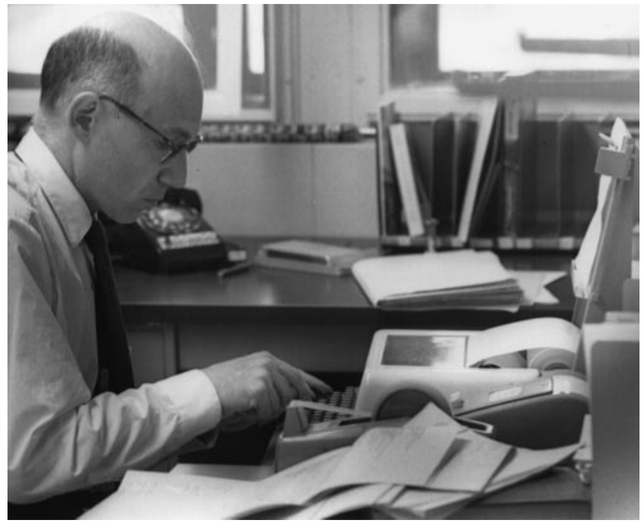
The work that led to Lederberg's Nobel award had begun in 1945, when he was 20 and still a medical student. Work published in 1944 inspired Lederberg to investigate the possibility of sexual reproduction in bacteria, which at the time were thought to reproduce only asexually. Lederberg's work, which formed the basis for his doctoral dissertation, demonstrated that bacteria can in fact reproduce through sexual recombination, and opened up the genetics of micro-organisms to the traditional methods of the field. These methods are now central to the fields of biotechnology and genetic engineering.

Lederberg's name is today mentioned in most textbooks of genetics and microbiology, not only for his demonstration of bacterial conjugation but also for his discovery of transduction and for his coining of the term *plasmid* to denote extrachromosomal genetic material.

His scientific reputation was already established by the caliber of his work, but the Nobel Prize extended his reputation to the general public and allowed him to pursue his interests at the intersection of science, policy, and society. Between 1966 and 1971, he wrote a series of more than 200 articles for the Washington Post, which converted into lay terms many of the pressing scientific issues of the day. Typical articles addressed such themes as "Russian computers are having delusions—or are they?" and the premonitory "Congress should examine biological warfare tests," a topic that continues to interest and engage him at the national level to this day.

The Stanford Years

In parallel with his genetics research, Josh has had a long interest in formalizing the intellectual processes of scientific investigation and discovery and using computer models to represent these activities. He had used card-sorting techniques to help analyze the bacterial reproduction observations that led to his Nobel Prize years earlier, but on coming to Stanford in 1959 to found the Stanford Department of Genetics, his study of informatics took on new dimensions.



Joshua Lederberg demonstrating that his informatics work was "hands on," even in the era of Model 33 teletypes.

He took classes in programming and began a series of systematic investigations of computer use in biomedicine. He introduced one of the first LINC computers into the Stanford Medical School and became principal investigator of a school-wide computer resource called ACME (Advanced Computer for Medical Experimentation). By the mid-1960s, Lederberg was working on ways to systematize the enumeration of chemical structures, especially molecular isomers—families of molecules with the same chemical formulas but with different arrangements of the atoms. This early work would find expression later on in the DENDRAL project.[1]

> *Comments by Donald A. B. Lindberg, MD, Director, National Library of Medicine:* It's hard to remember when I really first met Josh, but I can remember vividly the first serious conversation I ever had with him. It was 1964, it was my first meeting as a Markel Scholar, the group was in the Ahwanee Lodge in Yosemite, and the speakers had been Jacob Bronowski, Le Dubridge, and Josh Lederberg. And I remember seeing Josh in the sun at a coffee break, sitting alone at a table, sort of lost in thought, and I decided that he probably wouldn't mind being interrupted. So I asked him what was he thinking about? It seemed like not too impertinent a question, although it probably was. And he said well, he was thinking about this: that if he had a chemical molecule in his mind, and he could see it clearly, what would be the next one that logically should be considered? Well, that struck me as kind of an amazing thing to be thinking about.

With the post-Sputnik start of the U.S. space program in earnest, one of Lederberg's new areas of attention was the exploration of outer space in search of life forms. He coined the term "exobiology" to describe this new field, dealing with the search for the origins of life, and established the Instrumentation Research Laboratory to explore the synthesis of analytic instrumentation and computers to devise ways of detecting molecular signatures of life-related processes in planetary environments.

This required inventing new instruments, such as fluorescence-activated cell sorting, and applying evolving instruments, such as gas chromatographs and mass spectrometers, in novel ways. It is also required inventing new ways to represent information about molecular structures such as proteins in the computer so that they could be analyzed and compared with each other. Work he performed in 1964 may have represented the first time that a protein sequence was stored in a computer for analysis.

Comments by Edward A. Feigenbaum, PhD, Former Co-investigator of DENDRAL *and* SUMEX-AIM, *Professor of Computer Science, Stanford University:* In 1964, when I first met Josh, he was writing his first large-scale computer program, in a dialect of ALGOL. The program was called DENDRAL, for dendritic algorithm, but the reference to dendrites was metaphoric. Josh was creating a complete system for the enumeration of acyclic chemical structures. The system included:

- A notation in which each unique chemical structure was represented by a unique symbolic expression in canonical form.
- A graph manipulation algorithm for generating structures from molecular formulas. The DENDRAL algorithm was capable of generating structures exhaustively and without redundancy.
- A computer program to put the DENDRAL algorithm to work.

In 1964, this was a landmark program not only for graph manipulation by computer but for application of computers to chemistry and life sciences, especially application of symbolic processing methods.

When I arrived at Stanford in January 1965, I discussed with Josh my goal of creating models of the thinking processes of scientists, especially the processes of empirical induction with which hypotheses and theories were inferred from data. What I needed was a specific task environment in which to study these issues concretely.

Josh suggested a task of inferring organic chemical structures from mass spectral data and enthusiastically entered a collaboration. He suggested a generate-and-test approach, in which the DENDRAL program he had done would serve as the "legal move generator" in the heuristic search for plausible structural hypotheses. The program we built, with the help of Bruce Buchanan, was called Heuristic DENDRAL. We extended its performance greatly in both breadth and depth when Carl Djerassi joined the collaboration. Josh brought this about, conveying to Carl the importance of the work. Josh's enthusiasm was infectious.

My sketch of Lederberg and DENDRAL is necessarily brief and leaves out many of his tangible contributions to our work in artificial intelligence and its applications. And it leaves out almost all his intangible contributions—the texture, if you will, of the day-to-day interactions with Josh. These interactions were full of wisdom and guidance on how to run substantial interdisciplinary projects of the kind we were doing in expert systems, insights on the nature of knowledge and cognitive processes in science, and creative ideas for getting us over today's hurdle and tomorrow's hill—in truth, a flood of creative ideas. If we had had the energy and resources, we could probably have done ten other projects with these ideas.

Working with Josh on our common problems in modeling expert thought in science and biomedicine gave me the privileged opportunity to witness a great intellect at work, applying superb analytic thinking skills with the enthusiasm of a young hacker for how much fun it was to program a computer to do these things. And it was not only fun, but it was important in the history of computer science.

The DENDRAL collaboration led Josh to devote a great part of his energies to computer science, including the underlying languages and tools necessary to express his ideas about molecular biology and genetics.

Comments by Donald A. B. Lindberg, MD, Director, National Library of Medicine: I was walking across the lovely campus of Stanford with him, following a morning in which I don't remember the real circumstances, but a whole lot of very nice research was reported at Stanford. I must confess, while I liked it all, I was sick and tired of hearing about LISP and list-processing languages. So as we hiked off to the commissary or the faculty club or wherever we ate, I said to Josh, "Josh, what would happen if some malignant deity just all of a sudden destroyed, eliminated, LISP. There isn't any more LISP. What would you do?" And he said "Oh, well, I guess, Don, I'd take FORTRAN and write LISP."

Lederberg insisted on participating first-hand in the development, testing, and use of the computing systems that fascinated him, no matter how cumbersome the early human–computer interfaces were.

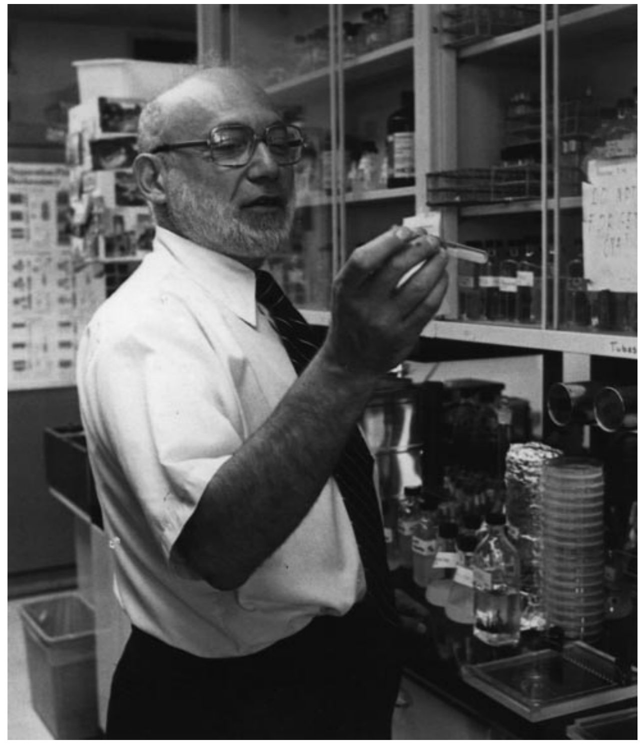
The breadth of Josh's biological, chemical, and computer science knowledge allowed him to serve as the bridge or glue among the interdisciplinary teams he assembled.

Comments by Carl Djerassi, PhD, Former Co-investigator of DENDRAL, *Professor of Chemistry, Stanford University:* When I came to Stanford in 1960, Josh was probably my closest friend, professionally and personally, and within about two years we started to collaborate, I as a chemist, he as a person who works in every field conceivable, including even chemistry. At that time he was interested in exobiology and unmanned flights to outer space and felt that whatever information we needed would get back through telemetry and that, in order to determine if there was any life in outer space, we had to look for *certain* evidence of life out there. I was involved at that time in the area of mass spectrometry and, precisely, materials that would be related to things in outer space. Steroids, of course, are an example of something that you are unlikely to find up there.

So we got together with Ed Feigenbaum, who had been working on artificial intelligence in general, to see what we could apply to chemistry, in particular the structural elucidation of organic compounds, which was precisely my area. And then they came to me and said, "Look, can we just simply try to simulate in a computer how Carl Djerassi thinks?" And that was really the beginning of a very fruitful collaboration for well over a dozen years in which a lot of other people were involved—Tom Rindfleisch, Bruce Buchanan, and others. And we published a lot of papers in which we did try to simulate a chemist's thinking in computer-intelligible language.

Note that was just the time when other chemists were working on simulating synthetic strategies. We did just the opposite. We talked about how one *establishes* structure rather than how one synthesizes it. That was unique at that time in the chemical field. What was most interesting, it stimulated us to do a lot of chemical research to fill gaps, intuitive gaps, in our own thinking, which the computer picked up because the computer program wasn't that intuitive yet. And that was really quite interesting, because chemists don't really wonder *how* they are thinking, they just do it. That really taught us an element of intellectual rigor. And it was a wonderful collaboration of people in medicine and computer science and chemistry. That was a large part of what at that time was called the DENDRAL program, which led to a number of other important algorithms and software projects that were eventually even commercialized.

Josh's contributions included very practical insights, such as designing a highly efficient table-based system for computing the atomic compositions of molecular fragments whose weights are precisely known from mass spectrometry.[2] As the DENDRAL programs evolved and made progress toward modeling the thought processes of chemists working to unravel molecular structures, Josh's interests returned to the broader possibilities of intelligent computer applications.



Lederberg has maintained an active laboratory to this day.

Anticipating the Role of the Internet in Science

The maturing of powerful time-shared computer systems in the 1960s and the beginnings of wide-area computer networks around 1970 under DARPA led Lederberg to imagine communities of scientists, and not just computer scientists, collaborating on research by means of these new tools. The SUMEX-AIM resource that he created in the early 1970s produced many systems that were the "seeds" of artificial intelligence research for years to come.

Comments by Thomas C. Rindfleisch, MS, Former Director of SUMEX-AIM *and CAMIS, Director, Lane Medical Library, Stanford University:* For me, one of the most important of Josh's many contributions to medical informatics was his use of digital technology—computing and communication—to help support scientific communications and collaboration. I came to Stanford in 1971 to join the DENDRAL project. At that time there was already a computer resource called ACME that Josh, Ed Feigenbaum, and Gio Wiederhold had developed in the mid-1960s. This resource was promoting the use of computers for applications in biomedicine and for the sharing of software and information among the researchers and clinicians. This system was innovative for its time. It had a very

Joshua Lederberg accepts the Morris Collen Award at the AMIA Symposium in November 1999.

friendly user interface to a time-sharing system and an incremental compiler that assisted in the development and debugging of the programs that were part of the ACME environment.

About a year after I came to Stanford, a small group of us, led by Josh Lederberg, synthesized the ideas from ACME and DENDRAL in the form of a proposal to the NIH called SUMEX-AIM. SUMEX-AIM contained two fairly outlandish ideas at the time. The first was to develop a national community of artificial intelligence in medicine application groups, and the second was to interconnect them using networking technology that was just then under development by the ARPA community in the form of the ARPANET.

I have to tell you that the defense of this proposal was very difficult in front of the peer reviewers at NIH. It took a year and a half with several rewrites and two very long and arduous site visits for us to get support for this project. Nevertheless, at the end of 1973 we did get support for SUMEX-AIM and the initial community of five activities—DENDRAL,[3] MYCIN,[4] a set of projects at Rutgers University,[5,6] PARRY,[7] and a protein-structure project.[8]

SUMEX-AIM was the first non-DOD research resource on the ARPANET. In fact, we were node number 56 among a total of 63 on the entire network in that current implementation. SUMEX lasted for about 19 years and was the nursery for more than 20 artificial-intelligence-in-medicine projects that were the best known in that era. Clearly, the SUMEX work was the result of lots of individual coordinated activities, but Josh's role was absolutely seminal in the formulation of the SUMEX community. Josh as a visionary, as a prototyper of technology, as a tester of that technology, and as a scientist and collaborator par excellence was absolutely indispensable.

Josh wrote a very important essay in 1977 that described these ideas. It was entitled "Digital Communications and the Conduct of Science: The New Literacy."[9] I'd like to quote a brief passage from this essay. In it he discussed how 'the convergence of economical digital communications with computer-aided tools would facilitate the interconnection of users separated both in time and space, and how this medium will increase the thoughtfulness of communication, return literacy in the efficient and precise use of language, and enhance scientific discourse in many ways."

As I scan the e-mail in my inbox today, I'm not sure how much progress we have made in the literate use of language, but it is absolutely clear with the current form of the Internet that digital communications have profoundly changed the way science is conducted.

The Move to Rockefeller University

Josh left Stanford in 1978 to become president of The Rockefeller University in New York. Despite the heavy burdens of administering one of the nation's leading research universities, he continued to explore the synthesis of computing and communications tools with scientific research on many fronts and to serve as a role model for young researchers in the emerging field of medical informatics.

Comments by Edward H. Shortliffe, MD, PhD, Former Principal Investigator of MYCIN, SUMEX-AIM, *and CAMIS, Professor of Medicine and of Computer Science, Stanford University:* When I first came to Stanford in the early 1970s, I was immediately aware of Josh Lederberg's remarkable influence there. I came with an interest in biomedicine and computer science and so was naturally drawn to folks such as Josh, Ed Feigenbaum, and Bruce Buchanan who had the dynamic DENDRAL project under way at that time. Josh, as principal investigator on the brand new SUMEX grant, was of course having a remarkable influence on the environment, not only because of his work on DENDRAL itself but also because he provided a kind of scientific leadership from which I as a young student learned a great deal.

I would say that Josh created an environment in which I was taught to understand computer science as *science*, and not simply the construction of artifacts. The scientific approach to computing, whereby one seeks to be sure that the results can generalize, where the experimentation is itself subjected to proof, and where there is an emphasis on methodologic development—these were key elements in the environment and the way in which I learned computer science because of the influence of people like Josh.

I was very aware of Josh's presence, even though I did not work with him personally. It was everywhere. He taught me about the notion of collaboration among scientists of widely varying backgrounds, and their ability to come together to create new and innovative ideas that are even greater than the sum of their parts. He showed me that there are still renaissance men. In spite of his background in molecular biology and genetics, he was making major contributions in computer science. He was creating an environment in which great computer science work could be done and was also aggressively experimenting with the newest technologies and beginning to understand how they might affect science in the future. I saw a brilliant mind in action, and I saw in him an interest in basically all things intellectual and an ability to innovate in very diverse fields. There was much to emulate in Josh Lederberg and his mentoring.

After he moved to Rockefeller University, Josh continued to have a big influence here at Stanford on our computing environment and on the scientific activities that we chose to undertake. He was always an advisor to the SUMEX-AIM project, even after his departure, as well as on the subsequent CAMIS resource. He has always been a source of personal mentoring for me. Sage advice is always available from Josh, and I have found that I can always turn for him for just the right incisive observation that can help me make a decision.

He has played an advisory role for years in companies and organizations that are dedicated to online information retrieval. I don't know anyone who is more adept at using online resources for finding information, sharing it with others, and coming up with ideas for how a vibrant environment ought to allow annotation and other similar techniques. Josh continues to be a master in that area.

To this day, Josh Lederberg brings a sensitivity and appreciation of computing and communications to his advisory work and to his influence at the national and international level. The informatics community owes him a great debt.

Josh was highly influential in guiding the Rockefeller University toward its role in the new era of basic biological science, genomic research, and the attendant revolution of clinical medicine. He became president emeritus of Rockefeller in 1990 and returned to his roots as a citizen scientist. His insights have been under increasing demand as he serves as an advisor to government, industry, scholarly foundations, and academia—and a continuing source of wise counsel to his friends and colleagues. He has been particularly active in applying scientific understanding to the direction of research, to public health, and to policy. As was mentioned, he has worked hard at educating the public and government about the risks of biological terrorism and warfare and advising on detection and counter measures.

Josh has continued his active research laboratory at Rockefeller, working on the genetics, biochemistry, and evolution of DNA; on the computer modeling of scientific reasoning; and on exobiology and the search for the origins of life.

Roles at the National Library of Medicine

A former member of the NLM's Board of Regents, Joshua Lederberg has been honored by the Library in their "Profiles in Science" program. The NLM is collecting and organizing Josh's voluminous papers as a contribution to the history of science. Most befitting Josh's career in medical informatics, this collection is available on the NLM's Web site.[10]

Josh's work in informatics is a lasting legacy, and he continues to pose challenges for the field.

Comments from an oral history interview with Joshua Lederberg, PhD, Former Principal Investigator of ACME, DENDRAL, *and* SUMEX-AIM, *President Emeritus, The Rockefeller University*: One principle I'd like to have my name attached to, Lederberg's Principle, is that machines will become really smart only when they can directly (1) read the literature and (2) spend some time living in the real world, where the survival of the fittest is what will determine who's out there. As long as we have to spoon feed them, datum by datum, they're going to be evolving in a very cumbersome and costly way indeed.

It is with pride, admiration, and affection that the American College of Medical Informatics salutes Joshua Lederberg as recipient of the 1999 Morris F. Collen Award.

The authors appreciate the technical assistance of Bindu Madhava, Bill Leonard, and Dan Masys in the preparation of this award description and the accompanying videotape. Thanks also to Barbara Hyde and the American Society for Microbiology for allowing them to quote from their extensive oral history of Dr. Lederberg, additional portions of which are available on the NLM's Web site, at http://www.profiles.nlm.nih.gov/BB/.

References ■

1. Lederberg J. How DENDRAL was conceived and born. Proceedings of the ACM Conference on the History of Medical Informatics. 1987:5–24.
2. Lederberg J. Computation of Molecular Formulas for Mass Spectrometry. San Francisco, Calif: Holden-Day, 1964.
3. Lindsay RK, Buchanan BG, Feigenbaum EA, Lederberg J. Applications of Artificial Intelligence for Organic Chemistry: The DENDRAL Project. New York: McGraw-Hill, 1980.
4. Buchanan BG, Shortliffe EH. Rule-based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project. Reading, Mass: Addison-Wesley, 1984.
5. Weiss S, Kulikowski CA, Safir A. Glaucoma consultation by computer. Comput Biol Med. 1978;8(1):25–40.
6. Kulikowski CA. Expert medical consultation systems. J Med Syst. 1983;7(3):229–34.
7. Colby KM. Clinical implications of a simulation model of paranoid processes. Arch Gen Psychiatry. 1976;33(7):854–7.
8. Engelmore R, Nii H. A knowledge-based system for the interpretation of protein x-ray crystallographic data. Stanford University Department of Computer Science, 1977. Report CS-TR-77-589. Available at: http://elib.stanford.edu:80/Dienst/UI/2.0/Describe/stanford.cs%2fCS-TR-77-589?abstract=. Accessed Jan 1, 2000.
9. Lederberg J. Digital communications and the conduct of science: the new literacy. Proc IEEE. 1978;66(11):1314–9.
10. National Library of Medicine. Profiles in Science. NLM, 1999. Available at: http://www.profiles.nlm.nih.gov. Accessed Jan 1, 2000.