

The Basement Interviews

Freeing the scientific literature

Harold Varmus, Nobel laureate, former director of the US National Institutes of Health, and co-founder of open access publisher Public Library of Science, talks to Richard Poynder.

Published on June 5th 2006¹

Harold Varmus was born in 1939, on the south shore of Long Island, New York, a product, as he put it in an autobiographical note he wrote years later,² of "the early twentieth century emigration of Eastern European Jewry to New York City and its environs."

His paternal grandfather, Jacob Varmus, had arrived in the New World from a small village near Warsaw just after the turn of the century, settling first as a farmer in Newburgh, New York, and subsequently working as a hatter in Newark, New Jersey. His paternal grandmother, Eleanor, was a victim of the influenza epidemic of 1918, dying when his father was eleven. Varmus' maternal grandparents — Harry and Regina Barasch — came originally from farming villages around Linz, Austria. Settling in Freeport, New York, they ran a children's clothing store.

Three years before Varmus' birth, his parents also settled in Freeport, where his father — who had trained as a doctor — established a general medical practice. From there Varmus' mother, Beatrice, commuted to a social services job in New York City.

A lake with alligators

When the United States entered the War, Varmus' father was assigned to an air force hospital near Winter Park, Florida, where the family resided until 1946. "My first memories," Varmus recalled later, "were to be of long beaches, and bass fishing on a lake with alligators." It was in Florida that Varmus' only sibling, Ellen Jane, was born.

After the family returned to New York, Varmus attended local public schools in Freeport; schools, he later characterised as being "dominated by athletics and rarely inspiring intellectually".

In 1957, after graduating from Freeport High School, Varmus entered Amherst College³ to prepare for medical school. His experience at Amherst, however

¹ This interview took place on 4th October 2005.

² <http://nobelprize.org/medicine/laureates/1989/varmus-autobio.html>.

³ Amherst College is an independent liberal arts college in Amherst, Massachusetts. It is the third oldest college in Massachusetts. Founded in 1821, Amherst was intended to be a successor to both Williams College, which was then struggling to remain open, and Amherst Academy, a secondary school which educated, among others, Emily Dickinson.

led to a change of plan. As he put it: "The evident intensity and pleasure of academic life there challenged my presumptions about my future as a physician, and my course of study drifted from science to philosophy and finally to English literature."

He also became active in politics and journalism, and served as editor of the Amherst college newspaper. In 1961, after receiving an English degree — graduating *magna cum laude*⁴ — Varmus went to Harvard, courtesy of a Woodrow Wilson Fellowship.

Midway through his first year as an English literature postgraduate, however, Varmus dreamed that he had become an English professor and that he had missed a day of lecturing due to illness. Rather than being disappointed, however, his students were jubilant, pleased that their class had been cancelled. When he woke, it occurred to Varmus that if he were a doctor no one would be happy if he didn't show up for work — a thought that inspired him to return to his medical studies.

Varmus applied to study medicine at Harvard twice, but was rejected on both occasions. Concluding that he was not yet mature enough, the admissions committee recommended he first do two years army service. Undeterred, however, Varmus enrolled at Columbia College of Physicians and Surgeons (P&S).⁵

Happy accident

Varmus began medical school with strong interests in psychiatry and international health. After serving an apprenticeship in a mission hospital in Bareilly, India, however, he drifted towards basic medical sciences.

In preparation for a career in academic medicine, Varmus worked as a medical house officer at Columbia-Presbyterian Hospital.⁶ In 1968, however, the Vietnam War interrupted his plans for an academic career.

In part to avoid the draft,⁷ Varmus applied for a research training post at the National Institutes of Health (NIH),⁸ where he was accepted for a clinical associate position in the laboratory of molecular biologist Ira Pastan.⁹

<http://www.amherst.edu>

⁴ "with great honor". http://en.wikipedia.org/wiki/Magna_cum_laude.

⁵ <http://www.cumc.columbia.edu/dept/ps>.

⁶ <http://cpmcnet.cpmc.columbia.edu>.

⁷ In a recent profile of Varmus in Wired, Jamie Shreeve commented: "To avoid the Vietnam draft [Varmus] applied for and won a position as a 'yellow beret' in the NIH's Public Health Service." *Free Radical*, Jamie Shreeve, Wired, June 2006, <http://www.wired.com/wired/archive/14.06/varmus.html>. See also *P&S Journal*: Fall 1994, Vol.14, No.3 http://cumc.columbia.edu/news/journal/journal-archives/jour_v14n3_0045.html.

⁸ <http://www.nih.gov>.

⁹ Ira Pastan is chief of the Laboratory of Molecular Biology, Division of Basic Sciences. He established the lab in 1970. <http://ccr.cancer.gov/Staff/staff.asp?profileid=5782>.

At the time of Varmus' interview, Pastan was working on the hormonal influences on the thyroid gland. By the time Varmus turned up at the lab, however, Pastan's focus had switched to the study of the genetics of E.coli — an area with which Varmus was unfamiliar.

While this dispirited Varmus, it turned out to be a happy accident, since the change symbolised a revolution that was taking place in medical research. As James Fallows explains in a 1999 *New Yorker*¹⁰ profile of Varmus, the thrust of medical research was at that time beginning to shift "from the operations of organs, or the disorders of whole organisms, like human beings or mice, to the mechanics of cells, and within them, specific genes."

In other words, rather than trying to target individual illnesses, researchers were adopting a new gene-level approach to medical research. As Varmus explained to journalist Susan Stamberg, in a US National Public Radio interview in 1999,¹¹ the future of medical research now lies in "the development of a notion of the gene as a physical entity that we can understand, manipulate, dissect and use to advance the great themes in medicine."

As luck would have it, Varmus' new post at NIH put him at the centre of this new approach just fifteen years after Watson and Crick had discovered the structure of DNA.¹² And by taking advantage of the evening courses offered to incipient physician-scientists at NIH, Varmus was able to undertake further postdoctoral training in molecular biology. He later developed an interest in tumour viruses.¹³

In 1969 Varmus moved to the University of California San Francisco (UCSF),¹⁴ where he began collaborating with microbiologist and immunologist Michael Bishop¹⁵ — conducting research on bacterial gene expression¹⁶ and tumour virology. Specifically, he studied the behaviour of retroviruses,¹⁷

¹⁰ *The Political Scientist*, James Fallows, *New Yorker*, June 7th 1999

http://www.mskcc.org/mskcc/shared/graphics/Varmus/New_Yorker_Profile_1999.pdf

¹¹ Interview: Dr. Harold Varmus of the NIH discusses the best advances in his field for this century, National Public Radio, 29th December 1999.

<http://www.npr.org/templates/story/story.php?storyId=1068480>.

¹² In an article published in *Nature* on April 25, 1953, James D Watson and Francis Crick described the structure of the chemical deoxyribonucleic acid, or DNA, which contains the genetic instructions specifying the biological development of all cellular forms of life (and most viruses)

<http://www.nature.com/nature/dna50/archive.html>

¹³ In his profile of Varmus in *Wired*, Shreeve commented: "Once [at NIH Varmus] gravitated toward the study of tumour viruses, partly because his mother had recently been diagnosed with breast cancer."

¹⁴ <http://www.ucsf.edu>.

¹⁵ Michael Bishop is an American immunologist and microbiologist. He currently serves as an active faculty member and chancellor at the University of California, San Francisco.

<http://nobelprize.org/medicine/laureates/1989/bishop-autobio.html>;

http://en.wikipedia.org/wiki/J._Michael_Bishop

¹⁶ Gene expression is the multi-step process by which a gene's information is converted into the structures and functions of a cell. http://en.wikipedia.org/wiki/Gene_expression.

¹⁷ A retrovirus is a type of virus that has RNA instead of DNA as its genetic material. It uses an enzyme called [reverse transcriptase](#) to become part of the host cells' DNA. This allows many copies of the virus to be made in the host cells. The virus that causes AIDS, the human immunodeficiency virus (HIV), is a type of retrovirus. <http://en.wikipedia.org/wiki/Retrovirus>.

including aspects of their unusual life cycle, the nature and origin of their transforming genes, and their potential to cause genetic change, not least their ability to cause cancer.¹⁸

At the time, many scientists thought that retroviruses caused cancer by injecting their genes into the host's own genome. In 1975, however, Varmus and Bishop found that these viral genes had in fact been "captured" from the host animals in the first place.

Their discovery demonstrated that cancer genes (oncogenes¹⁹) can evolve from normal cellular genes, called proto-oncogenes. In short, retroviruses can transform normal cells into cancerous ones.²⁰

Nobel Prize

This new insight meant that many of the cellular genes involved in cancer could be isolated. "It's been known for a long time that cancer was in some sense a genetic disease," Varmus explained to the *Boston Globe* in 1989.²¹ "The importance of our findings is that one can identify explicitly the genes that play a role in cancer."

The groundbreaking work done by Varmus and Bishop was recognised as a major contribution to current understanding, and enabled the scientific community to conduct an aggressive and successful search for the genetic origins of cancer. In 1989, therefore, they were jointly awarded a Nobel Prize "for their discovery of the cellular origin of retroviral oncogenes".²²

As a Nobel Laureate, Varmus began to feel a growing expectation that he become an ambassador for science, not least from former UCSF colleague Marc Kirschner²³ — who recruited Varmus to serve on advisory groups about the future of research. In 1992 Varmus also joined a group of scientists supporting the Clinton-Gore ticket,²⁴ and the following year he was confirmed as the new director of NIH.

Having never chaired a college department, or indeed run anything bigger than a 36-person laboratory, Varmus faced a significant challenge in his new

¹⁸ Genetic changes and the resulting changes in cell behaviour are at the heart of cancer biology.

¹⁹ An oncogene is a gene that can cause a cell to develop into a tumour cell, possibly into a cancerous one.

²⁰ In his profile of Varmus in *Wired*, Shreeve described it thus: "Their work showed that certain cancer-inducing viruses do their damage by capturing normal genes in the host's own cells and turning them into cancer causers. This discovery led to the isolation of these so-called oncogenes."

²¹ *Two win Nobel for cancer research*, *Boston Globe*, October 10th 1989

<http://www.boston.com/globe/search/stories/nobel/1989/1989m.html>

²² <http://nobelprize.org/medicine/laureates/1989>.

²³ Kirschner had gone on to become the founding chair of the Department of Systems Biology at Harvard. <http://sysbio.med.harvard.edu/faculty/kirschner>.

²⁴ The US presidential election of 1992 featured a three-way battle between Republican George Bush, the incumbent President; Democrat Bill Clinton, the governor of Arkansas; and independent candidate Ross Perot, a Texas businessman. Clinton won election by a wide margin in the US Electoral College, despite receiving only 43% of the popular vote.

http://en.wikipedia.org/wiki/U.S._presidential_election,_1992.

job. He was now in charge of 16,000 people and a large bureaucratic organisation composed of 24 quasi-independent internal institutes. Being director of NIH also required winning financial support from Congress.

In the event, he proved highly successful, getting NIH funding up more substantially than any other category of federal spending — from \$11bn to \$16bn. He also succeeded in persuading Congress of the importance of basic research — particularly studying cellular-level functions — rather than simply granting money for crusades against high-profile diseases like prostate and breast cancer.

At NIH, argues Fallows, Varmus "achieved a series of political victories that will affect scientific policy for many years to come and, at least by implication, may change the entire understanding of health, disease, and the limits of the human life span."

When, in June 1999, Fallows asked him about the secrets of his political success Varmus replied by e-mail: "I have the sense that you wish my life had more drama. I don't. As I have said before, my strategy has been to try to minimise it."

Controversy

Ironically at the very time he e-mailed Fallows, Varmus had just embarked on a project that would maximise the drama in his life, and spark a long-running controversy that has still to run its course. In short, his new project was to prove the catalyst for the creation of the Open Access Movement.

A few months earlier Varmus had been sitting in a San Francisco coffee house with Pat Brown,²⁵ a Stanford University biochemist who had at one time collaborated with the Varmus Lab as a trainee, and remained a close colleague. Brown began telling Varmus about arXiv.org, the web site where physicists place pre-published versions of their papers for review,²⁶ usually prior to publishing them in prestigious journals. This allowed physicists to speed up the research process, and maximise the impact of research papers by making them freely available to everyone over the Internet. Brown posed the question to Varmus: could not a similar initiative in biology provide the same benefits?

The question resonated for Varmus, and he spent the next few months pondering on it. The end result was E-BIOMED: "a proposal for electronic publications in the biomedical sciences",²⁷ which Varmus distributed to the research community in May 1999. In effect, E-BIOMED envisaged a fundamental and very radical change to the way biomedical research is communicated.

²⁵ Patrick Brown is a Stanford geneticist who became one of the co-founders of Public Library of Science (PLOS). <http://www.plos.org/about/board.html#Brown>.

²⁶ <http://www.arxiv.org>.

²⁷ The original proposal for E-BIOMED (Draft and Addendum) can be found here: <http://www.nih.gov/about/director/pubmedcentral/ebiomedarch.htm>.

The problem, however, was that Varmus' proposal implied a diminished role for science publishers, and the likelihood, therefore, that scholarly publishing would become a somewhat less profitable business. As such, E-BIOMED was greeted with a firestorm of protest and outrage — a response that NIH colleagues felt Varmus should have anticipated. "Varmus wrote the proposal himself and just sent it out to a bunch of people as a rough draft," comments one still puzzled NIH insider. "For us it was like 'Oh, my God: he clearly doesn't know what he is getting into'.

"Sure enough," he adds, "the response from publishers was explosive. It caused all kinds of problems."

As Varmus now concedes, he certainly *didn't* realise what he was getting into. Consequently, he was startled by the vehemence of the reaction from publishers, and surprised and angry when he discovered that, instead of engaging with him directly on the issue, many publishers had gone straight to Congress and complained directly to his appropriators.

Over the next few months there was a heated public (and private) debate. In the end, however, the strength of opposition from publishers was so great that when, eight months later, E-BIOMED was launched (re-branded as PubMed Central²⁸) it was a pale shadow of the revolutionary new "electronic publications" system that Varmus had envisioned.

Significantly, he had had to concede that publishers would have the final say on whether papers they published were placed in PubMed Central. And publisher response was derisory: even today, papers from only 251 of a possible 6,000 biomedical journals are made available on PubMed Central, and the content of some of these journals is only released after an embargo of between six and twelve months.

In retrospect, Varmus agrees that he was naïve not to have anticipated the furore. "I must have known that I was not going to be at NIH for much longer," he joked to *New Scientist* in 2003, "because this caused a tremendous political argument: what the hell was I trying to do to destroy the publication industry."²⁹

Indeed, Varmus left NIH within months of publishing the proposal, taking over as president of the Memorial Sloan-Kettering Cancer Centre.³⁰ Ironically, he says, his last public act at NIH was to sign the press release announcing the launch of PubMed Central.

²⁸ <http://www.pubmedcentral.nih.gov>

²⁹ *Freedom Fighter*, New Scientist, 1st November 2003 <http://lists.essential.org/pipermail/random-bits/2003-November/001141.html>

³⁰ Sloan-Kettering is the world's oldest and largest private institution devoted to patient care, education, and research into cancer, based in New York. <http://www.mskcc.org/mskcc/html/44.cfm>

Public Library of Science

By now, however, Varmus was firmly committed to the cause of Open Access (OA) — the more so, it seems, in light of the virulence of publisher opposition to E-BIOMED. In November 2000, therefore, Varmus co-founded — with Pat Brown and Mike Eisen³¹ — the Public Library of Science (PLoS).

Initially an advocacy group, PLoS was founded on the principle that if publishers were not prepared to act in the best interests of science, then the research community would have to twist their arms. And the first act of the founders was to invite fellow scientists to sign an open letter in which they pledged to boycott any journal publisher that did not make the papers it published freely available online within at least six months of publication.

PLoS attracted nearly 34,000 signatures from scientists in 180 countries; but while a small handful of publishers complied, most blithely ignored the PLoS letter. Worse, most of the scientist signatories were happy to disregard their own ultimatum, and continued publishing in the very journals that had turned a deaf ear to their request.

Unprepared to give up the cause of freeing the scientific literature, in 2001 Varmus and his colleagues reinvented PLoS as a not-for-profit open access publisher — inviting scientists to pay to publish their research in new open access journals, on the basis that in doing so they could ensure that their research was made freely available on the Web from the date of publication. A number of PLoS journals were subsequently launched, and more are planned.

To Varmus' gratification, the journals have attracted some first-rate submissions, and quickly acquired a reputation as high-quality publications.³²

What is less certain, however, is whether the author-pays publishing model they utilise is sustainable in the long term.³³

In many ways, Varmus has taken his role as an ambassador for science to its logical conclusion. After all, if the effectiveness of scientific research can be

³¹ Michael B. Eisen is a computational and evolutionary biologist at Ernest Orlando Lawrence Berkeley National Laboratory and the University of California at Berkeley. <http://rana.lbl.gov>

³² In his profile of Varmus in *Wired*, Shreeve comments: "research published in the flagship journals *PLoS Biology* and *PLoS Medicine* often finds its way to *The New York Times* or the BBC. Last June, less than two years after the first issue of *PLoS Biology* went online, [Thomson Scientific](#), a firm that tracks citation rates, assessed the journal [as having] an 'impact factor' higher than such established journals as [Biological Reviews](#) and [Proceedings of the National Academy of Sciences](#). Indeed, in a phenomenally short time, it has become the most cited journal in general biology."

³³ Wisely, PLoS set its article processing charge at \$1,500 from day one. OA publisher BioMed Central, by contrast, initially set its fee at \$525. Subsequently realising that this was far too low, in July 2005 BMC more than doubled the charge — to \$1,400 — sparking a rebellion by many of its editors. It also helped that PLoS majored on creating high impact journals able to compete with the world's best biomedical journals. Nevertheless, as Varmus concedes in the interview, PLoS too has had to re-calculate its estimate of when it will achieve profitability. Sceptics question whether it ever can.

maximised by disseminating it in the most efficient way possible, then any ambassador worth his salt would surely do everything he could to make that happen?

Varmus' ambassadorial activities have not been confined to the US either. In March 2004, for instance, he appeared as a witness for the UK Select Committee enquiry into scientific publishing,³⁴ which subsequently recommended that the UK government mandate all publicly-funded researchers to make their papers available on the Web.³⁵

Undoubtedly, Varmus will also have been involved in the behind-the-scenes lobbying that, in May 2005, saw the introduction of the NIH OA Policy,³⁶ and the introduction into the US Senate in December 2005 of the CURES Act of 2005.³⁷ No doubt he played a part in lobbying for the recently announced Federal Research Public Access Act of 2006 (FRPAA) too.³⁸

Naivety

What sort of man is Varmus? Describing him as a "lean, energetic, and intense presence", Fallows pointed out that — for someone in such a public position — Varmus managed to maintain a surprisingly low profile at NIH. Indeed, his colleagues at the time often referred to him as "the invisible administrator."

So successfully did he maintain public invisibility in fact that when, in 1996, Varmus gave the commencement address at Harvard, the university newspaper, *Crimson*, reported widespread disappointment that an apparent nobody had been invited to speak.³⁹

Why then seek to play a leading role in such a controversial — and increasingly high profile — movement as Open Access? "I believe that science is one of those activities that improves the state of the world," replies Varmus, "and once you realise how important publication is in the series of acts that constitutes the doing of science, and once you understand the incredible transformation of that publication process that the Internet, and software, and the whole digital world, now promises it is hard not to be pretty

³⁴ http://www.parliament.uk/parliamentary_committees/science_and_technology_committee/scitech111203a.cfm

³⁵ To date, however, the recommendation has yet to bear fruit. See footnote 98.

³⁶ The NIH policy, formally known as the "Policy on Enhancing Public Access to Archived Publications Resulting from NIH-Funded Research", came into effect on May 2nd 2005. See also footnote 106, and: <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-05-022.html>.

³⁷ See footnote 109, and <http://www.earlham.edu/~peters/fof/newsletter/01-02-06.htm#cures>.

³⁸ The FRPAA was introduced on May 2nd 2006, by Senators John Cornyn (R-TX) and Joe Lieberman (D-CT). See footnote 109, and: <http://www.arl.org/sparc/resources/frpaa.html>.

³⁹ As *Crimson* put it, "Despite Varmus' reputation as a brilliant researcher and the advances he has made in the field of medicine, many seniors expressed regret that a better-known speaker was not chosen. 'Compared to previous years, this is going to be a big letdown. Especially in a presidential year, you'd think that we'd get one of the candidates,' said Meloney L. McGuire '96, echoing the sentiments of many of her classmates." <http://www.thecrimson.com/article.aspx?ref=104793>

passionate about trying to make that part of the scientific universe work more effectively."

How significant a role has Varmus played in the Open Access Movement then? Critics often claim that his naivety has hindered more than helped. For instance, they argue, the E-BIOMED proposal only encouraged publishers to dig their heels in and resist change.

Varmus' supporters, however, contend that, like the music and movie companies, science publishers would inevitably have done everything they could to resist the impact of the Internet — since the new distribution models it permits pose a significant threat to their profits. Moreover, they add, as a Nobel Laureate and former NIH director, Varmus has provided the movement with an authority and credibility that it would otherwise have seriously lacked.

Importantly, comments the NIH insider, once alerted to the issue Varmus was immediately able to appreciate what was required, and then took personal responsibility for the task of making it happen. "Harold deserves credit for seeing that this was a real problem that needed to be solved, because most institute directors, and most scientists, just don't have the insight."

Certainly Varmus is now widely viewed as a key figure in the OA Movement. In 2004 he and PLoS co-founders Eisen and Brown were given a *Wired* Rave award⁴⁰ "For cracking the spine of the science cartel". And in a recent profile of him, the magazine described Varmus as being "the prophet of open access." A prophet, moreover, who has become locked into a real-life struggle "that has turned this icon of the scientific establishment into a powerful subversive."⁴¹

For all that, Varmus has a somewhat parochial view about Open Access, and makes no claims to understand the wider free knowledge movement, or even to particularly care about the role of Open Access within that wider movement, although he does support the principles of Creative Commons, whose licences PLoS utilises for its journals.

This disinterest in the bigger picture is a little surprising perhaps, because there is undoubtedly a commonality in origin, and in purpose, between the Open Access Movement and the various other free and open movements. If nothing else, they are all creative responses to the possibilities of the Internet, not least its ability to enable much greater collaboration and information sharing.

What is clear is that, in the same way that large and complex software projects now cry out for an open source approach to development, gene-level medical research requires an increasing degree of openness to be effective. The greater need for "open data" was graphically demonstrated, after all, by the huge efforts that NIH put into competing with privately-held Celera⁴² to

⁴⁰ <http://www.wired.com/wired/archive/12.04/rave.html?pg=14>

⁴¹ *Free Radical*, Wired, June 2006. <http://www.wired.com/wired/archive/14.06/varmus.html>.

⁴² <http://www.celera.com>

sequence the human genome (by means of the Human Genome Project, or HGP).⁴³ The fear was that vital biological information would be privatised if Celera became the only game in town, an outcome that would have obstructed scientific research.⁴⁴

In an online environment research papers will also increasingly be viewed as inseparable from the data on which their conclusions are based. And if both the data and the evaluation of that data need to be more widely and freely accessible, then Open Access will have to be treated as part of a larger and broader development.

Likewise, the new-style licensing models being developed by organisations like Creative Commons will become an ever more important component of the infrastructure that will be necessary if the various free and open movements are to flourish; and many of the new software tools being developed to support these movements will undoubtedly need to be based on Open Source Software.

In other words, Open Access, the HGP, Open Data, the Free and Open Source Software movements, Creative Commons, and the many other free and open movements, are all logical adaptations to the realities of today's increasingly complex knowledge environment. More importantly, they all face a common threat: the threat that, rather than being made more freely available, information could be increasingly privatised.

Excessively draconian use of intellectual property laws, for instance, could significantly hamper these movements, and hinder the improved collaboration, more rapid development, and greater creativity — be it artistic or scientific — that today's digital world promises. The ability of scholarly publishers to resist OA, after all, is based on their insistence that, as a condition of publication, they are given copyright in the papers they publish. By working together to resist ever more aggressive use of intellectual property the various free and open movements could surely achieve their separate aims more effectively.

Interest in the arts

But it would be wrong to suggest that Varmus' vision of the world is in any way narrow. Despite his evident passion for science, Varmus retains a strong interest in the arts. In the talk he gave at his Nobel Prize ceremony, for instance, he included lines from *Beowulf*,⁴⁵ which he recited in Anglo-Saxon. He has also written film reviews for the *New York Times*, and invested in

⁴³ The Human Genome Project was coordinated by US Department of Energy, the NIH and, the UK-based Wellcome Trust http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml (See footnote 116)

⁴⁴ In the event it was a draw. However, as a result of the success of the HGP, in May 2005 Celera announced that it was closing its subscription-based database service and releasing its genomic data on humans, rats, and mice into the public domain.

⁴⁵ *Beowulf* is one of the oldest surviving epic poems in what is identifiable as an early form of the English language. The precise date of the manuscript is debated, but most estimates place it close to 1000. <http://en.wikipedia.org/wiki/Beowulf>

Broadway plays. And when he visits foreign cities he often proposes meeting up with associates and friends in art galleries.

Varmus is known for his casual clothes — often khakis and rumbled open-necked shirts — and is a keen sportsman who loves outdoor sports, particularly bicycling, running, backpacking, skiing, and fishing.⁴⁶

In 1969, he married Constance Louise Casey, then a reporter for *Congressional Quarterly* in her hometown of Washington, DC. They have two sons: Jacob and Christopher. Varmus' younger sister now works as a genetic counsellor in California, and his brother-in-law is the novelist John Casey.

##

Although this interview was conducted by telephone, I had seen Varmus in the flesh when he gave evidence to the UK Science & Technology Select Committee. A tall, slim man — with, as Jamie Shreeve put it in his *Wired* profile of Varmus, "sparse, reddish brown hair" — he sat in a slightly hunched position throughout, leaning forward on to the desk as he explained the merits of Open Access to British politicians.

Indeed, sitting next to BioMed Central founder Vitek Tracz, Varmus presented a striking contrast. A small restless man, Tracz spoke quickly, and at some length, causing the chairman at one point to ask him "to keep the answers a bit shorter and answer the question". Varmus, by contrast, delivered his views briefly and precisely.

And where Tracz was full of nervous energy, Varmus exuded a quiet confidence and unflappability. He also revealed a consummate skill for flattering politicians, a skill no doubt honed at NIH. His final words to the Select Committee were: "It has been a pleasure for me to see this high-minded conversation. I appreciate it greatly."⁴⁷

While entirely courteous, Varmus was less ingratiating with me. When I e-mailed him to ask for an interview I heard nothing for several weeks. Eventually his secretary invited me to suggest some dates, and asked how much time I needed for the interview. I requested two to three hours: I was offered 60 to 90 minutes.

The week before the interview was scheduled I was asked to send a list of questions so that Varmus could review them over the weekend. On the day of the interview, however, he declined to answer any questions not directly related to Open Access, arguing that in the time he had allotted for our conversation he was not able to address any other topic.

⁴⁶ As Shreeve put it in his profile of Varmus in *Wired*, "At 66, he is an avid cyclist, as well as a jogger, hiker, and rower, and he loves to talk about his physical pursuits."

⁴⁷ Uncorrected Transcript Of Oral Evidence, Monday 8th March 2004

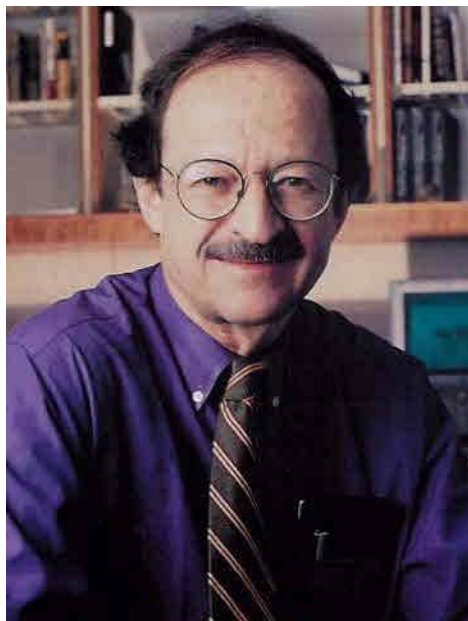
<http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/uc399-ii/uc39902.htm>.

And even though we finished our discussion before our scheduled time was up, Varmus still declined to answer questions on any topic other than Open Access. He suggested that if I wanted biographical information I could use what was already in the public domain.

He also appeared unusually controlling about the text of the interview, asking several times to have sight of it before publication.⁴⁸ At one point in the interview he complained that I was talking over him.

All in all, when I put down the phone I knew very little more about the kind of person Varmus is than before I picked it up — other perhaps than that he seems a very focused person, and apparently keen to protect his privacy.

So let me leave the last words to the NIH insider, who spoke to me on condition of anonymity. "Harold Varmus", he said, "is a visionary who has very little feel for how people are going to react to what he says and does. But he isn't generally concerned about that. He is also a man without guile."



The interview begins ...

RP: You are a proponent of Open Access. What is Open Access?

HV: It is important to remember that Open Access is a bit of moving target. When I first started thinking about the problem of how you make research literature more accessible to everybody, for instance, there really wasn't a coherent movement, and a definition was not agreed until the meeting we held at the Howard Hughes Medical Institute in April 2003.

⁴⁸ When however I later e-mailed the text — to both Varmus and his secretary — I received no feedback at all.

*RP: A definition that was formulated in The Bethesda Statement on Open Access Publishing?*⁴⁹

HV: That's right. The definition of Open Access that most of us use today was formulated there, and it is the definition that I give homage to.⁵⁰ Essentially it says that open access research articles are articles that are freely available in a digital form to anyone from the time of publication, that copyright of the articles remains with the author, and that they are made available on a fair-use basis — with appropriate attribution — by means of a Creative Commons licence.⁵¹

RP: Why is Open Access a good thing?

HV: To answer that we need to step back and look at the history of what scientists have tried to achieve when publishing their work. Essentially, the goal of publication in science is to get the information out to as many people as possible, and to do so as efficiently as possible. Today, the Internet has dramatically changed the way we can do that — so the public good here is in making scientific information (which is not exclusively, but largely, generated by public money) more accessible for others to use. Open Access is a good thing, therefore, because it can better serve the purpose of publishing scientific information.

RP: In other words, the Internet has made it possible to provide better access to scientific papers, to more people; and Open Access implies doing so without charging people for the information?

HV: Exactly. And the reason why Open Access is controversial is because it requires some big changes in operating principles. After all, when you have a system that has worked reasonably well people are loath to abandon it, especially where working well also means generating large revenues for publishers.

RP: You say that the definition of Open Access that you give homage to was formulated at a meeting held at the Howard Hughes Medical Institute in 2003. You

⁴⁹ <http://www.earlham.edu/~peters/fos/bethesda.htm>

⁵⁰ The Bethesda Statement says that an Open Access Publication (i.e. a research paper) is one that meets the following two conditions:

*The author(s) and copyright holder(s) grant(s) to all users a free, irrevocable, worldwide, perpetual right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship, as well as the right to make small numbers of printed copies for their personal use.

*A complete version of the work and all supplemental materials, including a copy of the permission as stated above, in a suitable standard electronic format is deposited immediately upon initial publication in at least one online repository that is supported by an academic institution, scholarly society, government agency, or other well-established organisation that seeks to enable Open Access, unrestricted distribution, interoperability, and long-term archiving (for the biomedical sciences, PubMed Central is such a repository).

⁵¹ Creative Commons (CC) licences enable copyright holders to grant some of their rights to the public while retaining others through a variety of licensing and contract schemes including dedication to the public domain or open content licensing terms. The intention is to avoid the problems current copyright laws create for the sharing of information. http://en.wikipedia.org/wiki/Creative_Commons.

had clearly been thinking about it prior to that: when did you first come across the concept of Open Access?

HV: My conversion from being an innocent person who went along with the system as it existed, to someone who became passionate about making better use of the Internet for scientific publishing occurred one specific day when I was in San Francisco.

RP: *When was that?*

HV: I think it was around December 1998. Anyway, I was in San Francisco for another purpose and — as I frequently did — I met with Pat Brown, who has been a scientific colleague for many years.⁵² We got together in a coffee shop in Parnassus and Carl [laughs], where Pat started to tell me about the physics preprint archive⁵³ that Paul Ginsparg⁵⁴ had developed at Los Alamos.⁵⁵

Pat said that this had inspired him to think about how something similar could be developed to help those working in biology rather than physics. So we sat and talked about the cultural differences between the community of physicists and the community of biomedical research — which, by the way, are very, very different. We also talked about how we might use the tools of digital communication to make the work of biomedical science proceed faster and more efficiently, and in a richer fashion.

RP: *And you personally became involved in developing the tools to enable that?*

HV: At that time I was director of the National Institutes of Health,⁵⁶ and so had probably more influence than I really understood. Anyway, as a result of my conversation with Pat I spent a lot of time over the next couple of months thinking about what he had said, looking at Ginsparg's site,⁵⁷ and imagining how the National Library of Medicine⁵⁸ could help to foster greater dissemination of published works. In doing so I thought also about some of the principles by which publishing is

⁵² Patrick Brown is a biochemist at the Howard Hughes Medical Institute and Stanford University School of Medicine, where he is currently a professor in the department of biochemistry.

<http://cmgm.stanford.edu/pbrown>

⁵³ arXiv.org (formerly known as the LANL preprint archive) is an archive for electronic preprints of scientific papers in the fields of physics, mathematics, computer science and biology which can be accessed via the internet. In many fields of mathematics and physics, almost all scientific papers are placed on the arXiv, with some work, including even very influential papers, remaining purely as e-prints and not published in peer-reviewed journals. Originally hosted at the Los Alamos National Laboratory, arXiv is now hosted and operated by Cornell University and mirrored worldwide. The archive was originally developed by Paul Ginsparg in 1991. As of February 2006, arXiv.org contains about 355,000 e-prints, with three to four thousand new e-prints being added every month.

www.arxiv.org

⁵⁴ <http://www.physics.cornell.edu/profpages/Ginsparg.htm>

⁵⁵ The Los Alamos National Laboratory. <http://www.lanl.gov>

⁵⁶ The National Institutes of Health (NIH), a part of the US Department of Health and Human Services, is the primary Federal agency for conducting and supporting medical research. <http://www.nih.gov>

⁵⁷ <http://www.arxiv.org>

⁵⁸ The National Library of Medicine (NLM), on the campus of the National Institutes of Health in Bethesda, Maryland, is the world's largest medical library. The Library collects materials and provides information and research services in all areas of biomedicine and health care. <http://www.nlm.nih.gov>

conducted, and I had conversations with David Lipman who was then, and still is, the director of the National Centre of Biotechnology Information.⁵⁹

RP: The history of how Open Access became a powerful movement for change is fascinating. When I spoke to open access publisher Vitek Tracz⁶⁰ he said that the idea of NIH helping to promote Open Access came initially from him: that he had suggested the idea to David Lipman,⁶¹ who then persuaded you to take action.

HV: Vitek may indeed have talked to David — I don't really remember because I didn't know Vitek at that point: I got to know him a little bit later in the story. But certainly Vitek is a very important figure in all of this and I credit him with a lot of imaginative thinking. He deserves recognition for the work he has done, and I am a great admirer of his.

At any rate, after talking to Pat I decided that I would try to introduce the idea of Open Access to the biomedical research community in a way that — in retrospect — I can now see was a little brash and naïve [small laugh].

E-BIOMED

RP: Specifically, you wrote the E-BIOMED proposal?⁶²

HV: Yes. In May 1999 I wrote what I now call a manifesto called E-BIOMED. I think it was actually quite a useful document, and there were two versions. The first version was published in May, and the second in June.

RP: What was the E-BIOMED manifesto?

HV: Frankly, the concept was more than just an incremental step towards Open Access: it was a plan to transform the entire workings of the scientific publishing world in our area. So it was a pretty dramatic proposal.

RP: What specifically did you propose?

HV: Part of the idea was to introduce the kinds of changes that we are now seeing being achieved. Essentially, to create a huge digital library in which we would place existing journals, plus possibly some new journals, and then make it all fully searchable and available at the click of a mouse.

⁵⁹ Established in 1988 as a national resource for molecular biology information, the National Centre for Biotechnology Information (NCBI) creates public databases, conducts research in computational biology, develops software tools for analysing genome data, and disseminates biomedical information — all for the better understanding of molecular processes affecting human health and disease

⁶⁰ The interview with Tracz in this series can be read at:

<http://poynder.blogspot.com/2006/05/interview-with-vitek-tracz.html>.

⁶¹ David Lipman is director of the US National Institutes of Health (NIH) National Center for Biotechnology Information (NCBI), and the National Library of Medicine (NLM), and the <http://www.ncbi.nlm.nih.gov/CBBresearch/Lipman>.

⁶² *E-BIOMED: A Proposal for Electronic Publications in the Biomedical Sciences*
<http://www.nih.gov/about/director/pubmedcentral/ebiomedarch.htm>

In addition, I suggested ways in which a lot of things that currently aren't ever published could be made publicly available. So my proposal was that we also make available a large collection of other kinds of research information, including online conversations with authors of previously published papers, and data sets that don't get published.

RP: You envisaged that some of the information in E-BIOMED would not be peer reviewed?

HV: That's right. Some of this information might not even be peer reviewed. It would, however, be flagged as such, and it would be surveyed by two people who had standing in the research community, and who met certain criteria. It would be their job to block information that we wanted to keep out — pornography and such like.

So all in all E-BIOMED was a radical proposal, and it generated a lot of heat.⁶³

RP: In particular, you attracted criticism from publishers, who saw E-BIOMED as a threat to their business.

HV: Indeed, and shortly after I published my E-BIOMED manifesto lobbyists for American publishers were lining up in Congress to talk to my appropriators about how I was trying to destroy the capitalist system!⁶⁴ I also had a lot of people representing scientific societies up in arms about what they saw as a misguided effort, and one that would be disruptive to their financial interests.

RP: You defended E-BIOMED fairly robustly. Chemistry and Engineering News reported at the time⁶⁵ that you described the reaction of publishers as "paranoiac."

HV: I don't think I said paranoiac. I wasn't completely surprised by the opposition. What I was surprised by was the vehemence of that opposition, and the fact that the publishers took their complaints straight to Congress. In doing so they ran the risk of turning what should have been an internal debate, even perhaps a friendly debate, into a debate that became very high-pitched, and potentially disruptive to the NIH, an organisation upon which everyone depends.

RP: You didn't see it as posing a commercial threat to publishers then?

⁶³ In his profile of Varmus in *Wired*, Shreeve commented: "Though the smaller, friendlier physics community had embraced the idea of freely sharing preprints, the more crowded and competitive world of biomedicine greeted it with open hostility, comparing it to socialism and lamenting what they saw as a loss of peer review. E-Biomed 'is among the very worst ideas I have ever heard,' one University of Wisconsin-Madison biochemist told *The New York Times*."

⁶⁴ As the OA movement has gathered momentum, so publishers have increasingly sought to neutralise it by means of lobbying. Reed Elsevier, for instance, spent \$12.5 million on lobbying in the US between 1998 and 2004, a figure that grew by another \$3.18 million in 2005 alone. This puts a number on the influence exerted by scholarly publishers on US government policy, and indicates that Reed Elsevier's annual spending on US lobbying activities increased by 695% between 1998 and 2005. For a detailed breakdown of how this money was spent see William Walsh's figures at: <http://www.library.gsu.edu/news/index.asp?view=details&ID=9945&typeID=62>.

⁶⁵ *Chemistry and Engineering News*, July 19, 1999, <http://pubs.acs.org/hotartcl/cenear/990719/7729gov1.html>

HV: Yes, of course it was a commercial threat to them. But there was still a debate to be had here, and I guess I was a little surprised by two things. First, I was surprised that people would actually go to Congress and try to deal with it at a congressional level, rather than debating it in the normal fora.

Second, I was surprised by how conservative many of my scientific colleagues proved. Many of them apparently felt that they would have a hard time distinguishing trash from high-quality work because the system would be broken.⁶⁶ They also didn't read my proposal carefully enough to see that I was not saying that NIH would now publish everything, and that everybody would be beholden to NIH. I wasn't arguing that at all.

RP: As you say, one of the concerns was that NIH would itself become a publisher. What did you do to address these concerns?

HV: The second version of E-BIOMED — which I issued a few months after the first — responded to a lot of the criticisms. I cut down the scope of the project, and I tried to make a number of things clear.

It is worth noting, by the way, that apart from the criticism we also had many, many useful comments; and so in the end I think the whole thing ended up being a very helpful exercise. It certainly made a lot of people who, like me, had not paid much attention to these issues before recognise that the Internet posed a tremendous opportunity for changing the way in which we deal with the biomedical research literature.

PubMed Central

RP: The end product of your E-BIOMED proposal was PubMed Central.⁶⁷ What did you hope to achieve with PubMed Central?

HV: PubMed Central was a way of trying to get a part of what I had envisaged with E-BIOMED created by means of urging journals to voluntarily provide their content to a digital library supported by the NIH. This was implemented by David Lipman, and launched around the time that I left the NIH — in fact I think the last thing I signed before I left at the end of 1999 was the PubMed Central press release. So it turned out that my last public act at NIH was getting PubMed Central in motion [laughs].

RP: PubMed Central was launched at the beginning of 2000. By then, as you indicated, the more radical aspects of E-BIOMED had been abandoned. The original

⁶⁶ See, for instance: <http://www.nih.gov/about/director/ebiomed/com0627.htm>.

⁶⁷ PubMed Central is a free open access online database of scientific papers. PubMed Central was developed as an online archive of biomedical journal articles. Participation by publishers in PubMed Central (PMC) is voluntary, although participating journals must meet certain editorial standards. Currently 251 journals supply papers, some of which are only available after an embargo of six or twelve months. <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pmc>

plan to require researchers to deposit their papers, for instance, was replaced by a request that publishers do so. It was also accepted that publishers would operate an embargo, and only deposit papers six months after publication. And while E-BIOMED had envisaged authors retaining copyright in their papers, it was agreed that publishers would continue to acquire copyright in papers published. In addition, the plan to include non-peer-reviewed material was discarded. All in all, you had to give up quite a lot?

HV: It was not a question of giving up on any of the goals of the E-BIOMED proposal. Rather the aim was to make PubMed Central a starting point — and the easiest way to start was to ask existing journals to provide their content after several months.

RP: *By agreeing to an embargo?*

HV: Yes. The data shows that people who subscribe to journals want those journals immediately, and so if they can only get them online for free after six months they are not going to give up their subscriptions. For this reason it seemed like a reasonable request to ask that journals provide their content after six months. That is still a very useful thing to have. And to understand why, you need to distinguish between the two modes in which the literature is used.

RP: *Can you expand on that?*

HV: Sure. The first mode is browse mode. Almost everybody has a few journals that they subscribe to and which they want to browse every week, or every month. Because people value those journals for browsing they are the journals that would most feel the impact of subscription loss if they were truly Open Access — as in available through the Internet from time of publication, or even as a preprint before the paper is published.

RP: *And the second mode?*

HV: The second mode is search mode. Because of the power that we now have to search the entire literature in digital form the search mode is becoming increasingly important. What we were trying to do with PubMed Central, therefore, was to create a deep library where, when you search the literature, you search everything, going back — we hoped — many years.

So while the open access definition assumes immediate access to the literature, part of the Open Access Movement is also focused on creating a deeper library — one that goes back to articles that were published in the last, say, 50 years.

RP: *This means digitising and inputting archival material?*

HV: We have done that with some journals yes. Consider, for example, the journals of the American Society for Microbiology,⁶⁸ which publishes several important journals and now contributes to PubMed Central. The articles in the ASM journals

⁶⁸ <http://www.asm.org>

that have been deposited in PubMed Central go back many years — I'm not sure how far, but certainly back to the 1970s. This is an incredible resource, and for those of us that use that literature it is really great to be able to get this stuff without having to go to the library.

RP: So even though it may not meet the strict definition of Open Access — as in providing immediate and free online access at the time of publication — making research information available on an embargoed basis is a valuable thing in itself?

HV: That's right.

RP: What do you think has been achieved with PubMed Central to date, and are you pleased with progress?

HV: As I indicated, I left NIH around the time PubMed Central was launched, although I was on the outside advisory committee for the first several months after it was launched.

That said, it was soon clear that we were not getting very many high ranking journals. In fact, there were only a handful — of which the *Proceedings of the National Academy of Sciences*⁶⁹ and *Molecular Biology of the Cell*⁷⁰ were the most prominent. There were a few others, but it was clear that there were no more than 50 to 100 journals at that point that were willing to provide content to PubMed Central. That was clearly a serious shortcoming, and it was why the Public Library of Science⁷¹ got launched.

Public Library of Science

RP: So the Public Library of Science was a response to the slow progress of PubMed Central?

HV: Yes. So some of us who were on the PubMed Central advisory board started trying to figure out how to bring to the attention of our fellow scientists, and to publishers, the aspects of science publishing that make deposition in a free digital library like PubMed Central not only important, but entirely reasonable.

RP: When it was launched in November 2000, PLoS was initially an advocacy group, and its first act was to publish an Open Letter and invite scientists to sign it. Can you talk me through the genesis of PLoS, and the aim of the Open Letter?

HV: Sure. Our view was that scientists contribute a tremendous amount for free in order to make the scientific literature work: we write articles for free; we edit for free; we do reviewing for free; and then we subscribe to the journals those papers are published in — which means we pay to get back what we have provided for free.

⁶⁹ <http://www.pnas.org>

⁷⁰ <http://www.molbiolcell.org>

⁷¹ www.plos.org

So we felt that given that print journals were not even taking the pretty conservative step of providing content to PubMed Central six months, or even a year, after publication, they simply weren't acting in the interests of the scientific community.

One way to highlight this, we decided, was for those of us who contribute to these journals to say that we would only continue to do so if the journals complied with a request to contribute the papers they published to PubMed Central within six months.

RP: *And that was what the Open Letter, or petition, called for? Essentially scientists signing the letter were saying that they would boycott any journal that did not comply with their request.*⁷²

HV: Indeed. By taking this stand we felt that we would be honouring our principles by only providing our services as authors, as editors, as reviewers, or as personal subscribers, to those journals that complied with our request. The aim was to favour those journals that did contribute their content to PubMed Central. What we called for, by the way, was not true Open Access, but greater access.

RP: *If it would not have provided true Open Access what was the point?*

HV: The intention was to popularise the idea of greater access. By signing this petition, or pledge — I did actually call it a boycott — we would be exemplifying the importance to the scientific community of greater access.

We also hoped that it would provide some muscle to push a lot of journals in the direction of Open Access. And in fact we did see more publishers over the course of time signing up to provide journal content to PubMed Central.

RP: *Nevertheless the Open Letter also failed in its objective didn't it? Even though they might have signed the letter, most scientists carried on as before?*

HV: True, and so most of us still had the sense that the movement was not achieving its objectives fast enough.

RP: *Why do you think the petition — which collected 34,000 signatures — failed to have the desired effect?*

HV: Mainly because my colleagues had not yet fully appreciated what Open Access publishing could be like: what it would mean to have greater access; and why it would be such a powerful thing to have a true digital library — one that was a great deal better than you could get by just going to PubMed⁷³ and hoping that by searching on

⁷² The letter ended with the threat: "To encourage the publishers of our journals to support this endeavour, we pledge that, beginning in September 2001, we will publish in, edit or review for, and personally subscribe to only those scholarly and scientific journals that have agreed to grant unrestricted free distribution rights to any and all original research reports that they have published, through PubMed Central and similar online public resources, within 6 months of their initial publication date." <http://www.plos.org/about/letter.html>

⁷³ As distinct from the PubMed Central *full-text* service that Varmus launched, PubMed is a free service of the National Library of Medicine that provides access to over 15 million *citations* from MEDLINE (which itself has 11 million citations and abstracts from health and medical journals and

the abstracts and titles you would be able to find the papers that you needed, and you could then obtain the full-text of all those the articles in your institution, assuming you work in an institution that is.

RP: In an attempt to move things forward more quickly, therefore, PLoS decided to become a publisher itself. More importantly, it decided to become an open access publisher.

HV: Right. When we saw the effect on publishers wasn't as dramatic as we had hoped we decided that another way to try and convince our colleagues of the virtues of open access publishing was to create open access journals ourselves. That was the point at which we decided that PLoS would be more than an advocacy group; that it would become a publisher.

That was a much bigger step, of course, because creating open access journals requires having a business plan; one, moreover, that would work over the long term.

RP: Was the aim to provide an alternative to traditional journals, or to offer a model for traditional journals to copy?

HV: It could be either. But what we were really trying to demonstrate was what open access publishing could look like. We also decided that we would start at the very high end, and try to publish journals that would compete with the most outstanding journals.

RP: A hard target for a start-up publisher?

HV: Sure. But we had watched with interest Vitek's experiment with BioMed Central,⁷⁴ and I think Vitek would acknowledge that one of the problems that BioMed Central had then — and to a certain extent still has — is that its journals do not compete with the very best in the field. They are just not thought to be in the top tier of journals.

So we decided that the one way to make open access journals much more acceptable to the scientific community — which views publication in certain journals to be a very strong criterion for appointment and promotion in jobs, and certainly a very important component of the regard with which one is held in a scientific field — would be to try to create open access journals that conferred on authors the same kind of prestige that they obtain when they publish in *Nature*,⁷⁵ *Science*,⁷⁶ *Cell*,⁷⁷ and the handful of other top biomedical journals.⁷⁸

other news sources), and other life science journals for biomedical research back to the 1950s.

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>

⁷⁴ Founded by Vitek Tracz in 1998, BioMed Central was the first "author-pays" Open Access publisher

<http://www.biomedcentral.com>

⁷⁵ <http://www.nature.com/nature/index.html>

⁷⁶ <http://www.sciencemag.org/magazine.dtl>

⁷⁷ <http://www.sciencemag.org/magazine.dtl>

⁷⁸ As Shreeve put it in his *Wired* profile of Varmus, "Varmus wanted the PLoS journals to have the credentials to immediately draw outstanding submissions. They lured the editor of the top journal *Cell* to serve as executive director. They hired staff who had worked at both *Nature* and *Science*, and they began wooing prestigious contributors."

RP: The first challenge was obtaining funding I guess?

HV: Absolutely. It is one thing to say you are going to become a publisher and create journals: it is quite another thing to actually do it. And we make no apologies for the fact that we didn't have Vitek's money [laughs] in our pockets to get started; and we didn't have the weight of a scientific society behind us.

So we needed financing to get going. To that end, we went from foundation to foundation until we found the Gordon and Betty Moore Foundation,⁷⁹ which believes in the principles that we were espousing.

RP: And in 2003 the Gordon and Betty Moore Foundation provided a \$9 million grant to get PLoS up and running.⁸⁰

HV: Right. So we hired some outstanding people and put together an editorial staff — and the necessary publishing people — to allow us to make the journals a reality.

The business model

RP: Let's turn to the business model. I guess Vitek Tracz did much of the groundwork when he created BioMed Central? But perhaps you could explain how the open access business model works?

HV: Essentially, it is a question of shifting the costs — moving the overall expense of publishing away from the reader to the author, who at the time when the paper is accepted for publication pays to be published.⁸¹ Some open access journals might also have a small submission fee, and then a larger fee at the time when the paper is accepted. Anyway, we believe that the author-pays model will enable us to cover our costs. In addition, costs are lower when you are publishing solely on the web.

RP: Right, because PLoS is an electronic-only publisher?

HV: Well, we do have some paper products,⁸² which we use basically for advertising. But in the long run we don't expect to have more than a few copies for archiving, and perhaps a small number for those few libraries that want paper. But in general we are for eliminating the cost of paper, and eliminating the cost of shipment and transportation. In that way, the overall costs will come down — although they are still significant. The point is, however, that these costs are shifted from institutional subscribers to the authors.

⁷⁹ Established in September 2000, the Gordon and Betty Moore Foundation seeks to develop outcome-based projects that will improve the quality of life for future generations. <http://www.moore.org>. Gordon Earle Moore is the cofounder of Intel Corporation and the author of [Moore's law](#).

⁸⁰ http://www.moore.org/news/2002/news_pr_121702plos.asp.

⁸¹ Referred to by BioMed Central as an article processing charge, in most cases this is expected to be paid by the researcher's funder or institution.

⁸² The annual print subscription for PLoS journals is \$160. <http://medicine.plosjournals.org/perlserv?request=get-static&name=contact>.

RP: By shifting the costs to the authors you are able to make their papers free at the point of use, and from the day of publication. This is because the charges are front-loaded?

HV: Right.

RP: Currently PLoS charges \$1,500 for each paper it publishes.

HV: Yes. There are other ways to cover our costs too, including advertising on the web site of the journal, and through various kinds of memberships and sponsorships. We also think that open access journals will be able to increase their revenues by the same kind of philanthropic mechanisms that national public radio uses in America.

RP: For those of us who don't live in America can you explain how that works?

HV: Well, you get good programming with US National Public Radio⁸³ largely because the public is supportive of it. In other words, many people — including me — provide donations to radio stations in order to ensure the quality of NPR. I see no reason why open access journals couldn't be supported in the same way.

The point to bear in mind is that while we can get costs down, journals are still expensive to produce. Moreover, we are providing a lot of so-called front matter — translations of scientific reports into lay language for students, and into simple language for patients and so on. We couldn't do these additional things without support from philanthropy.

RP: The front matter you refer to is paid journalism is it?

HV: We don't pay for it *per se*, but the editors need to solicit these articles, and then format them and prepare them for publication — because it's not a case of simply sitting back and waiting for the front matter to come in, in the way we do for research reports. This requires a lot of editorial attention, which means we have to hire people to develop the front part of the journal.

That said, we expect that some costs will go down considerably. The model also assumes, for instance, that people will submit their material in a way that is easy to publish. All in all, therefore, we firmly believe that over the next few years we will begin to break even.

RP: Can you say more about the journals, and how you plan to develop them?

⁸³ National Public Radio (NPR) is an independent, private, not-for-profit corporation that sells programming to member radio stations; together they form a loosely organised public radio network in the United States. NPR was created in 1970, following the passage of the Public Broadcasting Act in 1967, which established the Corporation for Public Broadcasting and also led to the creation of the Public Broadcasting Service. The network was founded on February 24, 1970, with 90 public radio stations as charter members. http://en.wikipedia.org/wiki/National_Public_Radio

HV: Sure. Currently we have two flagship journals: *PLoS Medicine*⁸⁴ and *PLoS Biology*.⁸⁵ We expect those two journals to be competitive with the very best journals in the world. In addition, we are developing a series of so-called community, or speciality, journals. The ones we have launched to date are called *PLoS Pathogens*,⁸⁶ *PLoS Computational Biology*⁸⁷ and *PLoS Genetics*,⁸⁸ and there will be additional ones in the near future.

While we want these community journals to be extremely high in quality, they are not as broad as *PLoS Biology* or *PLoS Medicine* since they confine themselves to the disciplines after which they are named.

Additionally, we plan to launch a journal that will be called *PLoS ONE*. This will cover everything in biology and medicine — any field is acceptable.⁸⁹

RP: *Can you say more about that?*

HV: What will also be different about it is that we will evaluate papers submitted to it for scientific quality, but not for impact. So we are not going to try to decide which the most interesting papers are, and we are not going to restrict the numbers of papers published.

Consequently we expect this to be a very large compendium of papers that have been vetted for scientific quality, but which will not be confined in terms of their likely importance. We will include papers that confirm earlier findings, for instance, and papers that make only a modest extension of prior findings.

RP: *This is possible because some of the limitations and costs of traditional publishing begin to go away on the Internet?*

HV: That's right. And the logic is that *PLoS ONE* will publish papers that will be read not by people picking up a journal (since it won't be printed) but by searching the journal database for articles in their own field. The journal will be highly visible, but visible only through search engines.

RP: *In other words, it will exploit the search mode enabled by the digital environment that you mentioned?*

HV: Yes, and for authors it means that they will be able to publish work without going through the very frustrating process of trying to find a commercial or society

⁸⁴ <http://medicine.plosjournals.org/perlerv/?request=index-html&issn=1549-1676>

⁸⁵ <http://biology.plosjournals.org/perlerv/?request=index-html&issn=1545-7885>

⁸⁶ <http://pathogens.plosjournals.org/perlerv/?request=index-html&issn=1553-7374>

⁸⁷ <http://compbiol.plosjournals.org/perlerv/?request=index-html&issn=1553-7358>

⁸⁸ <http://genetics.plosjournals.org/perlerv/?request=index-html&issn=1553-7404>

⁸⁹ *PLoS ONE* is expected to be launched at the end of 2006. In the meantime, PLoS has developed another journal — *PLoS Clinical Trials* — which was launched on 19th May 2006. The publisher describes this as, "an international peer-reviewed, open access journal that publishes results of randomised trials from all medical and public health disciplines." It adds, "The journal's aim is to increase the breadth of clinical trials reporting and thus ensure that all trials on human participants are reported in the peer-reviewed literature and accessible to all." <http://journals.plos.org/plosclinicaltrials/information.php>.

publisher that thinks the work is important enough to publish. As I say, we are not going to worry so much about how important a paper is, but publish it on the basis that it is good and deserving of public notice.

RP: PLoS deposits all its papers in PubMed Central. The expectation, then, is that users will find these papers by searching PubMed Central?

HV: That is where we expect people to do their searching. They can also search through our own web site, but it is much more efficient to search through PubMed Central, and that is where we expect everyone to do their searching; there or in the other digital libraries — of which there will be several ultimately.

International phenomenon

RP: Right. The Wellcome Trust, for instance, plans to build a European PubMed Central.⁹⁰ This reminds us that, like science itself, the Open Access Movement has to be viewed in an international context. It is also attracting the attention of politicians around the world. In the UK, in 2003, for example, an all-party group of British politicians sitting on the Science and Technology Select Committee⁹¹ launched an enquiry into science publishing.⁹² You appeared before that Committee as a witness. Why?

HV: Because I believe in the importance of open access publication, and the UK is obviously a very important source of scientific discovery. For that reason I would like to see the United Kingdom exemplify what open access publishing could do.

RP: When the UK Select Committee published its report⁹³ it called on the British government to create a network of institutional repositories and mandate all publicly-funded researchers to deposit a copies of their articles in those repositories, thereby making their research accessible to all “free of charge, online.” Was that the kind of conclusion you were looking for?

HV: Sure, I think the Committee did quite a good job. They took it very seriously and I was impressed by the members of the panel.

RP: The UK government, however, declined to intervene,⁹⁴ arguing that it was “not aware that there are major problems in accessing scientific information”.

⁹⁰ http://www.wellcome.ac.uk/doc_WTX022826.html

⁹¹ The select committee system plays an important role in enabling Parliament to scrutinise the work of Government.

http://www.parliament.uk/parliamentary_committees/science_and_technology_committee.cfm

⁹² In December 2003, the UK Science & Technology Select Committee announced an inquiry into science publications

(http://www.parliament.uk/parliamentary_committees/science_and_technology_committee/scitech111203a.cfm)

⁹³ In July 2004 the Select Committee published a report *Scientific Publications: Free for all?*

<http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/1200/120002.htm>.

⁹⁴ <http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/1200/120006.htm#a1>

HV: Yes, although I am a great admirer of David Sainsbury⁹⁵ I thought his response was limited.

RP: *What the UK Government did in effect was to pass the buck, saying that as Research Councils UK was currently working on a policy framework for the dissemination of scientific information,⁹⁶ the Government would review its own position after that policy was finalised?*

HV: Right, and I think it hoped that RCUK would make a more conservative statement than it did. In the event, I was pleased to see that the UK research councils⁹⁷ pretty much subscribed to the views of the committee, and gave much more robust support for Open Access. Certainly I expect that when it is finalised the RCUK policy will be an improvement on the US NIH policy.⁹⁸ However, the system is different in England because the major funder of medical research is the Wellcome Trust.⁹⁹

RP: *I'd like to come to the NIH policy in a moment. First, can you say what the implications are of the Wellcome Trust being the major funder of medical research in the UK?*

HV: The Wellcome Trust is not a governmental organisation and so it has the power to do what it likes. Fortunately it has a very strong leader in Mark Walport,¹⁰⁰ who is a fervent believer in Open Access, and who showed true leadership by making a very clear policy statement.

RP: *You are referring to the Wellcome Trust mandate,¹⁰¹ which is now in place, and which requires all its grant holders to make their research freely available in PubMed Central within six months?*

⁹⁵ Lord Sainsbury is Parliamentary Under-Secretary of State for Science and Innovation, with responsibility for the Office of Science and Technology, Research Councils and space matters. <http://www.dti.gov.uk/about/dti-ministerial-team/page8414.html>.

⁹⁶ <http://www.rcuk.ac.uk/access/index.asp>

⁹⁷ Research Councils UK (RCUK) is a strategic partnership through which the UK's eight Research Councils work together to champion the research, training and innovation they support. <http://www.rcuk.ac.uk>

⁹⁸ Despite the publication of a position statement on OA in June 2005, Research Councils UK has yet to introduce an OA policy. <http://www.rcuk.ac.uk/access/index.asp>. This follows intense lobbying by publishers, particularly The Association of Learned and Professional Society Publishers (ALPSP). See, for instance: <http://www.alpsp.org/news/rcuk.htm>.

⁹⁹ The Wellcome Trust is an independent charity funding research to improve human and animal health. Established in 1936 and with an endowment of around £11 billion, it is the UK's largest non-governmental source of funds for biomedical research. <http://www.wellcome.ac.uk>

¹⁰⁰ Dr Mark Walport is director of the Wellcome Trust. <http://www.kcl.ac.uk/175/biogs/walport.html>.

¹⁰¹ From October 1st 2005 it became a condition of Wellcome funding that researchers must deposit papers resulting from funded research into the PubMed Central open access repository within 6 months of publication. Specifically, the Wellcome Trust policy:

"Expects authors of research papers to maximise the opportunities to make their results available for free and, where possible, to retain their copyright.

Will provide grant holders with additional funding to cover the costs of page processing charges levied by publishers who support the Open Access model.

"Requires electronic copies of any research papers that have been accepted for publication in a peer-reviewed journal, and are supported in whole or in part by Wellcome Trust funding, to be deposited into PubMed Central (or UK PubMed Central once established). Note that this requirement will apply to all grants awarded after 1 October 2005, and from 1 October 2006 to all grants regardless of award date.

HV: Exactly. And this has provided a good model for RCUK, which will surely recognise that publicly funded researchers shouldn't be more restricted when it comes to providing access to their research than those who are supported by a private funder like the Wellcome Trust.¹⁰² So that has helped a lot.

The Howard Hughes Medical Institute has taken a very similar role in the US: supporting Open Access, offering extra money to investigators who use Open Access publications so they can cover their authors fees, and so on. This provides a strong inducement for researchers to embrace Open Access.

The problem, however, is that while it is a highly respected, institution, the Howard Hughes Institute is a very small fraction of the total research effort in the US. In England, the Wellcome Trust is a very much larger fraction of the total medical research effort.¹⁰³ That is what I meant when I said that the system is different in England.

The NIH policy

RP: Let's turn to the NIH policy then: am I right in thinking that after you left NIH very little happened in the way of open access initiatives until July 2004, when the US House Appropriations Committee instructed the NIH to develop a policy requiring free online access to articles based on NIH-funded¹⁰⁴?

HV: Yes. That's right. Until then there was no advocacy within the NIH that I am aware of, apart from David Lipman — who has always been a very strong proponent of freeing up the literature.

RP: So the pressure then started to come from Congress itself. That is interesting, given that publishers had turned to Congress in order to fend off the perceived threat from the E-BIOMED proposal. I guess OA advocates took a leaf out of the publishers' book and began their own lobbying?

HV: Well, there was also a two-year lapse of leadership at NIH, between my leaving and Dr. Zerhouni taking office.¹⁰⁵

"Affirms the principle that it is the intrinsic merit of the work, and not the title of the journal in which an author's work is published, that should be considered in making funding decisions and awarding grants." http://www.wellcome.ac.uk/doc_WTD002766.html

¹⁰² It remains to be seen if RCUK follows the model of the Wellcome Trust. After the intense public lobbying by publishers, in April 2006 RCUK announced that it would undertake [more research](#) prior to finalising its position on OA — to the acute frustration and anger of open access advocates: <http://openaccess.eprints.org/index.php?archives/81-guid.html>.

¹⁰³ The NIH annually invests over \$28 billion in medical research.

¹⁰⁴ <http://www.earlham.edu/~peters/fos/nihfaq.htm>

¹⁰⁵ Elias Zerhouni was appointed to replace Varmus in 2002, two years after Varmus left. <http://www.nih.gov/about/director>

RP: And Dr. Zerhouni obviously had his work cut out when he arrived. There would, however, have been no ignoring Congress' request that NIH develop a public access policy?

HV: It should be said that Dr. Zerhouni has always had an interest in publication issues. He came to the meeting we held at the Howard Hughes Institute; and he listened to me and many other people talk about how to make literature more accessible. But, as you say, he had to pay attention to Congress' request for tax-payer access.

RP: Which of course led to the NIH policy, the final version of which was published¹⁰⁶ in February 2005. What are your views on that policy?

HV: I would argue that the policy he came up is not as strong as it ought to be.

RP: Right and no doubt because publishers lobbied hard to protect their financial interests. The end result was that the draft proposal — which would have mandated researchers to deposit all NIH-funded research papers in PubMed Central within six months — was replaced with a policy in which they are only encouraged to deposit their papers, at least within 12 months. In short, the original proposal was watered down in the face of publisher criticism?

HV: Actually, the policy doesn't say encourage; it says request. Obviously it doesn't say mandate, or demand, or require — which may have been more than Dr. Zerhouni could have done. However, people are quite sensitive to what NIH wants because it is a major source of grant funding, so it wouldn't be too hard to be one step back from requirement, to say encourage or expect, and still get compliance. In any case, right now NIH is not getting compliance with its public access policy, and I think there will be repercussions from that.

RP: As you say, the exact wording used is important if you want people to comply. But what sort of repercussions do you anticipate as a result of non-compliance?

HV: Congress is going to say: "What is in this database?" And when people take a look they will see that very little is going into PubMed Central that was not already being contributed before the new policy came into effect.¹⁰⁷

RP: So Congress will likely demand a stronger policy?

HV: Well, they are obviously being told by Open Access advocates that this isn't working, and so when Dr. Zerhouni goes to Congress to testify in the appropriations process I am sure Congress will ask him: "How has this policy worked? What has been done?"¹⁰⁸

¹⁰⁶ The Policy on Enhancing Public Access to Archived Publications Resulting from NIH-Funded Research, <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-05-022.html>

¹⁰⁷ In fact compliance has been pitiful. In a January 2006 report to Congress, Elias Zerhouni said that less than 4% of eligible articles have been added to PubMed Central since the introduction of the NIH Policy in May 2005. http://publicaccess.nih.gov/Final_Report_20060201.pdf

¹⁰⁸ In April 2006, a key advisory panel to the NIH — the Public Access Working Group (PAWG) — reaffirmed its continued support for reforms to the NIH Public Access, stressing its November 2005

RP: So what would your advice be to him?

HV: It's not what I *would* advise: it is what I *do* advise. I think he should give a clear picture of what his expectations are, and encourage grantees to submit their accepted work to PubMed Central. The problem is that while I believe Dr Zerhouni is a proponent of Open Access, he is caught in a political battle — and it may be more difficult for him than it was for me to resist the publishers.¹⁰⁹

RP: Are you still involved with PubMed Central?

HV: I remain an advocate for it, but I'm not on the advisory committee any more.

Green or Gold?

RP: PLoS is an OA publisher. The main thrust of OA initiatives today, however, seems to be focused on mandating authors to self-archive papers they publish in traditional journals, rather than insisting that they publish in open access journals. Indeed, there has been a long-standing and vigorous debate within the OA movement about the respective merits of the so-called Green and Gold roads.¹¹⁰ The debate seems to revolve around the issue of whether it is better for OA advocates to put all their energy into the creation of new open access journals, or to focus on lobbying research funders and governments to require researchers to self-archive the papers they have published in subscription journals. What are your views on that debate?

HV: My views are very clear: at this point self-archiving is not Open Access.

recommendation that the policy be made *mandatory*, and that all NIH-funded works *must* be made publicly available in PubMed Central within six months of publication.

<http://www.taxpayeraccess.org/media/Release06-0413.html>.

¹⁰⁹ Indeed, it has become increasingly clear that political direction was required. On December 14th, therefore, the American Center for CURES Act of 2005 was introduced into the US Senate. Currently "in committee", this would create a new agency within the NIH, The American Center for Cures, whose primary mission would be to translate fundamental research into therapies. Its aims are not limited to OA, but it would "require" deposit of research papers at the time of publication. CURES also extends the OA policy beyond the NIH to the other agencies within the [Department of Health and Human Services](#), including the [Centers for Disease Control and Prevention](#) and the [Agency for Healthcare Research](#). <http://www.earlham.edu/~peters/fos/newsletter/01-02-06.htm#cures>.

And on May 2nd 2006, Senators John Cornyn (R-TX) and Joe Lieberman (D-CT) introduced The Federal Research Public Access Act of 2006 (FRPAA). This would require that US Government agencies with annual extramural research expenditures of over \$100 million make manuscripts of journal articles stemming from research funded by that agency publicly available via the Internet. The manuscripts would also need to be maintained and preserved in a digital archive maintained by the agency concerned, or in another suitable repository that permitted free public access, interoperability, and long-term preservation. Each manuscript would need to be freely available to users without charge within six months of being published in a peer-reviewed journal.

¹¹⁰ The Green Road implies researchers continue to publish in subscription-based journals, in the way they always have done, but then self-archive their papers on the Web, often after an embargoed period. The Gold Road implies creating new author-pays OA journals in which the publisher charges the author to publish, but publishes the paper on the Web from the day of publication, and without any access restrictions. http://www.ercim.org/publication/Ercim_News/enw64/jeffery.html.

RP: *Why do you say that?*

HV: One of the important components of the definition of Open Access that we have all agreed on is that research information should be placed in a searchable database. Right now the only way to be confident that you can do that effectively is by using a large public digital library like PubMed Central.

RP: *On the Web, of course, it is possible to build a single virtual archive by aggregating multiple distributed archives. Self-archiving advocates argue, therefore, that it would be better to create a network of institutional repositories, and allow researchers to search them as a single archive. This, they say, would achieve the same results as a central database, but would roll out Open Access more quickly?*

HV: Sure, you can have researchers self-archive their papers, and then use Google to search across them. That might work at some point in the future, but the problem today is that you can't be sure of finding everybody's web site; moreover, in doing a search like that you will inevitably retrieve a lot of junk.

RP: *Self-archiving advocates, however, don't necessarily advise scientists to use Google to do their searching. Their argument is that if researchers were mandated to deposit their published papers into institutional repositories, and if those repositories were all OAI-PMH¹¹¹ compliant, then effective searches could be conducted on the entire OA corpus using specialist search engines like OAIster,¹¹² which are able to harvest papers from the entire population of OAI-compliant repositories.*

HV: That may be, but I just don't believe it works. My technical knowledge is not at a very advanced level, but I understand that it is necessary to have a search engine that searches across all the papers, that has the right lexicon, and which can accurately retrieve the information that investigators want. Some day perhaps we will be able to just self-archive, and it will all work fine; but we're not there today.

Given the current level of technical development, therefore, the only approach that makes ultimate sense to me is to use a central public database. In other words, I believe the gold standard for OA is immediate deposition into a fully searchable public library of science, and the only one that I would trust for the purpose of searching OA literature today is PubMed Central.

RP: *Of course while people are free to debate the respective merits of Green versus Gold OA, and distributed versus central archives, I guess the various different*

¹¹¹ OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) is a protocol developed by the [Open Archives Initiative](http://www.openarchives.org/). It is used to harvest the metadata descriptions of records in an archive so that services can be built using metadata from many archives, not least search services to allow cross-searching across all OAI-compliant archives. The protocol is usually just referred to as the OAI Protocol. OAI-PMH is based on client-server architecture, in which "harvesters" request information on updated records from "repositories". A number of software systems support the OAI-PMH, including [GNU EPrints](http://www.gnu.org/licenses/licenses.html) from the University of Southampton and [DSpace](http://www.dspace.org/) from MIT. A number of large archives support the protocol including arXiv and the [CERN Document Server](http://cds.cern.ch/).
<http://www.openarchives.org/OAI/openarchivesprotocol.html>.

¹¹² [OAIster](http://oaister.umdl.umich.edu/o/oaister) is a project of the University of Michigan Digital Library Production Service. Its stated purpose is to "create a collection of freely available, previously difficult-to-access, academically-oriented digital resources that are easily searchable by anyone." <http://oaister.umdl.umich.edu/o/oaister>

approaches are not mutually incompatible. Regardless of whether researchers publish in a subscription-based journal or an OA journal, for instance, they can still deposit that paper in PubMed Central. Likewise, they can deposit their papers in both PubMed Central and their institutional repository. The point is that, however it is achieved, the aim is to make scientific papers freely available on the Web right?

HV: That's right.

Moral imperative

RP: *When you appeared before the UK Science & Technology Select Committee you said that there was a moral imperative for scientists to make their research freely available to the public. There is, you said, an obligation on researchers to "ensure everyone in this country [Great Britain] who has a curiosity about the information should have access to it....."¹¹³*

HV: Actually, I would say that everyone in the world should have access to it. It is important to stress here that publishing is a fundamental part of the process of doing science. Moreover, as a scientist I am not writing for money — like my wife, who was a professional writer at one time — but I am writing for fame: I want everyone to read what I write.

So there is a basic distinction between writers who write for a living, and writers who write to make their thoughts and findings known, and who want to be famous as opposed to being rich. Everybody who writes a scientific paper is writing to be read, not to make money. For that reason we volunteer our services, and we don't get paid.

That is what makes Open Access a powerful concept for scientists. It means, for instance, that anyone anywhere in the world who searches for oncogenes, mouse models, or any other search term that applies to my own work, will find it, and will be able to have immediate access to it. That is the goal we are hoping to achieve.

RP: *Does this moral imperative apply only to publicly funded researchers?*

HV: No. I don't care if they are publicly funded or not. The same applies to those who are funded by voluntary groups, advocacy groups, or charities. And while the issue might be slightly different for people who are funded by commercial entities, as a general principle when they decide to publish they also do so in order for everyone to see their work. Moreover, they are not paid for that work either.

The point is that people who write research reports have a very different goal to people who write for a living. This means that for scientific work we can now create a publication system that covers the cost of publication, and yet takes advantage of the power of the Internet to make the scientific literature unbelievably useful; much more useful, in fact, than it is even with those journals that are already Open Access today.

¹¹³ Uncorrected Transcript Of Oral Evidence, Monday 8th March 2004
<http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/uc399-ii/uc39902.htm>

RP: *When the publishers gave evidence to the UK Select Committee¹¹⁴ they argued against making scientific information available to the general public, on the grounds that lay people would be unable to understand it. They even suggested it might be dangerous to provide wider access.¹¹⁵ You do not agree?*

HV: That simply isn't true, and is a very self-serving argument on the part of the publishers. When you divide up the public, and think about whom it comprises, you find you are talking about physicians, you are talking about nurses, and you are talking about disease advocacy people; you are also talking about folks who are scientists but who are not connected to the major research institutions that have subscriptions to all the journals. In the US, therefore, the public includes literally millions of people who would like to have access to the scientific literature.

RP: *You are saying that millions of people want to access scientific literature?*

HV: Well, if you ask how many people are going to read a single article that is highly technical in nature, maybe you are only talking about a few people. But that is the case when you consider how many scientists read any scientific article: the number of readers varies from zero to thousands. But certainly in this country when you list the number of people who are interested in the scientific literature — whether they are science journalists, educators, students, or simply people interested in health issues — the numbers are vast, and people do want access.

RP: *As you said, your premise when writing E-BIOMED was that in an all-electronic environment the entire publishing system can be re-structured. When I interviewed Vitek Tracz he said he believes that peer review will become "increasingly irrelevant, ineffective and too time consuming; and, in practice, simply not done." Given the radical views you expressed in E-BIOMED, I am assuming you agree with that?*

HV: I don't see our abandoning peer review. However, I do think that there are ways to enrich discussions about papers that go well beyond peer review.

RP: *Can you give me an example?*

HV: Sure. For example at some point I would like to see PLoS have a way to have comments attached to papers. We want to be sure we don't end up with scurrilous and irresponsible commentary, but having, say, a blog attached to every paper, and allowing a to and fro between the authors and other scientists working in the same area could be a very enriching experience — like having a permanent meeting in which people are standing around a poster and talking about a piece of work, and having conversations. In some cases this could lead to new insights. So there are ways in which we could move the review process into a more public setting, and do so in a profitable and enjoyable way.

¹¹⁴ Uncorrected Transcript Of Oral Evidence, Monday 1st March 2004

<http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/4030101.htm>

¹¹⁵ The managing director of Wiley Europe, John Jarvis, commented: "This rather enticing statement that everybody should be able to see everything could lead to chaos ... [since there is] ... a lot of information in the world which most of us need help with." <http://www.parliament.the-stationery-office.co.uk/pa/cm200304/cmselect/cmsctech/uc399-i/uc39902.htm>.

RP: Making it interactive, and a more transparent process.

HV: Yes. The other component here — as I said earlier — is that it is going to be very important build a digital library of past publications. You may not want to go to go all the way back to 1900 in biology, as biology builds on itself, but something published in chemistry 50 or 75 years ago is probably as true today as it was then. In fact, even in biology we will want to have the important journals going back as far as we possibly can, if only for historical purposes. I'm thinking here about those journals that would account for 95% to 98% of citations — say the 300 leading journals going back 30 years.

RP: That will be difficult to achieve though won't it?

HV: Actually, we have costed it out, and for roughly \$10 million we could capture at least 95% of the literature that has been cited, and put it into a searchable, retrievable digital form. \$10 million is not a lot of money when you consider the amount that you can spend on research over 30 years.

RP: But you are talking about material that is still in copyright. That means you would need the cooperation of publishers wouldn't you?

HV: Yes, unfortunately it would be necessary to work with the publishers. However, while it may be difficult to get some of them to the bargaining table, I would hope that society journals would come around quite quickly, and recognise that there are advantages to their members.

The societies have got to understand that their primary responsibility is to their members, and having access to the scientific literature — especially the literature of their own society journals — at the flick of a mouse would be incredibly useful for society members.

RP: This goes to the point you made earlier about how the Web now enables both search and browse modes?

HV: Absolutely. I gave a talk recently, for example, that involved some retrospective analysis of work done in the late 1970s and early 1980s, most of which was published in the *American Society of Microbiology* journals. Because the ASM has digitised all of its publications and put them into PubMed Central it was remarkable just how easy it was to get the information.

To do that research using journals that are not digitised would have been extraordinarily difficult, particularly now that libraries are cutting back on their space, and on their archive volumes, and putting those they do hold in remote storage. The fact is that doing research using paper resources is now extremely difficult.

RP: Do you think that the race between the publicly-funded Human Genome Project and privately-held Celera to sequence the human genome offers any insights into

*Open Access, or tells us anything about the necessity in modern science of making information more freely available?*¹¹⁶

HV: The precepts of the publicly funded genome project made a very big impression, and it does have an impact on people's approach to scientific information, and their attitudes about access to it. There is also no doubt that virtually everyone working in the field of science that I work in depends very heavily on publicly available annotated sequences of humans, of mice, of worms and many other organisms. That principle is one that heavily informs the Open Access Movement, because it demonstrates how important search and retrieval is. It shows how important it is not just to have pieces of sequence floating around and the literature somewhere else, but to have them actually amalgamated into one site where you can do a search and get the information you want. That is a metaphor that we are trying to develop with the scientific literature.

However, beyond the fact that a public project ultimately achieved a moral victory over a private corporation, the race between Celera and the Humane Genome Project is irrelevant here.¹¹⁷

The end game

RP: What is the end game for PLoS?

HV: There is no end game. Where we would like to be in, say, four or five years is to have a very successful set of journals of the kinds I have described, and enough revenue to feel we are on a very firm financial footing. As I said, we expect to achieve this through authors' fees, some selective advertising, memberships, sponsorships, and philanthropy. We also want to be a leader in making Open Access *the* mode of publishing for all journals in biomedical science.

RP: Traditional subscription-based publishers argue that open access publishing is not a long-term business model. When you gave evidence to the UK Science & Technology Select Committee you said you expected PLoS to be self-funding in two

¹¹⁶ Launched in 1986, The Human Genome Project (HGP) was developed to map the human genome down to the nucleotide (or base pair) level and to identify all the genes present in it. It was coordinated by the US Department of Energy, the NIH and, the UK-based Wellcome Trust. In 1998, an identical, privately funded quest was launched by the American researcher [Craig Venter](#) and his firm [Celera Genomics](#). The \$300 million Celera effort was intended to proceed at a faster pace and at a fraction of the cost of the roughly \$3 billion taxpayer-funded project. Initially the rivals agreed to pool their data, but the agreement fell apart when Celera refused to deposit its data in the unrestricted public database [GenBank](#). Celera had incorporated the public data into their genome, but forbade the public effort to use Celera data. On [14 April 2003](#), a joint [press release](#) announced that the project had been completed by both groups, with 99% of the genome sequenced with 99.99% accuracy.

¹¹⁷ In 2005, in a decision that essentially marked the end of the genome wars, Celera Genomics announced it would release its formerly proprietary human, mouse, and rat genome sequences to the public domain. The decision went into effect on July 1st 2005. <http://www.bioworld.com/newsitems/2005/06-05/06-09-05-news-celera>.

and a half years (which would take us to the end of 2006).¹¹⁸ Is that still the expectation?

HV: Actually, we are reworking the business plan at the moment. It is taking a little longer to become self-funding than I would have wanted, so I don't think it is likely that we will become self-funding in 2006. It will, however, be some time in the next couple of years.

RP: *What changed?*

HV: The estimate I gave to the Select Committee was based on the original business plan we developed in 2003 with the Moore Foundation. It is always hard to predict exactly when things will happen, and progress has been a little slower than we would have liked. But I don't think our experience falls short of what you might expect — almost all start-ups have that kind of lag. Overall we have done very well: the journals have been extremely successful and we are happy with our progress.

RP: *How many journals might PLoS eventually publish?*

HV: In the most extreme situation we would not expect to publish more than about 5% of the total number of papers that appear in biomedical research journals, and there are around 6,000 journals in our field. So we don't plan to cover the whole waterfront. We would, however, like to see all the existing journals convert to Open Access.

RP: *How likely is it that existing journals will all convert to Open Access?*

HV: Well, we are already seeing some good signs. The *Journal of Clinical Investigation*,¹¹⁹ for instance, has in effect become an open access journal. If you look at the way they operate now you will see that there are significant page charges, and while they do still sell subscriptions, the journal is freely available at the point of publication on both PubMed Central, and through the journal's own web site. So you could argue that that journal has already made the transition.

There are other journals like the *Proceedings of the National Academy of Sciences* that offer an open access option: for an extra \$1,000¹²⁰ you can have your article published Open Access instead of waiting six months for the article to appear in PubMed Central. Currently around 40% of PNAS authors are taking advantage of that option.

RP: *And in July 2004 Springer launched Open Choice,¹²¹ which allows authors to have their papers freely available on the Web if they pay \$3,000. A year later Oxford*

¹¹⁸ Uncorrected Transcript Of Oral Evidence, Monday 8th March 2004

<http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/uc399-ii/uc39902.htm>

¹¹⁹ <http://www.jci.org>

¹²⁰ http://www.pnas.org/cgi/content/full/101/23/8509?maxtoshow=&HITS=10&hits=10&RESULTFOR MAT=&fulltext=open+access&searchid=1132745160579_16138&stored_search=&FIRSTINDEX=0&journalcode=pnas

¹²¹ <http://www.springer.com/sgw/cda/frontpage/0,11855,1-40359-0-0-0,00.html>

University Press launched Oxford Open,¹²² with prices ranging from \$1,500 to \$2,880.

HV: So journals are beginning to look to the future; and not just society journals, but for-profit journals too. They are recognising that authors are increasingly going to favour Open Access.¹²³ That means that if they want authors to submit papers to them they are going to have to change the way they do business.

RP: *And they can make a good business out of it?*

HV: Certainly. PLoS is a not-for-profit, but Vitek has pointed out to me many times that it is perfectly possible for publishers to make a profit with Open Access. It may not be a 40% profit,¹²⁴ but it could be 10% — which would be more in line with the way most industries operate.¹²⁵

RP: *The problem for commercial publishers presumably is that this means telling their shareholders that they have got to lower their expectations. What is your advice to these commercial companies?*

HV: They need to realise that they are going to get priced out of the market by open access publishers if they don't change. When authors see how available their papers are in open access journals they are going to largely prefer the open access mode — because researchers want to be read. That is a very, very clear message and publishers have no choice but to respond to it.

Valuable and achievable

RP: *There are now a whole range of "open" movements, including Open Source Software, Open Biology, Open Spectrum, and Creative Commons. The Internet is undoubtedly an important factor in the development of these movements, but there seems also to be a growing debate about the relative merits of proprietary versus non-proprietary modes of doing business. What are your views on that debate?*

¹²² <http://www.oxfordjournals.org/oxfordopen>

¹²³ In addition, in February 2005 Blackwell Publishing announced [Online Open](#); and in May 2006 Elsevier announced that "From May onwards some Elsevier journals will be offering to their authors the option to pay a sponsorship fee to ensure that their article, already accepted for publication, is made freely available to non-subscribers via ScienceDirect."

http://www.earlham.edu/~peters/fos/2006_05_21_fosblogarchive.html#114848427961388480.

¹²⁴ PLoS founders frequently argue that Reed Elsevier earns margins of nearly 40% (<http://www.dailycal.org/sharticle.php?id=13201>). In giving evidence to the UK Select Committee enquiry Crispin Davis explained that the company's profit margin was 34%. He added that he was referring to the gross profit margin "before goodwill, amortisation, tax and so on". He continued: "If you take all of those off, our net reported profit margin is 17%. [and] We invest in excess of £150 million a year on new technologies, functions, ScienceDirect and so on, and that is capitalised over a five to ten year period, but essentially the 17% would be before that investment."

¹²⁵ In fact, Tracz's BioMed Central has yet to achieve profitability. In July 2005, therefore, the company doubled its article processing charge, from \$525 to \$1,400 (£750). See footnote 33, and <http://povnder.blogspot.com/2006/05/interview-with-vitek-tracz.html>.

HV: Look, there are still things that are done on the Internet that make money for people, and there are various ways of doing that: by selling advertising, by selling stocks, by using pay-per-view and so on. That doesn't mean that things should not be on the Internet, just that you need different kinds of business plans for them on the Web.

None of us are saying that the Internet has to be used in a certain way: what we are saying is that for the kind of science that we do, having an open access literature is extremely valuable, and it is achievable.

RP: And beyond the kind of science you do?

HV: I don't follow the debates beyond biomedical research well enough to answer that question in an informed way. However, I am certainly a strong proponent of the kinds of things that Larry Lessig talks about with Creative Commons¹²⁶: the reuse of information, the creation of new ideas by having access to previously published work, the right to reformulate information in ways that will allow us to have new insights — all of us who contribute to discovery want to see this.

RP: You are saying that it is not so much a battle between open and proprietary models, but simply a matter of finding the right model for the networked environment. And the right model may involve making a profit from providing open information services, or in some cases it may mean continuing to sell proprietary information?

HV: Exactly.

RP: Often when governments intervene in order to prevent commercial organisations from enjoying information monopolies they in turn are accused of trying to create governmental monopolies. That was the criticism levelled at the original E-BIOMED¹²⁷ proposal; it was the criticism levelled at the NIH policy on public access.¹²⁸ It is also a criticism¹²⁹ that has been levelled at PubChem.¹³⁰ Do you have views on that?

HV: That really is a false issue, because there is no monopoly here. Monopoly occurs when an organisation owns the copyright or intellectual property on a commodity that can't be accessed any other way. So every journal that operates by obtaining copyrights from its authors is basically operating a monopoly on a set of papers that

¹²⁶ The Lessig interview in this series is at: <http://poynder.blogspot.com/2006/04/interview-with-lawrence-lessig.html>.

¹²⁷ *Just a Minute, Please*, Floyd E Bloom, Science, 9 July 1999.
<http://www.sciencemag.org/cgi/content/summary/285/5425/197>

¹²⁸ *Socialized Science*, Rudy M Baum, Chemical & Engineering News, September 20, 2004,
<http://pubs.acs.org/cen/editor/8238edit.html>

¹²⁹ In a public statement in May 2005 the American Chemical Society claimed that as "a taxpayer-supported resource" PubChem will over time "pose an insurmountable threat to CAS' survival" since it is "a mini-replica of the CAS Registry, and a replica poised to expand." It added: "We believe that taxpayers should not fund the entry of NIH into the information industry more broadly than is necessary to disseminate the information whose creation it funds."
http://www.cas.org/acsnih/acscas_statement.pdf

¹³⁰ PubChem is a free database of chemical structures of small organic molecules and information on their biological activities. <http://pubchem.ncbi.nlm.nih.gov>

are unique, since they aren't going to be published anywhere else. That means that the only way you can see them is by paying for them.

And that is why some of these journals get away with [laughs] highway robbery when they sell subscriptions for \$5,000 or \$20,000 per issue to an institution.¹³¹ Since the only way you can see the information is by buying the journals, then that is a true monopoly.

What PubMed Central is doing, by contrast, is simply providing access to scientific content that — at least if the journals are open access journals — can be used under terms that are spelled out by the Creative Commons licensing procedure. Effectively this says: "Do what you like with it so long as you provide appropriate attributions." This has nothing to do with a monopoly.¹³²

RP: So while copyright will continue to be relevant in an open access environment, it will not be asserted in the proprietary manner of traditional publishers, who generally insist on acquiring copyright in the papers that they publish, and then prevent anyone else from re-publishing them. With Open Access the copyright likely will remain with the author, and it will not be used to prevent others from re-publishing the material?

HV: Well, copyright does exist, and PLoS allows the authors to hold on to it — they just deed the article to the journal under the terms of a Creative Commons licence.

RP: If you look back to that day in 1998 when you sat in a San Francisco coffee house with Pat Brown, and consider how the Open Access Movement has developed, is there anything you would do differently if you could start over?

HV: Questions like that are always hard to answer. I guess I might have started out in a more conservative mode! On the other hand, that might have been wrong: perhaps people needed to be bolted awake by the potential of open access technology. I don't know. What I have learned, however, is just how much effort it takes to convince the scientific community to make a change in their publication practices.

RP: And that was a big surprise to you?

HV: Well, there is an irony here. Most people when pressed understand how important publication is to a scientific career. In fact, the celebrations of a scientific life are frequently more festive when a paper has been accepted by a leading journal than it is when the discovery itself is made.

¹³¹ One of the early drivers for Open Access was the extremely high price inflation that STM journals have undergone since the war. In some cases, critics have pointed out, the annual subscription is now as much as a diamond ring or a new automobile. The example most often cited is the cost of Elsevier's *Brain Research* — which at the time of writing is available on Amazon for an annual subscription of \$23,622.65.

<http://www.amazon.com/exec/obidos/tg/detail/-/B00007KXQW/103-1286164-5675819?v=glance>

¹³² PLoS utilises CC Attribution Licence 2.5. This allows others to copy, distribute, display, and perform the work, to make derivative works, and to make commercial use of the work — so long as the original author is credited. <http://creativecommons.org/licenses/by/2.5>.

But the point that deserves a lot of emphasis is that publication is part of the process of doing science. The reason I emphasise that is because there are many people in the world — including some scientists — who, although they know that publication is part of the scientific process, make the assumption that money that goes to the doing of science in laboratories should not be used for publication. The reality, however, is that the act of writing up your work and publishing it is not only part of the discovery process, but it is also an essential feature in the scientific process taken more broadly.

RP: This goes to the issue of needing to persuade funders that when they award research grants they should permit grantees to use some of that grant to pay to publish the results of the research in OA journals?

HV: That's right. After all, if we don't publish our work it is quite useless to anybody else. One of the ways in which science works is to have discovery confirmation, and we also need to enable others to extend and build on earlier research in an orderly way. That can only happen by communicating findings. And this means that the publication process needs to be very tightly interwoven into the whole process of doing science.

RP: Why is it taking so long, and proving so hard, to make Open Access a reality?

HV: You have to take a longer-range view on this. In fact, if you start the clock not from the opening of PubMed Central — which is not really Open Access — but from the launch of the first BioMed Central or PLoS journals, you can see that we have made a tremendous difference in just a few years.¹³³

From that perspective events are actually moving reasonably quickly, and by the end of the decade I believe the world is really going to look very different. At that point it will have been ten years from my politically naïve E-BIOMED [laughs] statement, during which we will have seen a transformation of the biomedical research literature. When you think of it like that it's not slow at all. Things just always seems slower when they are happening than they do in retrospect.

RP: What has been your biggest disappointment over the way that OA has developed?

HV: One of the things I have been most disappointed about is the relatively small number of scientific societies that have been supportive of Open Access. To my mind that says there is something seriously wrong with the way in which scientific societies operate — and is an issue I have been troubled about over the last several years.¹³⁴

RP: What is the ultimate goal. Are you looking for 100% of all research to be freely available on the Internet?

¹³³ Other OA advocates would date the birth of the Open Access Movement much earlier. Ginsparg's [arXiv](#), for instance, was established in 1991, and Stevan Harnad's *Subversive Proposal* was first published in 1994. For an overview of the development of the Movement see [Peter Suber's Open Access Timeline](#).

¹³⁴ The most egregious example of the negative response to OA by learned societies is probably the attitude of the Royal Society. See for example: <http://www.scidev.net/News/index.cfm?fuseaction=readNews&itemid=2498&language=1>.

HV: Certainly 100% of the biomedical research. But it will be different in different fields. What happens depends on the culture and the economics of the field in question. While my focus is the biomedical literature, I am also interested in chemical information. Chemistry, however, is going to be a harder nut to crack: we are having trouble getting the chemists to see the virtues of Open Access.

RP: And if you consider the situation in the humanities and social sciences it looks different again I guess?

HV: Right. In other fields there are going to be different ways to operate. The social sciences, for example, have a very different funding structure, and it is not yet possible to imagine supporting all the social science journals with authors' fees alone. That suggests that subscriptions will continue to be sold.

RP: I'm curious about your motivation for supporting Open Access so vigorously. I can see that when you were director of NIH you may have felt that you had a responsibility to look for better ways to communicate research. But your decision to continue supporting the movement after leaving NIH, and to do so with such dedication, goes beyond the call of duty doesn't it?

HV: I believe that science is one of those activities that improves the state of the world, and once you realise how important publication is in the series of acts that constitutes the doing of science, and once you understand the incredible transformation of that publication process that the Internet, and software, and the whole digital world, now promises it is hard not to be pretty passionate about trying to make that part of the scientific universe work more effectively.

RP: You also have strong views on stem cell research, which is another issue that came to the boil while you were at the NIH, and which you have continued to ... ?

HV: Frankly it was a live issue when I arrived at the NIH, although we didn't actually have any stem cells at that point. But what has that to do with open access publishing?

RP: I was hoping to talk about your interest in Open Access in the context of other areas of your life, both professional and personal.

HV: I don't think we have time for that. I have spoken on and written on the topic of stem cells in so many places. You will see from the many public pronouncements I have made, and the hearings I have attended, that I am a supporter of stem cell research. But this is irrelevant to the Open Access Movement.

RP: You are also involved with the Campaign to Defend the Constitution (DefCon),¹³⁵ an online grassroots movement aimed at combating the religious right. One could perhaps argue that there is a connection here with Open Access....?

¹³⁵ The DefCon website describes the organisation as "an online grassroots movement combating the growing power of the religious right. We will fight for the separation of church and state, individual freedom, scientific progress, pluralism, and tolerance while respecting people of faith and their right to express their beliefs." <http://www.defconamerica.org>

HV: I am not aware of any connection. Let's just stick to the topic. All these other things take time to discuss, and they would be given short shrift if we tried to squeeze them into this conversation.

RP: *OK, let's leave it there then. Thanks for your time.*

© 2006 Richard Poynder

This interview is published under the terms of the Creative Commons Attribution-Non-commercial-No-Derivatives Licence (<http://creativecommons.org/licenses/by-nc-nd/2.5/>). This permits you to copy and distribute it as you wish, so long as you credit me as the author, do not alter or transform the text, and do not use it for any commercial purpose.

If you would like to republish the interview on a commercial basis, or have any comments on it, please email me at richard.poynder@journalist.co.uk.

Please note that while I make this interview freely available to all, I am a freelance journalist by profession, and so make my living from writing. To assist me to continue making my work available in this way I invite anyone who reads this interview to make a voluntary contribution. I have in mind a figure of \$8, but whatever anyone felt inspired to contribute would be fine. This can be done quite simply by [sending a payment](#) to my PayPal account quoting the email address aotg20@dsl.pipex.com. It is [not necessary](#) to have a PayPal account to make a payment.

I would like to acknowledge the help of the [Open Society Institute](#), which provided a small upfront grant to enable me to get started on *The Basement Interviews* Project. Further information about *The Basement Interviews* can be found at the [Open and Shut?](#) site.