

---

DAVID BALTIMORE

## Baltimore's Travels

---

*How a routine scientific dispute literally became a federal case but revealed important lessons about the policing of science.*

---

For the past three years, I have been at the focal point of an Investigation into the accuracy of a scientific paper (published April 1986 in the journal *Cell*) of which I was a coauthor. The investigation, which wound up before a subcommittee of the U.S. House of Representatives and was reported on the front pages of newspapers, continues to this day. My purpose in writing this article is to explain how a routine scientific dispute came to be adjudicated in such an unusual setting, how we chose to defend ourselves, and the scientific and political issues involved. One central issue was apparent to me from the beginning, and it remains central today: Who should judge science? Along the way I came to identify several other issues:

- What is the worth of scientific collaborations, and what are the duties and responsibilities of those involved in them?
- Is error in science inherently bad and should it always be corrected where found?
- How does one distinguish between error and fraud?
- And does science do an adequate job of policing itself?

### **The background**

In the early 1980s, I began a series of experiments to explore the production and control of antibodies—proteins made by immune cells that participate in the recognition and destruction of invading microorganisms. The experiments employed an exciting new technology, the insertion of new genes into mouse embryos to produce “transgenic mice.” The work required the collaboration of several laboratories with expertise in different areas of molecular biology and immunology.

Our initial goal was to produce transgenic mice carrying a new antibody gene. The experiments worked quite well and were described in two papers that predated the *Cell* paper involved in the current controversy. Studies of the expression of the inserted “transgene” provided new insights into the regulation of antibody production. A next step in this work was to examine how the new antibody gene affected the activity of natural antibody genes. For these experiments, I approached Professor Thereza Imanishi-Kari at the MIT Cancer Center. We designed a two-pronged

study: Her laboratory would analyze the antibody proteins produced by the transgenic mice, and my laboratory would characterize the active antibody genes. One purpose of this parallel design was to provide an internal mechanism for cross-checking results. For example, if the protein analysis—always a tricky and difficult task—should give inconsistent results, the genetic analysis might reveal the source of the problem.

One key observation was that the inserted antibody gene had a far greater influence than previously expected on antibody production by the resident mouse genes. Much of our analysis focused on the region of the antibody molecule that enables each antibody to recognize a specific foreign structure; scientists refer to this region as having an “idiotype.” Our previous studies had shown that many immune cells from the transgenic mice were secreting antibodies with idiotypic regions matching those of the new gene. We had assumed that these antibodies were products of the transgene, but the new experiments told a different story. They showed that most of the relevant antibodies were derived from normal antibody genes inside the mouse cells. Through some unknown mechanism, the presence of the inserted gene had induced the cells to adopt a particular idiotypic pattern.

The *Cell* paper describes two possible explanations for this mimicry. It is important to recognize, however, that the paper and the experiments it describes are just one link in the chain of scientific evidence. We fully expected that other laboratories would extend the experiments into new directions and publish their own interpretations. Each new effort would represent a small step forward in understanding how the immune system responds to external challenges. Future studies might offer new insights into derangements of the immune system that cause autoimmune diseases such as arthritis, lupus, and certain life-threatening anemias.

### **The paper is challenged**

At the outset, the substance of the dispute was not unlike others that occur regularly in biology labs. It was simply a disagreement over scientific matters between two scientists: Professor Thereza Imanishi-Kari, now at Tufts University, a leading immunologist who specializes in the analysis of antibodies; and Dr. Margot O'Toole, a postdoctoral researcher on a one-year appointment in the Imanishi-Kari laboratory, who was not a participant in the study but worked on a related question.

Dr. O'Toole had reviewed the draft of our paper, made a number of helpful suggestions, and indicated no concerns prior to publication. Shortly after the paper was published, however, O'Toole expressed reservations in a most unusual fashion: Rather than discuss her objections with the authors (the usual procedure), she complained to an outside authority—not MIT, which employed both O'Toole and Imanishi-Kari, but Tufts University, which was considering whether to appoint Imanishi-Kari to its faculty. The gist of her complaint then—it has changed from time to time over the intervening three years—was that the conclusions of the paper were not supported by the scientific evidence it presented.

Tufts appointed a panel to review O'Toole's charges about the paper and also appointed a special committee to review Imanishi-Kari's qualifications and reconsider her appointment. The scientific review panel, headed by Tufts professor Henry Wortis, found that O'Toole's complaints involved matters of interpretation and, because the paper's central thrust was unaffected, did not recommend any corrective action. Following its two reviews, Tufts appointed Imanishi-Kari to a tenure-track position.

O'Toole then took her complaints to MIT, which appointed Professor Herman Eisen, an eminent immunologist, to review them. He found that she had made an appropriate expression of scientific concern, but the errors he observed in the paper were not of sufficient importance to require correction. Further, he concurred with the Wortis panel's decision that although O'Toole's interpretation might differ from that of her supervisor, the paper's conclusions were sound.

In these reviews, completed by early summer of 1986, all the issues were scientific. No one had accused anyone of unethical or criminal behavior; O'Toole simply said she thought the conclusions of the paper were not borne out by the data generated in the Imanishi-Kari laboratory. After the second review, I thought that the matter was closed.

Meanwhile, Charles Maplethorpe, who received his Ph.D. for work done with Dr. Imanishi-Kari but who had left her laboratory before the paper was published and had not been involved in the study, got himself involved. Maplethorpe contacted two scientists at

---

*The words “Secret Service” had terrifying import and carried the implication that we must have done something awful.*

---

the National Institutes of Health (NIH), Walter Stewart and Ned Feder, who had made reputations for themselves by publishing papers analyzing cases of previously demonstrated fraud in science. They received from O'Toole copies of 17 pages of laboratory notes taken from the Imanishi-Kari laboratory. These pages, of more than a thousand pages collected during the study, included data from a number of failed experiments. On the basis of the 17 pages, plus conversations with O'Toole and Maplethorpe, Stewart and Feder mounted a challenge of their own.

The Stewart and Feder challenge soon developed into a *cause célèbre* because of the manner in which they conducted it. First they wrote a lengthy manuscript clearly charging that our paper was consciously misleading. The Stewart-Feder manuscript was submitted to a number of journals, all of which rejected it. Frustrated by their inability to publish what journal editors told them was not a scientific article that could be refereed, Stewart and Feder went public. They circulated the manuscript widely to scientists, asking for comment. They also began speaking about their “investigation” on university campuses and at scientific meetings and offered to send a complete file of correspondence to anyone asking for it.

That file contained a large number of letters and memoranda they had exchanged with me, my fellow authors, other scientists, editors, and NIH officials. It was a highly selective file, containing, for example, little or nothing of the correspondence that by this time they were having with Margot O'Toole and with staff members of congressional committees. In much of this correspondence, and in the speeches of Stewart, my name and those of my colleagues were frequently juxtaposed with the names of people who have been found guilty of fraud. In other words, Stewart and Feder utilized a classic propaganda technique—guilt by association. Because of their hostile acts, I asked the NIH to conduct a formal review to determine the accuracy of the *Cell* paper.

This activity by Stewart and Feder continued throughout 1987 and early 1988, at which time I learned from a newspaper reporter that two congressional committees investigating fraud and misconduct would shortly hear from Stewart and Feder about our paper. One committee was the Oversight and Investigations Subcommittee (O&I) of the House Energy and Commerce Committee, chaired by Congressman John Dingell (D-Mich.); the other was the Human Resources and Intergovernmental Relations Subcommittee of the House Government Operations Committee, chaired by Congressman Ted Weiss (D-N.Y.).

We were not notified of these hearings nor were we permitted to answer the charges against us. The Weiss subcommittee hearing referred only obliquely to our dispute with Stewart and Feder and did not mention by name the authors of the paper. At the O&I hearing, however, my name was used liberally and the words “fraud,” “misconduct,” and “misrepresentation” were casually bandied about. Although I was not personally accused of having committed fraud, the implication was clearly made at the hearing that at least one of the authors had.

Meanwhile, NIH (which funded the research) was convening a panel, consistent with my earlier request, to review the dispute. The panel's start was delayed by inappropriate appointments, however—two of its initial members had coauthored papers with me—and was slow getting under way.

### **Defending against attack**

In the wake of the congressional hearings of April 1988, it became clear that my own reputation and those of my coauthors were being sullied by continuing acrimonious attacks from Stewart and Feder, who had allied themselves with Peter Stockton, a policy analyst on the staff of the O&I subcommittee. Soon after, Stewart and Feder were detailed from NIH to work with Stockton. Clearly, the authors needed to mount a defense.

For a time we got by with just my own staff, consisting of the administrator in my office and a part-time public relations person who does work for the Whitehead Institute, plus assistance from my personal

lawyer. Before long it became apparent that we needed help from people familiar with the procedures of Congress and the workings of Washington politics. We retained Washington counsel and spoke with others knowledgeable in the ways of that city.

We learned something about the nature of such congressional investigations: The effort to resolve the dispute would not be primarily a scientific or a legal one. Rather, it would be largely directed at press coverage. We would not have a “day in court” in the sense of being informed in advance of any charges or complaints about our work. We would not have the opportunity to confront and question our accusers, to present our own evidence or witnesses, or have any assurance that a record would be developed and a decision based on it. There would likely be no final ruling from a congressional panel; instead, we would be judged in the “court of public opinion.”

I was advised that although the alleged errors (or fraud or misconduct, as they came to be referred to) did not take place in my laboratory, I was nevertheless the real target of Stewart, Feder, and Stockton because of my visibility. Thus I agreed to lead the defense. One of my first defensive actions was to write a letter to about 400 of my scientific colleagues, acquainting them with my side of this matter. This action was reported in some newspapers.

Meanwhile, the third scientific review of the *Cell* paper was in progress. The NIH Review Panel deliberated through the last half of 1988. Like the two reviews before it, the NIH panel reported that there were errors but the findings of the paper were sound. Fraud, misconduct, or misrepresentation were explicitly rejected. Unlike the other two reviews, this one suggested corrections in the journal. Although we disagreed with the panelists—as did *Cell's* editor—corrections about the errors were submitted and published.

Meanwhile, unbeknownst to us, Stewart, Feder, and Stockton had asked the Secret Service to examine laboratory notes, most of which had been subpoenaed from the authors. We learned of this accidentally—and only a couple of weeks before the next hearing. Again, our source of information was a newspaper reporter, John Crewdson of the *Chicago Tribune*, who warned some scientists that they should not defend us because the Secret Service had discovered incriminating evidence showing that data had been tampered with.

We were, of course, annoyed that the involvement of the Secret Service was not shared with us, but was known to a small number of newspaper reporters. Further, Crewdson's reaction was typical of the “guilty until proven innocent” attitude displayed by some reporters and by some of those directly involved in the investigation. The mere mentioning of the words “Secret Service” had terrifying import, and carried the implication that we must have done something awful to warrant its involvement. We also started to hear rumors that the O&I subcommittee would soon hold hearings (official notification came just a week before they began).

In the week prior to the hearing, the subcommittee and the Secret Service did allow us to see some of their findings, although they gave us no written report (and, to this writing, no report has appeared). Only a few days before the hearing did we get a few sheets of outlined results of their inquiry. Once we saw the nature of their analysis, it was evident that they had uncovered no secrets and no surprises: They had found that experimental notes were not in sequence, implying poor maintenance of records from experiments. They also found that one illustration was a composite of several photographs. The way the composite was represented by the Secret Service, it sounded as though this accepted and common practice was meant to be misleading, which it wasn't.

A simple inquiry to the authors would have elicited the information it took them nine months of high-tech analysis to gather. Their failure to ask us created an atmosphere in which it seemed that we had something to hide, which we didn't. But a news reporter learning of this development, and not having a period for reflection and further inquiry, could believe that the Secret Service report had uncovered a smoking gun. We had to help educate the press before the hearing, but we did not want to be the source of a leak.

True to form, the subcommittee staff had apparently started a leak; we deduced this because we had earlier heard from others about the material we'd been shown. Thus we felt free to discuss the Secret Service information with the press before the hearing. This was very useful because when reporters heard both sides, they could put things into perspective.

Going into the hearing, we decided that because the charges against us were voluminous and constantly

---

*If the investigation shows nothing else, it is that uninformed or malinformed outsiders cannot effectively review the progress of scientific activity.*

---

shifting, we needed to individualize our defenses. I concentrated on the effect such congressional reviews might have on science. Dr. Imanishi-Kari, whose work was under direct attack, responded in a personal manner. She focused on the Secret Service findings implying that she was a messy note-taker, which she readily admitted. She also addressed the question of motive, and in dramatic fashion: What possible reason, she asked the congressmen, could she have to cheat? The most frequent citations of the *Cell* paper were by medical researchers investigating autoimmune diseases, particularly lupus. This often-fatal disease took the life of her sister, and Imanishi-Kari herself, she revealed, is likewise afflicted. If she were to deceive lupus researchers, it would be self-destructive.

### **Who judges science?**

In my defense, I tried to show how this sort of procedure could seriously harm the functioning of the remarkably successful American biomedical enterprise. One of the reasons that biomedical research has progressed so well in this country is that we have an efficient and effective system based upon peer review. Scientists receive grants only after their proposals are reviewed by other scientists who are expert in their field. The results are likewise reviewed by peers, first when the paper reporting on the work is offered to a refereed journal, and later, after publication, when other scientists verify the results.

This verification process is the cornerstone of American science, which Congressman Dingell, as the son of one of the congressmen who helped to found NIH and brother of a biomedical researcher, clearly wants to protect (while also ensuring that the public gets its money's worth). I would like to help him in this effort, and I am sure that most other scientists would also like to help in an honest examination of the process by which science is reviewed and verified. But fundamentally, I do not believe it is possible to evaluate science in any other way.

The peer-review system came under attack from the moment O'Toole took her accusations outside the university-NIH review process and handed it over to Stewart and Feder, and then on to Stockton. These three unqualified outsiders tried to impose their own review upon the scientific ones that had preceded, even suggesting that the ways in which scientists take their notes need to be regulated to make "auditing" such as their own more convenient. If the sad history of this investigation demonstrates nothing else, it shows that uninformed or malinformed outsiders cannot effectively review the progress of scientific activity.

In a more appropriate manner, our immunological study was followed up with complementary work by, among others, Professors Leonard and Lenore Herzenberg of Stanford University, who found as we had that many cells in the experimental system we studied did not express the transgene. They also found "double-producers," cells that expressed the transgene as well as the "natural" gene, which we did not find. There are reasonable explanations of this discrepancy, a type of refinement that often occurs in science—indeed, it is at the very heart of the scientific process. But it is important to keep in mind that without our study, the Herzenbergs would have been unlikely to do theirs; in turn, their work amplified ours.

### **The nature of collaboration**

Another pillar of American science affected by these proceedings was collaboration. Questions were raised repeatedly as to whether I as a coauthor had responsibility for the work performed in a collaborating laboratory and whether I had faithfully discharged that responsibility.

When two laboratories decide to collaborate, two or more professors make the decision and maintain the formal link. They communicate periodically to check on progress, new ideas, or new directions, and from time to time there may be joint meetings of the whole team to review the research. But it is the day-to-day contact between trainees, who are supervised independently, that generally makes the science progress. When writing up results, this network of interactions continues while the draft manuscript goes back and

forth, and the final manuscript is an amalgam.

Like many other researchers, I have been involved in numerous scientific collaborations, which are essential for bringing complementary skills to bear on important problems. Trust is at the heart of successful collaboration, and this one was no different. If the scientists in a collaboration are knowledgeable enough to judge the details of each other's experiments, there is little purpose in collaborating. I fear that the type of investigation to which we have been subjected, because of its intimidating nature, could make scientists more wary of collaborating; it could thus undermine an important method for advancing knowledge.

### **Error versus fraud**

The errors that have been identified in the *Cell* paper were inconsequential to the conclusions; had this paper not become the focus of such intense scrutiny, the errors would not have been noticed by anyone other than perhaps immunologists seeking to expand upon the results. But now, because of the spotlight, the nature and role of errors have become of increasing concern.

In this connection, it is important to remember that no study is ever complete. The last word is never written. Investigators make the somewhat arbitrary and personal decision to write a paper when they believe that a story—one that makes sense and that others will want to read and build on—can be told. The scientific literature is a conversation among scientists; we describe to each other what we have done, and we implicitly ask our colleagues to try to build on our results, testing their value and their ability to support a growing edifice.

In fact, a scientist's first response upon reading a new paper may be to see a new interpretation, to conceive a different experiment that could provide a wholly new thrust for the paper. This is a thought process that is encouraged in all young scientists, and it is one of the driving forces in science.

I should emphasize that the errors of which I speak must be distinguished from fraud, which involves the conscious misrepresentation of data or the conscious use of data that do not emanate from a described experimental setup. The word "conscious" is crucial here because *unconsciously* either of these forms of error could be made and should not be considered fraud. If as a result of this type of investigation American scientists become afraid to make a mistake, then all of science will be slowed, to everyone's detriment.

### **Policing science**

None of the issues discussed here stands alone, and certainly the policing of science, which means the policing of *scientists*, is related to all the others. Although I believe that errors—and even purposely altered data—will ultimately be discovered by scientific peers, there is obviously the need for an alternative mechanism: One scientist who suspects wrongdoing on the part of another should be able to take appropriate action.

We are currently developing a mechanism at the Whitehead Institute that will go beyond the requirements of a draft regulation issued by the U.S. Department of Health and Human Services. At this point, the procedure is as follows:

- A question about the possibility of scientific misconduct can be raised by anyone in an entirely confidential manner.
- Once a question has been raised, the director appoints a committee of inquiry, composed of appropriate and knowledgeable people; selections are made confidentially and with consideration of real or apparent conflicts of interest. The committee must conclude its work within 60 days, submitting documentation and a written report.
- If evidence of misconduct is found, the director immediately names a committee of investigation to conduct a thorough and authoritative review of the evidence within 120 days. On the strength of the committee's recommendations, the director can take action ranging from reprimand to termination of employment and tenure. Likewise, if the committee finds that the complaint was intentionally dishonest and malicious, the director can take similar action against the accuser.
- Both at the outset of an investigation and after it has been concluded, funding agencies are fully informed.

I believe that the developing Whitehead policy adequately protects the rights of both accused and accuser while providing an effective and speedy procedure for dealing with fraud or misconduct. The policy

delineates in no uncertain terms the duties of each person in the Institute for upholding ethical standards, and it keeps adjudicatory responsibility in-house, where it belongs.

Because NIH has ultimate responsibility for the use of its grant funds, it too needs a mechanism for considering charges of misconduct. NIH should determine whether proper procedures have been followed, but not second-guess appropriately constituted reviews. It should also decide what monetary restitution is needed. If it finds that a local review was procedurally faulty or if it believes that only a national review can be truly objective, it should then take over the process.

If the requests of whistleblowers for reviews are kept secret, and the proceedings and verdict published only if significant error or fraud is found, protection of the rights of the accuser and the accused might be more likely. This protection is built into the Whitehead policy, and it should apply universally. Without it, others could be subject to public attacks of the sort directed against me by Stewart, Feder, and Stockton (who need to reflect on the meaning of due process and the American tradition of protecting the rights of the accused). But unlike me, others may lack the resources or the resolve to successfully fight back. They deserve protection. The progress of science depends upon it, and American society should insist upon it.

The scientific community needs to deal promptly and severely with cases of fraud. We have often been loath to do so, perhaps because as honest men and women we not only don't know much about it but are uncomfortable contemplating it. The growing belief that this attitude must change could be a positive result of the publicity surrounding the *Cell* paper investigation.