KIMIMASA MAYAMA/EPA/ALAMY

CELL-INDUCED STRESS

As a much-hailed breakthrough in stem-cell science unravelled this year, many have been asking: **'Where were the safeguards?'**

BY DAVID CYRANOSKI

It seemed almost too good to be true — and it was. Two papers[1,2] that offered a major breakthrough in stem-cell biology were retracted on 2 July, mired in a controversy that has damaged the reputation of several Japanese researchers.

For scientists worldwide it has triggered painful memories of a decade-old scandal. In February 2004, South Korean researcher Woo Suk Hwang announced that he had generated stem-cell lines from cloned human embryos[3], creating a potential source of versatile, therapeutic cells that would be genetically matched to any patient. A frenzy of excitement followed this and a subsequent publication[4], but that didn't compare with the media firestorm when the results were revealed to be fabricated. The two main cloning papers were retracted[5], and the careers of some dozen scientists were devastated.

In the soul-searching that followed, 'research integrity' became a hot topic, scientists re-evaluated the responsibilities of authorship, and institutions vowed to improve the way that they police their staff. *Nature* and other journals also made promises, saying that they would vet manuscripts more thoroughly. In an Editorial at the time, *Nature* wrote[6]: "Keeping in mind the principle that extraordinary claims require extraordinary proof, *Nature* may in rare cases demand it."

A year later, when Shoukhrat Mitalipov of the Oregon Health & Science University in Portland claimed to have cloned embryonic-stem-cell lines from monkeys[7], *Nature* required independent tests to verify that the lines came from the monkey donors. This verification was published alongside the cloning paper[8]. "I applaud what they did," says Alan Trounson, the outgoing president of the California Institute for Regenerative Medicine in San Francisco, who helped with the testing.

Then came Japan's stem-cell case. This January, Haruko Obokata, a young biochemist at the RIKEN Center for Developmental Biology (CDB) in Kobe, Japan, reported in *Nature*[1,2] that she had converted mouse cells to an embryonic-like state merely by subjecting them to stress, such as physical pressure or exposure to acid (see *Nature* **505,** 596; 2014). The process, labelled stimulus-triggered acquisition of pluripotency (STAP), was so contrary to current thinking that some scientists said they accepted it based only on the reputation of Obokata's co-authors, who were some of the most trusted names in stem-cell research and cloning.

But the paper[1] that set out the fundamental technique was soon shot full of holes. There was plagiarized text in the article. Figures showed signs of manipulation, and some images

© 2014 Macmillan Publishers Limited. All rights reserved

Haruko Obokata tearfully faces the media after she was found guilty of misconduct in April.

were identical or nearly identical to those used later in the same paper and elsewhere to represent different experiments. More damning were genetic analyses that strongly suggested the cells were not what they were purported to be. And although deriving STAP cells was advertised as simple and straightforward, no one has yet been able to repeat the experiment.

Within the space of six months, Obokata was found guilty of misconduct by her institution; well-respected scientists, including RIKEN head Ryoji Noyori, bowed their heads in apology; and both papers were retracted[9]. In the end, the evidence for STAP cells seemed so flimsy that observers began to ask where were the extra precautions and the 'extraordinary proof' that had been promised post-Hwang.

The case has reopened difficult questions about the quality of research and peer review, and the responsibilities of co-authors, institutions and journals. It is also making its mark as an example of how not to do things. The episode has already become a "parable in my lab for teaching students about scientific ethics", says Jeanne Loring, a stem-cell biologist at the Scripps Research Institute in La Jolla, California.

In this article, the news team at *Nature* — which is editorially independent from the journal team that reviewed and published the STAP papers — attempts to find out what went wrong and what can be learned from the case.

CAUGHT IN ISOLATION

The STAP saga has its roots in a contentious hypothesis more than a decade old. In 2001, Charles Vacanti, an anaesthesiologist at the Brigham and Women's Hospital in Boston, Massachusetts, said that he had found "spore-like cells" in virtually every type of mammalian tissue[10]. According to Vacanti, these cells were pluripotent — that is, they could develop into any type of cell in the body — and seemed to lie dormant until activated, perhaps by injury or disease, to regenerate tissue.

Vacanti told *Nature*'s news team in January that by 2006 his laboratory could grow the cells in large numbers, but that they still "were not exceptionally well characterized by us". That is, the team had not demonstrated pluripotency. This was a job he gave to Obokata, a graduate student who had joined his lab in 2008.

Proving pluripotency is often done by injecting cells into a developing mouse embryo — creating a 'chimaera' — and tracking their fate. It is a difficult experiment, and Obokata needed help. "I was looking for the god's hand of chimaeric-mouse generation," she said back in January. A Google search led her to famed mouse cloner Teruhiko Wakayama at the CDB, whose lab she entered in 2011 as a visiting professor. After hundreds of failures to get cells derived from adult mice to show up in chimaeras, she and Wakayama switched to newborn mice as the source of the cells — and the process worked.

By that point, both Vacanti and Obokata were convinced that the stress of the isolation process was creating the pluripotent cells. Obokata said that the idea had come to her while she was taking a bath and reflecting on the stress in her own life.

In the experiments at RIKEN, she used acid to stress spleen cells from newborn mice, and

> "THERE HAS TO BE **CONTROL**, BUT ALSO **TRUST** IN SCIENCE, OTHERWISE THE SYSTEM BREAKS DOWN."

she carried out further experiments to characterize their conversion with Yoshiki Sasai and Hitoshi Niwa, two highly regarded stem-cell biologists at the CDB. With the two key characteristics of STAP cells now demonstrated — they were pluripotent and were created using stressful conditions — she had enough data to publish two papers in *Nature* on 30 January[1,2].

Obokata became an instant celebrity in Japan, where the media picked up on details such as the 'Moomin' cartoon stickers on her lab equipment and the traditional Japanese cooking apron, given to her by her grandmother, that she wore instead of a lab coat.

But within weeks, anonymous observers began noting mistakes in the papers, including evidence of image manipulation, duplications and plagiarism (see go.nature.com/e4dwry). Researchers also started to report that they could not reproduce the supposedly simple experiment.

On 1 April, a RIKEN investigative committee concluded that Obokata had committed scientific misconduct. She maintained that the results were real, but one by one her co-authors withdrew support for the findings. In principle, *Nature* retracts articles only when all co-authors agree, although in rare cases papers can be retracted even if one or more of the authors dissent. In June, Obokata relented and agreed to retract both papers (see go.nature.com/wsfox5). She has not responded to multiple requests for interview since April. She has, however, been invited to participate — under surveillance — in ongoing efforts at RIKEN to verify the original findings.

Should the papers have been published in the first place? Critics have argued that many of the flaws could have been identified beforehand by *Nature* — the easiest, in theory, being a 17-line passage that was taken almost word for word from a 2005 paper[11] by another group.

To detect signs of plagiarism, most journals use a service called CrossCheck. It can compare a submitted manuscript with some 40 million published articles from around 100,000 titles, looking for text matches.

Nature editors did use CrossCheck and did not find the match. But the journal from which the text was lifted, *In Vitro Cellular & Developmental Biology — Animal*, had not been indexed by the service at the time. "Although the databases are very large and growing, there are limitations," explains Rachael Lammey, a product manager at CrossRef in Oxford, UK, which provides the CrossCheck service. Such misses get flagged "a couple times a year", she says, but there is no way to know how many instances of plagiarism fall through the cracks.

Moreover, identifying the match probably would not have halted publication. Many instances of copied text do not constitute plagiarism and just require citation of the original source. Indeed, the RIKEN investigative committee concluded that the passage — a methodological description — should have cited the original, but that the failure to do so was not misconduct.

The committee was more vexed by instances of manipulated and duplicated images in the STAP papers. Obokata had spliced together gel lanes from different experiments to appear as one. And she had used an image of cells in a teratoma — a tumorous growth that includes multiple types of tissue — that had also appeared in her PhD dissertation. The captions indicated that the image was being used to represent different types of cell in each case. The committee judged that in both instances, although she might not have intended to mislead, she should have been "aware of the danger" and therefore found her guilty of misconduct. Obokata claimed that they were mistakes and has denied wrongdoing.

PICTURE IMPERFECT

Image manipulation and duplication within the same manuscript can be detected, and journals are increasingly checking for such problems. Jana Christopher analyses images in every manuscript before they are accepted by EMBO Press, a journal publisher based in Heidelberg, Germany. She uses a set of automated adjustments created for the image software Photoshop by the US Office of Research Integrity that change attributes such as contrast and colour to make manipulations easier to spot (ori.hhs.gov/actions).

At the behest of the chief editor of *The EMBO Journal*, Bernd Pulverer, Christopher ran tests on the STAP papers without knowing their background. She spotted three problems:

© 2014 Macmillan Publishers Limited. All rights reserved

the gel manipulation that was ultimately attributed to misconduct, a seemingly innocent duplicated image mistake and a composite image of cell colonies most probably done to save space.

"The aberrations we saw are fairly typical," says Pulverer, who reports that around 20% of scanned manuscripts have been found to have such issues since the journal started looking for them in 2011. *The Journal of Cell Biology* (*JCB*), published by Rockefeller University Press in New York, has been systematically scanning figures and images in all accepted papers since 2002 and finds about the same rate.

But the journals do not immediately consider a problematic image fraudulent. Spliced gel lanes, for example, are often attempts to present data more clearly and concisely. In most cases, these manipulations are done naively, to create a 'prettier', more informative image, says Pulverer. But, he says, "in cases where we don't obtain plausible explanations and source data for the figure in question, we dig deeper". At the *JCB*, acceptance is revoked for about 1% of papers, according to the journal's executive editor, Liz Williams.

Such scanning methods are far from foolproof. Some worry that alerting authors to problems with their images allows would-be fraudsters to improve their forgeries. And although manipulated images might be easy to spot, it is harder to identify duplications, especially when they come from other articles. "Cross-literature comparisons would require very high-powered search algorithms and probably a supercomputer," says Pulverer. "This has been discussed for a number of years but never moved forward." In the STAP case, current image-checking procedures would not have caught the problem with the teratoma image — or several other problems with the main paper[1] that surfaced under closer scrutiny.

Philip Campbell, editor-in-chief of *Nature*, says: "We have concluded that we and the referees could not have detected the problems that fatally undermined the papers." But scientists and publishers say that catching even the less egregious mistakes raises alarm bells that, on further investigation, can lead to more serious problems being discovered.

Many say that the tests should be carried out on all papers. Christopher says that it takes about one-third of her working week to check all accepted manuscripts for the four journals published by EMBO Press. At *Nature* and the *Nature* research journals, papers are subjected to random spot-checking of images during the production process. Alice Henchley, a spokeswoman for *Nature*, says that the journal does not check the images in all papers because of limitations in resources, and that the STAP papers were not checked. But she adds that as one outcome of this episode, editors "have decided to increase the number of checks that we undertake on *Nature*'s papers. The exact number or proportion of papers that will be checked is still being decided."

In the face of extraordinary claims, checking images and text is hardly sufficient to test the veracity of results. Independent genetic analyses[12,13] proved that the world's first sheep cloned from adult cells, Dolly, was genetically identical to the mammary-gland cells from which she was cloned. And Snuppy, created by Hwang's lab, was similarly confirmed to be the first cloned dog[14,15]. Both verifications took place after publication of the initial results in order to quell debate, but such verification also could be done pre-publication, before a controversy arises.

"REPUTATION IN SCIENCE IS EVERYTHING. ONCE GONE, IT'S EXTREMELY HARD TO GET BACK."

Nature took that step in 2007, when it solicited an independent test of Mitalipov's monkey stem-cell lines, which showed them to be true clones. Those tests required coordination between researchers who had to navigate restrictions on sending cells across borders. It "dragged on for months", says Mitalipov. He subsequently chose another journal, *Cell*, to publish his next big stem-cell cloning paper[16] — demonstrating the process using human cells — partly for that reason. That paper was published 12 days after acceptance. An anonymous critic pointed out mistakes in the final manuscript including image duplication and mislabelled figures, but the authors proved that they were harmless errors (see *Nature* http://doi.org/mnk; 2013).

Some researchers say that it would be best if collaborators on a project carried out more stringent verification tests of their own before submitting a paper — something that Wakayama now woefully acknowledges. In the STAP case, however, it is not immediately obvious how this could have been done. Unlike for sheep, dogs or primates, genetically identical mouse strains and matched embryonic-stem (ES)-cell lines are widely available, making it easy to provide samples that look roughly correct. "There is no way to distinguish, genetically, ES cells from STAP cells originating from the same strain," says Rudolf Jaenisch, a stem-cell biologist at the Whitehead Institute for Biomedical Research in Cambridge, Massachusetts.

That said, post-hoc genetic analyses do seem to be unravelling the STAP riddle. In March, as problems with the papers mounted, Wakayama outsourced genetic sequencing of the purported STAP cells (see *Nature* http://doi.org/tf8; 2014). The work focused on a certain characteristic — the spot where a gene encoding a fluorescent protein inserts itself into the genome — to pinpoint the cells' origin. The results, which Wakayama announced in June, showed them to be different from the mice that had supposedly been used to make them (see *Nature* http://doi.org/tf9; 2014).

A step beyond genetic verification would be independent replication of the experiments. That would probably put undue burden on the scientists asked to do the work. But Richard Behringer, a developmental biologist at the MD Anderson Cancer Center in Houston, Texas, says that asking authors directly whether more than one person in the lab has reproduced the results would be one way for the journal "to ensure that all data and images in the manuscript were solid".

REPLICATION ISSUES

Nature does not disclose communications between editors and authors, but Campbell says that there were four independent groups on the two papers, and "it was our understanding that the work was independently replicated". When questions first arose about the STAP cells, moreover, co-authors on the papers were adamant that they had overseen replication.

Sasai says that his lab observed the generation of STAP cells. But in fact, he had asked Obokata to replicate only the first part of the STAP process — the expression of a protein, Oct4 — which he recorded with live imaging. At the time, he says, he did not consider "the possibility of a gap between these cells and the derivation of STAP stem cells".

Wakayama said that he "independently" produced STAP stem cells that looked exactly like ES cells — development beyond what Sasai witnessed — which convinced him that the process was solid. After problems first emerged, he told the *Nature* news team in an e-mail: "I succeeded at RIKEN independently, therefore, I know this result is absolutely true." Looking back now, though, he realizes that his replication was not completely independent — Obokata was at his side during the entire experiment "and oversaw every step in the process", he says. Because he moved to a new position at the University of Yamanashi soon after that test, he never characterized the cells and could not rule out the possibility that they had been switched or contaminated. The RIKEN investigative committee found that Sasai and Wakayama, although not involved in the misconduct, carried "heavy responsibility" for what happened.

The co-author who has most confused the issue of replication is Vacanti. Within a week of the STAP papers being published, he sent photos of what he claimed were human STAP

© 2014 Macmillan Publishers Limited. All rights reserved

YUYA SHINO/REUTERS/CORBIS

Nobel laureate and head of RIKEN, Ryoji Noyori, was moved to apologize for the mistakes made over the STAP research.

cells to the magazine *New Scientist*. As others failed to replicate the STAP experiment, he told the *Nature* news team in mid-February: "There really shouldn't be any difficulty. If I can do it, anyone should be able to do so." In mid-March, he published online a list of tips for making STAP cells "regardless of the cell type being studied". To date, however, he has produced no additional evidence that he has derived STAP cells in his laboratory. In a statement released on 2 July, Vacanti asserted that although he agreed to the retractions owing to errors in the manuscripts, he is confident that the "core concept" of STAP will be "verified by the RIKEN as well as independently by others".

For many stem-cell researchers, the most shocking part of the STAP controversy was the involvement of Niwa, Sasai and Wakayama in such troubled work. "Co-authors of a paper like that should have been certain that they can reproduce results independently and in this case they should share responsibility," says Davor Solter, a developmental and stem-cell biologist at the Institute of Medical Biology in Singapore. Wakayama takes the blame for not making more effort to check Obokata's work, such as looking at her notebooks, which the investigative committee found to be alarmingly disorganized.

Others sympathize with the researchers, who themselves were duped — whether through negligence or intention — by a junior colleague. "There has to be control, but also trust in science, otherwise the system breaks down completely," says Maria Leptin, a molecular biologist and director of EMBO. "I cannot watch over every step while they are pipetting. That's not the point."

But in addition to lax oversight, Janet Rossant, a stem-cell researcher at the Hospital for Sick Children in Toronto, Canada, and the outgoing head of the International Society for Stem Cell Research, points to "poor reviewing and editing by *Nature*, who were also too ready to publish without verification". Campbell disagrees. "*Nature* did not let down its guard," he says.

Some say that the journal should publish reviewers' comments to clarify the process. Campbell says that the publication of referees' comments has been considered, but that the disadvantages — which include potential misinterpretations and the desire of many referees to keep their comments confidential — have prevented the journal from embracing this.

"We have to accept that where there is research, there will be research misconduct," says Paul Taylor, a research-integrity adviser at the University of Melbourne, Australia. Efforts by institutions to train researchers and improve data-management infrastructure might help, "but no policy, no education or training, no administrative requirement, is going to stop misconduct".

Taylor adds that the focus should be on how an institution responds. In that sense, he says the STAP problem seems to have been a success. RIKEN has acknowledged flaws in its data management and exaggeration in its press release for the STAP papers. Taylor says that its response has been fast, effective and transparent. In the midst of the investigation into Obokata's work, Noyori instructed all RIKEN labs to review their published work — totalling tens of thousands of papers — for similar types of errors.

Among stem-cell researchers, STAP has become another cautionary tale to add to Hwang's, with its own set of lessons. For Loring, the story stresses the importance of good record-keeping and the need to enter collaborations with caution. "I lecture my lab members that being an author carries responsibility for the validity of all of the work in the paper. I really try to live by that." She says that she has removed her name from authorship lists in cases when she could not vouch for the quality of a manuscript.

But for many it is a lesson hard-learned, once again. "Reputation in science is everything," says Trounson, in a statement that applies no less to journals and institutions than individual scientists. "Once gone, it's extremely hard to get back." ■

David Cyranoski *is* Nature*'s Asia-Pacific correspondent, based in Shanghai, China.*

1. Obokata, H. *et al. Nature* **505,** 641–647 (2014).
2. Obokata, H. *et al. Nature* **505,** 676–680 (2014).
3. Hwang, W. S. *et al. Science* **303,** 1669–1674 (2004).
4. Hwang, W. S. *et al. Science* **308,** 1777–1783 (2005).
5. Kennedy, D. *Science* **311,** 335 (2006).
6. *Nature* **439,** 243 (2006).
7. Byrne, J. A. *et al. Nature* **450,** 497–502 (2007).
8. Cram, D. S., Song, B. & Trounson, A. O. *Nature* **450,** E12–E14 (2007).
9. Obokata, H. *et al. Nature* **511,** 112 (2014).
10. Vacanti, M. P., Roy, A., Cortiella, J., Bonassar, L. & Vacanti, C. A. *J. Cell. Biochem.* **80,** 455–460 (2001).
11. Guo, J. *et al. In Vitro Cell. Dev. Biol. Anim.* **41,** 278–283 (2005).
12. Ashworth, D. *et al. Nature* **394,** 329 (1998).
13. Signer, E. N. *et al. Nature* **394,** 329–330 (1998).
14. Parker, H. G., Kruglyak, L. & Ostrander, E. A. *Nature* **440,** E1–E2 (2006).
15. Seoul National University Investigation Committee, Lee, J. B. & Park, C. *Nature* **440,** E2–E3 (2006).
16. Tachibana, M. *et al. Cell* **153,** 1228–1238 (2013).

© 2014 Macmillan Publishers Limited. All rights reserved