REVIEW

Interactive learning: Lessons from two hybrids over two decades

Stanley Fields

Howard Hughes Medical Institute, Departments of Genome Sciences and Medicine, University of Washington, Seattle, WA, USA

The initial yeast two-hybrid experiment – published in 1989 – described an approach to detecting protein–protein interactions that has flourished over the last two decades, leading to the assembly of large-scale data sets of these interactions. Yet the yeast assay originated because of the laboratory's interests in technology development, not because of its need to identify partners of any protein then under study. In addition to such motivating forces, other features of the process of originating a technology can be revealed by considering the lessons of the two-hybrid approach. These include the value of timeliness in a method's development, the willingness of an investigator to try experimental approaches that prove fruitless, the ability of biological macromolecules to display surprising attributes, the benefits of a community expending efforts to expand the uses of a technology platform, and the role of scientific training of those who work in technology.

Received: April 13, 2009
Revised: June 10, 2009
Accepted: June 13, 2009

Keywords:
Protein arrays / Protein interaction / Technology development / Two-hybrid / Yeast

1 Introduction

While 2009 marks two decades since the publication of the *Nature* paper [1] describing the yeast two-hybrid assay, the point in time when the method came to mind was – perhaps not too surprisingly – 2 years earlier. In the intervening 22 years, the idea has grown from a hand-drawn sketch (Fig. 1) to a staple of biology laboratories, fastening together proteins derived from a veritable zoo of organisms. Its central principle – that the two functional domains of a transcriptional activator can be split apart and each fused to one of a pair of partner proteins in order to reconstitute the activator's ability to turn on a reporter gene – has spawned a brood of related technologies, as well as displays of cellular networks with thousands of interactions, and the algorithms to decipher the implications of these networks. The two-hybrid history has also been personally instructive, providing lessons about the nature and nurture of technology development that I have carried with me throughout my research career.

Correspondence: Dr. Stanley Fields, Howard Hughes Medical Institute, Departments of Genome Sciences and Medicine, University of Washington, Box 355065, Seattle, WA 98195, USA
E-mail: fields@u.washington.edu
Fax: +1-206-543-0754

2 The origins of technology

It is often said that new technologies arise in response to a specific biological problem, but the two-hybrid approach refutes the universality of this claim, reflecting a different reality of how methods often come to be. The yeast assay was not developed because of our need to find partners for any protein we were studying two decades ago, nor did it spring from our awareness that a simple method to do so did not then exist and had to be devised. In fact, my laboratory at the time could not have reasonably been said to even have a focus on protein interactions, although it did have an interest in technology.

The immediate motivation for the two-hybrid system was a request for grant applications that would lead to results with commercial potential, a request that would seemingly be met by a proposed assay system whose use in the biotech and pharmaceutical industries could be licensed and whose products – plasmids, strains, libraries and other reagents – could be sold by companies that supply biomedical research

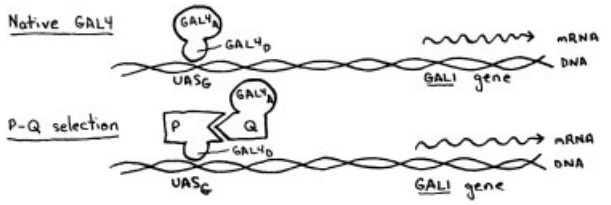

Figure 1. The first sketch of the two-hybrid assay, as provided in a grant application submitted at the end of 1987 to the Procter and Gamble Company. The native yeast Gal4 protein is shown as having DNA-binding ($GAL4_D$) and activation ($GAL4_A$) domains, with the DNA-binding domain recognizing a site on DNA known as the Upstream Activation Sequence for the *GAL* genes (UAS_G). Proteins P and Q form hybrids with the DNA-binding and activation domains, respectively, and reconstitute transcriptional activity, leading to expression of the *GAL1* gene.

laboratories. Despite these now obvious means to royalty income, the application outlining what became the two-hybrid assay was not funded, a consequence of the review panel not being persuaded of the technology's feasibility. Yet, in spite of the negative funding news, we immediately initiated experiments to develop the method, based solely on the beauty and elegance of this idea: if this assay worked, surely it would have tremendous utility to address problems being tackled by biologists. A few months later, we resubmitted the grant application to Procter and Gamble's University Exploratory Research Program, which funded the application and provided us with a small 3-year grant.

In fact, an argument can plausibly be made that it is the questions biologists pose that follow directly from the technologies available to them, rather than the converse. For instance, it makes little sense to purify vanishingly small quantities of a protein complex if mass spectrometric methods do not exist to identify the constituents; it is not worth postulating how the expression of thousands of genes changes in response to a cellular stress if DNA arrays had not come into fashion; and it is not logical to propose surveys of the entire human genome for variations linked to disease if an extensive set of polymorphisms and the instrumentation to rapidly detect DNA sequence polymorphisms are not available. Biologists who do not adjust their thinking – and even more, the aims of their research – in response to continuing technological innovations soon find themselves left in the dust of their more intrepid colleagues.

3 The value of timing

The two-hybrid method, which followed directly from the state of knowledge about transcription in the mid-1980s, demonstrates the critical nature of timeliness for a new technology. On the one hand, the two-hybrid assay could not have been developed much earlier because it required specific knowledge of the structure and modularity of eukaryotic transcription factors. On the other hand, the method fortunately predated the remarkable developments in genome sequencing and in MS instrumentation that enabled the rapid identification of proteins present in a complex by a determination of the sequence of short peptides from these proteins. Had these developments already been in play, the yeast method might have had far fewer takers.

Timeliness looms large in most technologies. In the early days of the yeast method, biochemists often said, perhaps ruefully, that yeast two-hybrid took a task – the purification of protein complexes – that had been the provenance of experts and replaced it with another task – a microbial genetic selection – that could be carried out by undergraduates new to biology. In the parallel universe of biochemistry, the advent of the TAP tag [2] 10 years ago provided an early means to conveniently purify a protein and its interacting partners. It was this tag, and the many other simple protein purification strategies it spawned, that ultimately brought the task of purifying protein complexes into the undergraduate laboratory. Such is always the nature of important technological discoveries: the unthinkable becomes, first, thinkable and then, not too much later, routine.

A lesson associated with timeliness is the interconnectedness of technological advances. The yeast two-hybrid assay arrived just as sequencing of whole genomes – of bacteria, then yeast, the nematode, the fruit fly, and the more recent numerous representatives of all the kingdoms of life – began to take off. The ramifications of the sequencing revolution for two-hybrid methodology were many: genome sequences enabled the small inserts in two-hybrid libraries to be assigned to proteins; they allowed array approaches in which full-length open reading frames fused to DNA-binding or activation domains could be made and analyzed [3, 4]; they galvanized scientists and funding agencies to consider functional studies on complete sets of proteins; and they brought in computational biologists to devise new algorithms, including those that analyzed protein interaction data.

4 The likelihood of failure

Another enduring lesson came as the realization set in of all the reasons why the yeast two-hybrid method might not or even – according to early skeptics – should not work. To name just a few of these reasons: most proteins need not have a modular structure that allows the construction of usable activation domain hybrid libraries; the affinity of most interactions might not be sufficient to trigger transcription in the two-hybrid configuration; most non-nuclear proteins might not be readily directed to the nucleus; and the eukaryotic transcription complex might have stereochemical constraints that require defined protein constructs

to position the DNA-binding and activation domains in precise locations. What we came to appreciate was that undue focus on any of these failure scenarios, most of which we considered only well after the assay had become established, might easily have persuaded us not to embark on the approach. Naiveté can sometimes be helpful.

More generally, any experiment – if thought about long enough – can be construed as highly unlikely to impossible to succeed. That biologists continue to do experimental science reflects our positive nature. For technology developers in particular, the willingness to try out numerous experimental approaches that prove, ultimately, to be doomed to failure is just part of the luggage that they carry. It is the occasional success along the way that makes these technologists willing to lug this baggage.

5 The surprising talents of proteins

One unexpected lesson of the two-hybrid and related methods pertains to the almost infinite malleability of proteins. As molecular biology began its rise to prominence with the advent of DNA cloning in the 1970s, its quintessential belief in the ability of bits of DNA to be disassembled in one place and reassembled somewhere else came to be orthodox. A similar conviction, however, regarding matchmaking among protein domains was heretical. In retrospect, the results of experiments performed in Mark Ptashne's laboratory [5, 6], published in the few years immediately prior to the two-hybrid assay, provided ample clues to the ability of protein domains – from LexA and Gal4 in this case – to be diced and spliced to each other to result in hybrid transcription factors that could function in yeast. Even earlier, fusions of proteins to β-galactosidase [7, 8] foreshadowed the widespread tagging of proteins with green fluorescent protein or other peptides or protein domains to follow their whereabouts within the cell or to act as purification hooks. For the two-hybrid assay, the extreme flexibility of eukaryotic transcription factor domains which enables them carry out their functions when fused to nearly any other protein domains allowed a rapid demonstration of the feasibility of this approach; similar efforts with a prokaryotic regulator in a bacterial assay system might not have been successful.

6 The power of community

The yeast two-hybrid method provides notable support for the proposition that a technology is proved worthwhile only when it is adopted, and often adapted, by a wider community. In many cases novel adaptations, for uses originally unforeseen in the initial application, add much value to the core technology. In this way, developments to extend the yeast assay mirror similar evolutions in technologies like DNA microarrays, protein purification and localization tags, PCR and now high-throughput DNA sequencing. For the two-hybrid method, these adaptations included new selection schemes, improved vectors and strains, ''reverse'' methods that select against interactions, approaches to assay DNA-protein, RNA–protein or small molecule–protein interactions, strategies that detect protein interactions in cellular compartments other than the nucleus, numerous other reporter proteins that can be split into two such as ubiquitin, green fluorescent protein, luciferase, and dihydrofolate reductase, and implementations of the two-hybrid approach that use host organisms beyond yeast [9–11]. Some of these related methods have gained considerable currency in their own right; others quickly vanished because they provided little additional benefit beyond technologies already in use, or suffered because they were unfamiliar and not easily implemented.

Along with widespread adoption came an increasing interest within the genomics community to scale up the assay to handle ever-larger complements of proteins. This scaling entailed engineering the assay to search proteins in parallel, by some combination of automation, pooling and array strategies [12]. Increases in scale of course increased the number of interactions reported in a single publication. But early efforts in this direction had another, more unfortunate outcome: high-throughput data could not be purged of the false positives that inevitably accompany any two-hybrid search [13]. The assay itself became tainted in the mind of some biologists: if the large scale studies led to results that could not be trusted, then maybe none of the assay's results should be accepted.

The solution to this lack of faith in the method's trustworthiness has, again, come from efforts of the community, both in experimental and computational directions. Numerous algorithms have been developed that compare two-hybrid data to other data types in an effort to classify interactions based on their *in vivo* likeliness, and many improvements to the two-hybrid protocol have been implemented in order to eliminate false positives (for example, Ref. [14]). In a similar vein, complicating issues of high-throughput MS analyses of protein complexes have also been dealt with by both experimental and computational approaches (for example, Ref. [15]). These analyses often rely on an initial immunoprecipitation for the purification of proteins, which can lead to the MS-based identification of contaminating proteins and publications that include false positives.

7 An appropriate training

The two-hybrid assay can also be viewed in the context of an individual scientific career. In 1976, about a decade before the two-hybrid assay originated, I arrived in Cambridge, England, for a 2-year fellowship that turned into 5 years and doctoral work at the MRC Laboratory of Molecular Biology. It was in Cambridge that I first realized the power of technology to propel biological science, largely from

seeing the impact of Fred Sanger [16]. As a student in George Brownlee's laboratory, and working closely with a postdoctoral fellow, Greg Winter – an earlier convert to technology development – I completed a thesis project focused on the sequence analysis of the RNA segments of influenza virus. But it could hardly be said that the experience made me into a virologist; rather, it succeeded in turning me into a technologist, using and developing methods of molecular biology in the early days of cloning and DNA sequencing.

I sought to remedy my lack of grounding in biology by carrying out postdoctoral training on the pheromone response pathway of the yeast *Saccharomyces cerevisiae* with Ira Herskowitz at the University of California, San Francisco. In the Herskowitz laboratory, I learned to think like a yeast geneticist, doing yeast crosses, conducting yeast screens and selections, and analyzing yeast mutants. Along the way, I also became aware of the current findings in yeast transcriptional regulation.

The two-hybrid assay – the progeny of a marriage of technology development and yeast genetics – amply reflects its dual heritage. Ideas in science typically emerge in the interstices between research areas, a compelling argument for the value of interdisciplinary training and the virtue of completing graduate and postdoctoral stints in widely differing environments.

8 A good name

The yeast two-hybrid assay benefits from many attributes of its name, which manages to convey in a short phrase both the host organism and its major moving parts. The name allows a convenient conversion to the readily identifiable acronym Y2H, and effortlessly scales down or up to one-hybrid and three-hybrid configurations. Above all, the name immediately captures the essential idea of the method and what is required in the way of plasmid constructions to effect it. But the assay nearly appeared in our 1991 publication [17] without its memorable designation.

That the yeast method needed a good name was impressed upon me by my postdoctoral advisor, Ira Herskowitz, who himself popularized concepts such as the "cassette model" of yeast mating type interconversion and the "α1-α2 model" of yeast cell type determination. Shortly after we demonstrated that the assay worked on a known pair of interacting proteins and were in the midst of developing it into a library screen, I saw Ira in San Francisco and described our efforts. He asked me what we called the method, and I vaguely recall mumbling something along the lines of "Reconstitution of Gal4 transcriptional activity by the interaction of two proteins fused to different domains of…." Ira wisely shook his head, signifying that such an unwieldy name would not serve us well in popularizing the approach. Upon my return to the lab, we sought a short descriptor and came up with "a two hybrid system," helpfully changed by a copyeditor at the *Proceedings of the National Academy of Sciences USA* to "the two-hybrid system."

9 Final thoughts

You may have enough foresight to be able to predict what new technologies will sweep through biological research, but I – 20 years on from the publication of the two-hybrid method and mostly engaged in technology development during this period – have yet to develop any such prognostic ability. Yet I have acquired the certain knowledge that progress in biology will follow directly from our collective success in coming up with new technological innovations. The advice to university departments, tenure committees, funding agencies and the like is clear: encourage biologists – especially young ones – to tinker in the lab, trying to devise approaches that will change how we go about our business.

I thank Mark Johnston and Eric Phizicky for comments on the manuscript. I apologize that in writing this personal view of the two-hybrid assay, I have not cited the many primary publications that contributed to this work and the many applications that are related to or followed on from it. Research in my laboratory has been supported by NIH grant P41 RR11823 from the National Center for Research Resources. I am an investigator of the Howard Hughes Medical Institute.

The author has declared no conflict of interest.

10 References

[1] Fields, S., Song, O., A novel genetic system to detect protein-protein interactions. *Nature* 1989, *340*, 245–246.

[2] Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* 1999, *17*, 1030–1032.

[3] Uetz, P., Giot, L., Cagney, G., Mansfield, T. A. *et al.*, A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* 2000, *403*, 623–627.

[4] Ito, T., Chiba, T., Ozawa, R., Yoshida, M. *et al.*, A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. USA* 2001, *98*, 4569–4574.

[5] Brent, R., Ptashne, M., A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* 1985, *43*, 729–736.

[6] Ma, J., Ptashne, M., Converting a eukaryotic transcriptional inhibitor into an activator. *Cell* 1988, *55*, 443–446.

[7] Silhavy, T. J., Casadaban, M. J., Shuman, H. A., Beckwith, J. R., Conversion of beta-galactosidase to a membrane-bound state by gene fusion. *Proc. Natl. Acad. Sci. USA* 1976, *73*, 3423–3427.

[8] Silhavy, T. J., Shuman, H. A., Beckwith, J., Schwartz, M., Use of gene fusions to study outer membrane protein

localization in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 1977, *74*, 5411–5415.

[9] Stagljar, I., Korostensky, C., Johnsson, N., te Heesen, S., A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins *in vivo*. *Proc. Natl. Acad. Sci. USA* 1998, *95*, 5187–5192.

[10] Remy, I., Michnick, S. W., Clonal selection and *in vivo* quantitation of protein interactions with protein-fragment complementation assays. *Proc. Natl. Acad. Sci. USA* 1999, *96*, 5394–5399.

[11] Vidal, M., Legrain, P., Yeast forward and reverse 'n'-hybrid systems. *Nucleic Acids Res.* 1999, *27*, 919–929.

[12] Cusick, M. E., Klitgord, N., Vidal, M., Hill, D. E., Interactome: gateway into systems biology. *Hum. Mol. Genet.* 2005, *14*, R171–R181.

[13] Fields, S., High-throughput two-hybrid analysis. The promise and the peril. *FEBS J.* 2005, *272*, 5391–5399.

[14] Yu, H., Braun, P., Yildirim, M. A., Lemmens, I. *et al.*, High-quality binary protein interaction map of the yeast interactome network. *Science* 2008, *322*, 104–110.

[15] Gingras, A. C., Gstaiger, M., Raught, B., Aebersold, R., Analysis of protein complexes using mass spectrometry. *Nat. Rev. Mol. Cell Biol.* 2007, *8*, 645–654.

[16] Sanger, F., Sequences, sequences, and sequences. *Annu. Rev. Biochem.* 1988, *57*, 1–28.

[17] Chien, C. T., Bartel, P. L., Sternglanz, R., Fields, S., The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl. Acad. Sci. USA* 1991, *88*, 9578–9582.

