CellPress

Scientific rigor through videogames

Adrien Treuille[1] and Rhiju Das[2,3]

[1] Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15206, USA
[2] Department of Biochemistry, Stanford University, Stanford, CA 94305, USA
[3] Department of Physics, Stanford University, Stanford, CA 94305, USA

Hypothesis-driven experimentation – the scientific method – can be subverted by fraud, irreproducibility, and lack of rigorous predictive tests. A robust solution to these problems may be the 'massive open laboratory' model, recently embodied in the internet-scale videogame EteRNA. Deploying similar platforms throughout biology could enforce the scientific method more broadly.

Introduction

A growing spate of controversies, retractions, and fraud cases highlight the susceptibility of modern biology to untruths. Despite an elaborate peer review system, issues such as data manipulation, lack of reproducibility, lack of predictive tests, and cherry-picking among numerous unreported data occur frequently and, in some fields, may be pervasive (e.g., [1–3]). Some of these issues are further intensified by what is now a hallmark of modern biology, the use of high-throughput experimental techniques. A single nucleic acid sequencing run can generate billions of data points. Following up such experiments with cycles of hypothesis generation and testing is critical to establish scientific truth but can be expensive and time-consuming. It is particularly tempting to skip the extra work if a massive initial data set can be cherry-picked into a publishable manuscript without the additional effort.

In comparing the rapid growth of experimental throughput to the seemingly fixed rate of 'conventional' hypothesis generation, some have worried that our definition of scientific understanding may have to change [10], while others argue that experimental design must be automated [4]. In our view, however, hypothesis generation and experimental tests by creative scientists are what make science exciting and truthful. Therefore, 3 years ago we began to explore a new idea: rather than allow growing experimental data to overwhelm hypothesis generation, perhaps a larger number of people could be recruited to deeply analyze and design experiments. The result is what we term a massive open laboratory model for science (Figure 1).

Corresponding authors: Treuille, A. (treuille@cs.cmu.edu); Das, R. (rhiju@stanford.edu).
Keywords: RNA design; crowdsourcing; reproducibility; transparency; videogame; high-throughput experiments.

0968-0004/
 http://dx.doi.org/10.1016/j.tibs.2014.08.005

The first massive open laboratory

The first implementation of a massive open laboratory, EteRNA, and its application to an unsolved RNA engineering problem was published earlier this year [5]. The problem we chose involves noncoding RNAs, whose critical roles in cellular and viral machinery are inspiring many groups to pursue novel therapies that disable or manipulate the molecules involved. When we started our efforts, existing algorithms to design RNAs that fold into specific structures had been largely untested in the laboratory [6]. Accelerations in a technique called single-nucleotide resolution structure mapping [9] enabled us to test numerous designs and we found that they typically did not fold correctly *in vitro*. We were clearly missing some of the rules for robust RNA engineering, but the way forward was not obvious to us or other experts. RNA design therefore offered an important problem for which existing *in silico* models were available but incomplete and for which experimental feedback could be obtained rapidly, all features necessary for launching a massive open laboratory. Further motivated by the successes of prior scientific gaming projects [7,8], we developed a videogame interface for creating and viewing RNA design experiments and sought to recruit 'citizen scientists' to join us in these puzzles.

Since its release in 2011, EteRNA has registered over 150 000 participants who have contributed over 2 million human-hours of puzzle solving and generated 13 000 designs that have been experimentally synthesized as part of the game. In their first experimental tests, the community fared no better than existing algorithms at RNA design. Within months, however, experimental feedback on previously unseen puzzles enabled the EteRNA community to substantially outperform these algorithms in challenges of growing complexity. Furthermore, the design rules proposed by EteRNA's participants – most of which were unknown in prior expert literature – enabled the development of an automated algorithm, EternaBot, that also outperformed prior algorithms in designing numerous previously unseen RNA folds [5].

Scientific rigor through a massive open laboratory

EteRNA presents a scientific model leading to empirically grounded publications, similar to conventional science. Unlike conventional science, however, EteRNA's massive open laboratory embodies several unconventional solutions to difficult issues in modern scientific practice, including data manipulation, irreproducibility, lack of predictive tests, and cherry-picking.

Trends in Biochemical Sciences November 2014, Vol. 39, No. 11

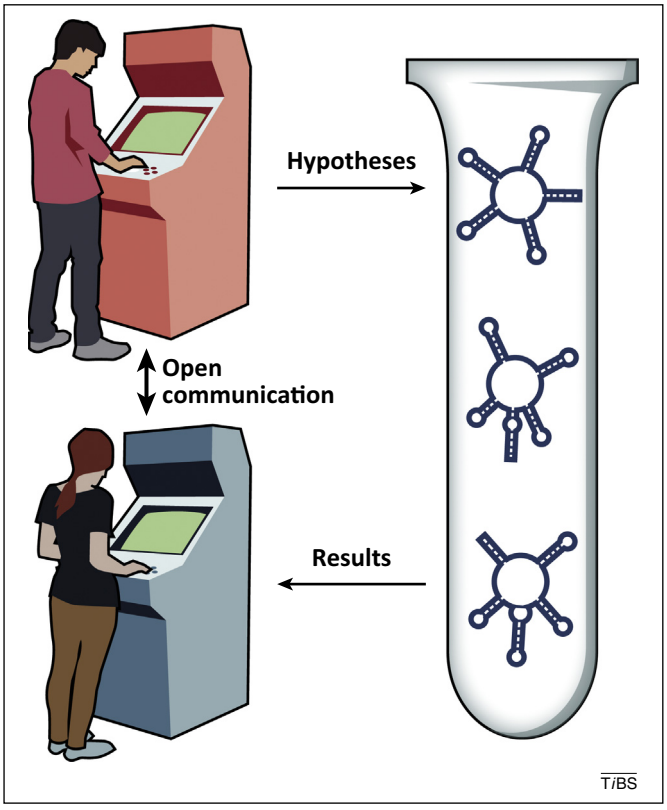

Figure 1. The massive open laboratory model for science. EteRNA participants (left) are incentivized through a videogame-like interface to make specific hypotheses about RNA engineering and to design experiments that rigorously test them. Actual RNA synthesis and *in vitro* structure mapping occurs in a remote wet laboratory (right), which returns experimental feedback and game rewards to the participants, and the cycle continues. Data integrity, reproducibility, and truly predictive tests are intrinsic features of a massive open laboratory.

First, the EteRNA massive open laboratory model prevents many forms of data manipulation. In typical scientific practice, the same team both generates hypotheses and conducts experiments. Although not usually noted, this generates a conflict of interest. EteRNA instead separates these activities: participants generate hypotheses while our laboratory remotely conducts experiments, certifies their accuracy, estimates experimental errors, and releases all raw data to be publically available and searchable. This openness and the separation of hypothesis creator and experimentalist preclude many common forms of data manipulation.

Second, the platform allows rapid tests of reproducibility. In conventional scientific research the burden of additional experiments and analysis can de-incentivize additional cycles of predictive tests before publication. In EteRNA, however, experiments are defined through a straightforward interface online, occur monthly, and generate point rewards for the participant. This approach encourages participants to conduct more tests or to challenge each another's hypotheses by creating novel experimental tests. Our remote experimental pipeline ensures that these challenges occur under uniform experimental conditions. Compared with conventional science, a massive open laboratory thus strongly encourages reproducibility and error checking.

Third, EteRNA requires rigorous adherence to the scientific method: a nontrivial prediction or hypothesis must precede each experiment. In conventional research, scientists who develop models that rationalize prior biological data are not always obliged to conduct further rigorous experimental tests. Calculations that are presented as matching data are not necessarily predictions but instead can be 'postdictions', whose accuracy may not reflect predictive power and whose presentation may reflect the scientist's bias. The actual order of events – prediction versus postdiction – cannot typically be assessed by others. In EteRNA, however, a participant's hypothesis is time-stamped in the project in the form of a specific RNA sequence or design rule and no experiment is conducted unless a hypothesis precedes it. It is important to note that creative models that rationalize data *post hoc* are still allowed and even incentivized through rewards for proposing design rules, but such models are tested in subsequent experiments. In addition, cherry-picking among unreported data or hypotheses would be caught easily, since both data and hypotheses are open to all and fully searchable. In these ways, time-stamping and openness ensure scientific rigor.

Prospects for more massive open laboratories

The massive open laboratory approach therefore provides a model of scientific practice that has unusual rigor and is robust to data manipulation, irreproducibility, and cherry-picking in ways that most modern scientific practices are not. We view the main current challenge of the massive open laboratory approach as a practical one of infrastructure development. On the one hand, the experimental throughput of EteRNA has grown from tens to thousands of sequences per experimental round and a wide range of RNA molecules and hypotheses can now be explored *in vitro*. On the other hand, testing these hypotheses' biological relevance requires separate experiments in viruses and living cells, systems for which massive open laboratories do not yet exist. Nevertheless, it is possible to envision compelling game interfaces and massively parallel platforms for genetic engineering of viruses and cells, live-cell fluorescence microscopy, microfluidic interrogation of cells with small molecules, profiling macromolecule content through sequencing and mass spectrometry, and high-resolution electron microscopy of purified macromolecular complexes. In other words, massive open laboratories should be possible for some or much of standard biological inquiry and deploying such platforms may enable the rigorous and creative practice of the scientific method to keep pace with experimental throughput.

What are the remaining challenges? Based on EteRNA's experiences, the engineering cost required to develop any new massive online laboratory, while significant, should be within the budget of a conventional grant in the life sciences. The major current barrier may instead be the career risks that these projects pose for their creators. In particular, videogames, which appear critical for recruiting scientifically engaged citizens, are generally viewed as incompatible with 'serious' or rigorous research. Nevertheless, massive open laboratories integrate videogames with transparent data access, experimental certification, and hypothesis time-stamping and thereby may surpass conventional

science in rigor. We hope that other life scientists and computer scientists will recognize the need for projects that enable large-scale research with scientific rigor and that a few will join us in taking the risks to create new massive open laboratories.

Acknowledgments

The authors are grateful to W. Kladwang and J. Lee for their close collaboration over several years, J. Salzman and O. Brandman for discussions, and the Burroughs–Wellcome Foundation and W.M. Keck Medical Research Foundation for supporting expansions of the EteRNA massive open laboratory and for open-access charges.

References

1 Ioannidis, J.P. (2005) Why most published research findings are false. *PLoS Med.* 2, e124

2 Begley, C.G. and Ellis, L.M. (2012) Drug development: raise standards for preclinical cancer research. *Nature* 483, 531–533

3 Sutherland, W.J. *et al.* (2013) Policy: twenty tips for interpreting scientific claims. *Nature* 503, 335–337

4 Naik, A.W. *et al.* (2013) Efficient modeling and active learning discovery of biological responses. *PLoS ONE* 8, e83996

5 Lee, J. *et al.* (2014) RNA design rules from a massive open laboratory. *Proc. Natl. Acad. Sci. U.S.A.* 111, 2122–2127

6 Bida, J.P. and Das, R. (2012) Squaring theory with practice in RNA design. *Curr. Opin. Struct. Biol.* 22, 457–466

7 Lintott, C.J. *et al.* (2008) Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. *Mon. Not. R. Astron. Soc.* 389, 1179–1189

8 Khatib, F. *et al.* (2011) Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nat. Struct. Mol. Biol.* 18, 1175–1177

9 Seetin, M.G. *et al.* (2014) Massively parallel RNA chemical mapping with a reduced bias MAP-seq protocol. *Methods Mol. Biol.* 1086, 95–117

10 Weinberg, R. (2010) Point: hypothesis first. *Nature* 464, 678