



Future Science
James A. Evans
Science **342**, 44 (2013);
DOI: 10.1126/science.1245218

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of October 4, 2013):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/342/6154/44.full.html>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/content/342/6154/44.full.html#related>

This article **cites 11 articles**, 6 of which can be accessed free:

<http://www.sciencemag.org/content/342/6154/44.full.html#ref-list-1>

Future Science

James A. Evans

The massive growth of global research activity in recent years has spurred studies exploring how productive this expansion has been and what the future may hold. Although the creativity and serendipity of individual discoveries will remain difficult to model, quantitative research has revealed regularities in the rates of discovery and the outcome of published findings over time (1). Some of these studies demonstrate that innovation has been decreasing, which may reflect the output of a scientific enterprise whose system of input has supported the pursuit of research focused on “low-hanging fruit” (2). Coincident with this decline has been the recent global recession, a circumstance that called for accountability of economic and social returns from public investments in research (3). There is now a great demand for insight into how the system of science works. On page 127 of this issue, Wang *et al.* (4) offer one approach to assess, and perhaps even augment, scientific productivity.

An emerging area of interest in research on the “science of science” is the prediction of future impact. Impact prediction is consequential for the evaluation of research grants, the dispensing of scholarly awards, and the determination of faculty salaries, among other decisions. As predictions improve, they will play a larger role in directing choices about what areas public and private capital will choose to research, develop, and produce. But how can we predict the future?

A number of recent studies have honed our understanding of the factors that influence future citations to an article or researcher. For example, citations accrue to articles published in scientific journals over time according to a well-behaved log-normal distribution, with a rise in citations at the point of publication followed by a gradual decay (5). There is also a strong “first-mover” advantage to the receipt of citations—that is, an early mediocre article on a topic will often receive more citations than a later excellent one (6). There are still other factors that predict citations to a scientist’s complete oeuvre, such as the subject diversity of journals in which the scientist’s early articles were published (7). Wang *et al.* advances these efforts by integrating

Department of Sociology, University of Chicago, Chicago, IL 60637, USA. E-mail: jevans@uchicago.edu



Winners foreseen. Methods to predict future citations of an article could improve the overall productivity of the scientific enterprise.

three key features into a generative model of an article’s long-term impact: preferential attachment, a temporal citation trend, and the underlying “fitness” of the paper. Their model links previous models of what accounts for the popularity of Web sites [the preferential attachment with which other Web sites hyperlink to it (8) and the “quality” of the Web sites’s content, reflected in its utility for Internet users (9)] and the log-normal probability with which an article’s influence rises then falls over time. The authors combined these elements into a three-parameter model with a number of clear, derivable, and empirically estimable quantities associated with an individual article. These include the article’s relative fitness (its importance relative to its peers), immediacy (the time required to reach its citation peak), longevity (its rate of citation decay), impact time (the characteristic time to attract the bulk of its citations), and ultimate impact (the citations it will acquire over its lifetime, which depends only on its relative fitness). The resulting model was used to predict later citations on the basis of early patterns. By accounting for these different features, all nonzero citation trajectories followed a similar path. Moreover, the authors generalize the model to the level of journals (as they could for researchers, departments, universities, countries, or the scientific system as a whole) to explain how a recent drop in the relative impact of a well-known journal can be solely accounted for by the rising impact time associated with its articles.

We should consume these explanations and the concepts underlying them with cau-

Can predicting an article’s success change science?

tion (as we should with any simplifying model). For example, “fitness,” which is well-named, is not merely quality—the “inherent” or timeless value of an article to science—but the fit between an article and its perceived importance to an audience at the time of publication. Fitness is neither “inherent” nor timeless. With relevant developments in other areas of science, an article’s ideas may become fit. The model of Wang *et al.* is also not magical; its estimation of the

future improves with more past data. Moreover, citations represent only a single dimension of impact, because nonacademic practitioners may be influenced by an article but cannot cite it. Nonetheless, the model assembles a number of important factors for predicting future citation success.

The success of this model and others that will build on it raises the question of how predicting an article’s success could change science. Widespread, consistent prediction of an idea’s future impact will necessarily speed the resolution of which ideas win and which lose in the competition for attention and resources. Because we cannot know everything about the future, knowing only the momentum of an article’s reception could act as a self-fulfilling prophecy—thus, scientists and funders could prematurely abort ideas that may yet have a second or third act to play. The classic work by mathematicians Erdős and Rényi on random networks, which much later became the foundation for popular approaches in network science, is such an example; Wang *et al.* admit that their model could not have predicted future citations of this seminal work. The ability to better predict an article’s success could translate into a faster scientific life cycle for the discovery—from time of publication to widespread acceptance. This might then translate into faster convergence to best practices that would boost the number of scientists with skills required to build on an impactful discovery. This could happen, however, only because rapid swarming around new ideas will increase competition for follow-on research and publication, as

CREDIT: C. BICKEL/SCIENCE

a larger proportion of scientists pursue the same anointed path.

These cautions, however, do not apply to all that more successful impact prediction portends. The ability to automatically extract scientific claims from research articles and reason across them should lead to the prediction or computational generation of promising new hypotheses. It likely will also expose common assumptions and practices of science to scrutiny and explicit evaluation (10). In this way, citation prediction represents one step on the path to creating algorithmic or

robot “scientists” (11) that are more creative, risky, persistent, and wide-reading than ourselves (12). By enabling scientists to consider not only the most fruitful hypothesis but also the most fruitful algorithm for generating hypotheses, future prediction methods would augment scientific ability, increase productivity, and multiply returns from science for society.

References

1. S. Arbesman, *The Half-Life of Facts: Why Everything We Know Has an Expiration Date* (Current, New York, 2012).
2. B. F. Jones, *Rev. Econ. Stud.* **76**, 283 (2009).

3. J. Lane, *Science* **324**, 1273 (2009).
4. D. Wang, C. Song, A.-L. Barabási, *Science* **342**, 127 (2013).
5. M. J. Stringer, M. Sales-Pardo, L. A. Nunes Amaral, *PLoS ONE* **3**, e1683 (2008).
6. M. E. J. Newman, *Europhys. Lett.* **86**, 68001 (2009).
7. D. E. Acuna, S. Allesina, K. P. Körding, *Nature* **489**, 201 (2012).
8. A.-L. Barabási, R. Albert, *Science* **286**, 509 (1999).
9. G. Bianconi, A.-L. Barabási, *Europhys. Lett.* **54**, 436 (2001).
10. J. A. Evans, J. G. Foster, *Science* **331**, 721 (2011).
11. R. D. King *et al.*, *Science* **324**, 85 (2009).
12. J. A. Evans, A. Rzhetsky, *Science* **329**, 399 (2010).

10.1126/science.1245218

PLANT SCIENCE

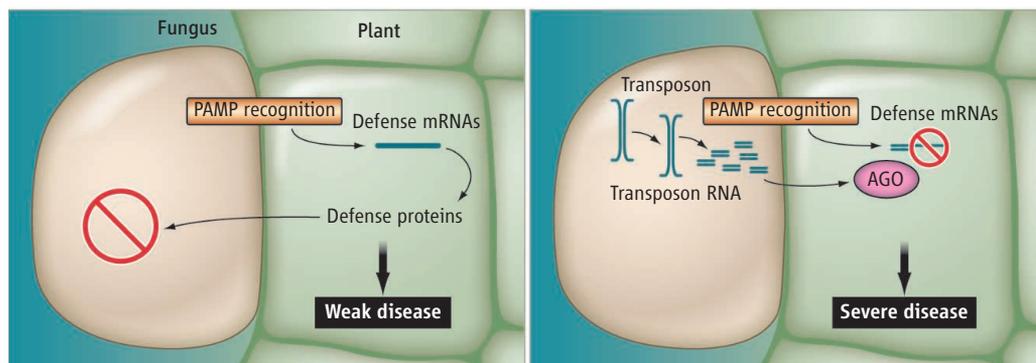
Small RNA—the Secret of Noble Rot

David Baulcombe

Benjamin Franklin should have included infectious disease as a third certainty of life alongside death and taxes. We will always be plagued by pests and pathogens because they generally have shorter life cycles than their hosts, allowing them to evolve rapidly to avoid host defenses. The use of antibiotics, for example, may select for drug-resistant strains of bacteria and, in agriculture, the widespread cultivation of a disease-resistant crop is often followed by the appearance of resistance-breaking strains of the pathogen.

Whenever a host acquires a novel defense, the pest will, eventually, evolve a corresponding counter-defense system. On page 118 of this issue, Weiberg *et al.* (1) describe a strategy in which a fungal parasite uses RNA to block a host's defense system

In plants, the evolutionary arms race involves a set of host receptors that recognize components on the surface of the pathogen (2). These pathogen-associated molecular patterns (PAMPs) include, for example, flagellar proteins of Gram-negative bacteria and chitin in the cell wall of fungi. Triggering these receptors leads to intracellular signaling within and biochemical changes in and around the infected cell such that the invad-



Fungal sRNAs affect disease progression. In the absence of fungal sRNA (left) the pathogen-associated molecular pattern (PAMP) recognition system ensures that the fungus is suppressed and disease is weak. By contrast, in their presence (right), the host defense system is blocked, and the disease progresses rapidly.

ing pathogen is suppressed. Counter-defense then involves pathogen-derived effectors that either block the initial recognition on the surface of the host cell, or are transferred into the infected cell where they suppress the signaling pathways or biochemical changes. However, the arms race does not end there. The plant has a second set of receptors that recognize the effectors and trigger additional layers of defense. The pathogen, correspondingly, has additional counter-defense systems (2).

Most of the characterized pathogen effectors are proteins, but Weiberg *et al.* (1) demonstrate that we can add RNA to the list of effectors with trans-kingdom activity. The new research involves *Botrytis cinerea*—a necrotrophic fungal pathogen that infects many plant species, including tomato and strawberry, on which it causes gray mold.

Small RNA from a fungal pathogen is transferred to cells of a plant host where it silences the mRNA for defense proteins.

Most notably, it is the noble rot that is so important for the production of exquisite dessert wines.

The new findings involve a class of small RNAs (sRNAs) that includes microRNAs and small interfering RNAs. These sRNAs are typically 20 to 24 nucleotides in length and they guide Argonaute (AGO) nucleases by Watson-Crick base pairing to coding or noncoding RNAs in either the nucleus or cytoplasm so that the targeted RNAs accumulate at a lower level than they would in the absence of the sRNA (3).

There are more than 800 fungal sRNAs that are induced when *Botrytis* is infecting plant cells and, of these, 73 were predicted from their sequence to silence host mRNAs, including some that have a role in immunity (1). This silencing potential could be due to