site IIIa is associated with Na+ release. High-resolution structures that accurately reveal Na+ coordination and associated hydrogen-bonding networks will be essential for a better understanding of the structure-function relations of ion exchange, transport, and specificity and how the mechanism is affected byregulation and disease-related mutations.

References and Notes

Acknowledgments: B. Vilsen and J. Petersen, Department of Biomedicine, Aarhus University, Denmark, are thanked for preparing enzyme for crystallization. We thank C. Schulze-Briese, T. Tomizaki, and V. Olieric (Swiss Light Source) for assistance with synchrotron data collection; B. Bjerring Jensen, A. M. Nielsen, and J. L. Karlsen for technical assistance; and J. P. Mort, L. Yatime, M. Laursen, H. Khandelka, and M. J. Clausen for valuable discussions. Support was provided by the Danscatt program of the Danish Natural Science Research Council. M.N. was supported by the Swedish Research Council, L.R. by the Danish Council for Independent Research in Medical Sciences, E.L. by a European Research Council starting grant (contract 209825), and P.N. by a European Research Council advanced grant (contract 250322). H.P. was supported by the Lundbeck Foundation, the Carlsberg Foundation, and I’Oréal/United Nations Educational, Scientific, and Cultural Organization. The authors made the following contributions: H.P. and P.N. performed the study design. M.N. crystallized the protein, collected and processed x-ray data, and determined and refined the structure, assisted by L.R. and P.G. The structural analysis was carried out by M.N., L.R., and P.G., assisted by P.N., whereas M.A. and E.L. performed the MD simulations, assisted by P.G. H.P. designed and performed the electrophysiological studies. N.F. designed and performed the deocclusion experiments. M.N., L.R., P.G., H.P., and P.N. wrote the paper. All authors discussed the results and commented on the manuscript. Coordinates and structure factors have been deposited in the Protein Data Bank (PDB) with accession no. 4hjq.

Supplementary Materials
www.sciencemag.org/content/342/6154/123/suppl/DC1
Materials and Methods
Figs. S1 to S12
Table S1
References (23–58)
17 July 2013; accepted 3 September 2013
10.1126/science.1243352

Quantifying Long-Term Scientific Impact
Dashun Wang,1,2* Chaoming Song,1,3* Albert-László Barabási1,4,5,6†
The lack of predictability of citation-based measures frequently used to gauge impact, from impact factors to short-term citations, raises a fundamental question: Is there long-term predictability in citation patterns? Here, we derive a mechanistic model for the citation dynamics of individual papers, allowing us to collapse the citation histories of papers from different journals and disciplines into a single curve, indicating that all papers tend to follow the same universal temporal pattern. The observed patterns not only help us uncover basic mechanisms that govern scientific impact but also offer reliable measures of influence that may have potential policy implications.

The many tangible measures of scientific impact, one stands out in its frequency of use: citations (1–10). The reliance on citation-based measures, from the Hirsch index (h) to the g-index (i11), from impact factors (IF) to eigenfactors (i12), and on diverse ranking-based

1Center for Complex Network Research, Department of Physics, Department of Biology, and Department of Computer Science, Northeastern University, Boston, MA 02115, USA. 2IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598, USA. 3Department of Physics, University of Miami, Coral Gables, FL 33124, USA. 4Center for Cancer Systems Biology, Dana Farber Cancer Institute, Boston, MA 02115, USA. 5Department of Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA 02115, USA. 6Center for Network Science, Central European University, Budapest, Hungary.

*These authors contributed equally to the work. †Corresponding author. E-mail: alb@neu.edu

power: A group of papers that within a 5-year span collect the same number of citations are found to have widely different long-term impacts (fig. S2B).
3) Paradigm-changing discoveries have notoriously limited early impact (3) precisely because the more a discovery deviates from the current paradigm, the longer it takes to be appreciated by the community (16). Indeed, although for most papers their early- and long-term citations correlate, this correlation breaks down for discoveries with the most long-term citations (Fig. 1B). Hence, publications with exceptional long-term impact appear to be the hardest to recognize on the basis of their early citation patterns.
4) Comparison of different papers is confounded by incompatible publication, citation, and/or acknowledgment traditions of different disciplines and journals.

Long-term cumulative measures like the Hirsch index have predictable components that can be extracted via data mining (4, 17). Yet, given the myriad of factors involved in the recognition of a new discovery, from the work’s intrinsic value to timing, chance, and the publishing venue, finding regularities in the citation history of individual papers, the minimal carriers of a scientific discovery, remains an elusive task.
In the past, much attention has focused on citation distributions, with debates on whether they follow a power law (2, 18, 19) or a log-normal form (3, 7, 15). Also, universality across disciplines allowed the rescaling of the distributions
by discipline-dependent variables (7, 15). Together, these results offer convincing evidence that the aggregated citation patterns are characterized by generic scaling laws. Yet little is known about the mechanisms governing the temporal evolution of individual papers. The inherent difficulty in addressing this problem is well illustrated by the citation history of papers extracted from the Physical Review (PR) corpus (Fig. 1A), consisting of 463,348 papers published between 1893 and 2010 and spanning all areas of physics (3). The fat-tailed nature of the citation distribution 30 years after publication indicates that, although most papers are hardly cited, a few do have exceptional impact (Fig. 1B, inset) (2, 3, 7, 19, 20). This impact heterogeneity, coupled with widely different citation histories (Fig. 1A), suggests a lack of order and hence lack of predictability in citation patterns. As we show next, this lack of order in citation histories is only apparent, because citations follow widely reproducible dynamical patterns that span research fields.

We start by identifying three fundamental mechanisms that drive the citation history of individual papers:

1. **Preferential attachment** captures the well-documented fact that highly cited papers are more visible and are more likely to be cited again than less-cited contributions (20, 21). Accordingly a paper’s probability to be cited again is proportional to the total number of citations, the paper received previously (fig. S3).

2. **Aging** captures the fact that new ideas are integrated in subsequent work; hence, each paper’s novelty fades eventually (22, 23). The resulting long-term decay is best described by a log-normal survival probability (Fig. 1C and supplementary materials S2.1), where $t$ is time; $\mu$ indicates immediacy, governing the time for a paper to reach its citation peak; and $\sigma$ is longevity, capturing the decay rate.

3. **Fitness**, $\eta$, captures the inherent differences between papers, accounting for the perceived novelty and importance of a discovery (24, 25). Novelty and importance depend on so many intangible and subjective dimensions that it is impossible to objectively quantify them all. Here, we bypass the need to evaluate a paper’s intrinsic value and view fitness $\eta$ as a collective measure capturing the community’s response to a work.

Combining these three factors, we can write the probability that paper $i$ is cited at time $t$ after publication as

$$\Pi_i(t) \sim \eta_i c_i^{s} P(t)$$  

Solving the associated master equation, Eq. 2 allows us to predict the cumulative number of citations.
citations acquired by paper \( i \) at time \( t \) after publication (supplementary materials S2.2)

\[
c_i^t = m \left[ \frac{\lambda \Phi \left( \frac{i - t}{\sqrt{2}} \right)}{\sqrt{2\pi}} - 1 \right] \approx m \left[ \frac{\lambda \phi \left( \frac{i - t}{\sqrt{2}} \right)}{\sqrt{2\pi}} - 1 \right]
\]

where

\[
\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-y^2/2} dy
\]

is the cumulative normal distribution, \( m \) measures the average number of references each new paper contains, \( \beta \) captures the growth rate of the total number of publications (supplementary materials S1.3), and \( A \) is a normalization constant (supplementary materials S2.2). Hence \( m, \beta, \) and \( A \) are global parameters, having the same value for all publications. We have chosen \( m = 30 \) throughout the paper, because our results do not depend on this choice (supplementary materials S2.3).

Equation 3 represents a minimal citation model that captures all known quantifiable mechanisms that affect citation histories. It predicts that the citation history of paper \( i \) is characterized by three fundamental parameters: the relative fitness, \( \lambda_i = \eta_i/\lambda \), capturing a paper’s importance relative to other papers; \( \mu_i \); and \( \sigma_i \). By using the rescaled variables \( \tilde{t} = (\ln t - \mu_i)/\sigma_i \) and \( \tilde{c} = \ln(1 + c_i/m)/\lambda_i \), we obtain our main result

\[
\tilde{c} = \Phi(\tilde{t})
\]

predicting that each paper’s citation history should follow the same universal curve \( \Phi(\tilde{t}) \) if rescaled with the paper-specific \( (\lambda, \mu, \text{and} \sigma) \) parameters. Therefore, given a paper’s citation history, that is, \( t \) and \( c_i^t \), we can obtain the best-fitted three parameters for paper \( i \) by using Eq. 3. To illustrate the process, we selected a paper from our corpus, whose citation history is shown in Fig. 1, D and E. We fitted to Eq. 3 the paper’s cumulative citations (Fig. 1E) by using the least square fit method, obtaining \( \lambda = 2.87, \mu = 7.38, \text{and} \sigma = 1.2 \). To illustrate the validity of the fit, we show (Fig. 1E) the prediction of Eq. 3 using the uncovered fit parameters.

To test the model’s validity, we rescaled all papers published between 1950 and 1980 in the \( PR \) corpus, finding that they all collapse into Eq. 5 (Fig. 1E, see also supplementary materials S2.4.1 for the statistical test of the data collapse). The reason is explained in Fig. 1G: By varying \( \lambda, \mu, \text{and} \sigma \), Eq. 3 can account for a wide range of empirically observed citation histories, from jump decay patterns to delayed impact. We also tested our model on all papers published in 1990 by 12 prominent journals (table S4), finding an exceptional collapse for all (see Fig. 1G, inset, for \( \text{Science} \) and supplementary materials S2.4.2 and fig. S8 for the other journals).

The model Eqs. 3 to 5 also predicts several fundamental measures of impact:

Ultimate impact (\( c^\infty \)) represents the total number of citations a paper acquires during its lifetime. By taking the \( t \rightarrow \infty \) limit in Eq. 3, we obtain

\[
c^\infty = m(c^\infty - 1)
\]

a simple formula that predicts that the total number of citations acquired by a paper during its lifetime is independent of immediacy (\( \mu \)) or the rate of decay (\( \sigma \)) and depends only on a single parameter, the paper’s relative fitness, \( \lambda \).

Impact time (\( T^I \)) represents the characteristic time it takes for a paper to collect the bulk of its citations. A natural measure is the time necessary for a paper to reach the geometric mean of its final citations, obtaining (supplementary materials S2.2)

\[
T^I \simeq \exp(\mu)
\]

Hence, impact time is mainly determined by the immediacy parameter \( \mu \) and is independent of fitness \( \lambda \), or decay \( \sigma \).

The proposed model offers a journal-free methodology to evaluate long term impact. To illustrate this, we selected three journals with widely different IFs: \( \text{Physical Review B (PRB)} \) (IF = 3.26 in 1992), \( \text{Proceedings of the National Academy of Sciences USA (PNAS)} \) (10.48), and \( \text{Cell (33.62)} \).

![Fig. 2. Evaluating long-term impact. (A) Fitness distribution \( P(\lambda) \) for papers published by \( \text{Cell, PNAS, and PRB} \) in 1990. Shaded area indicates papers in the \( \lambda = 1 \) range, which were selected for further study. (B) Citation distributions for papers with fitness \( \lambda = 1 \), highlighted in (A), for years 2, 4, 10, and 20 after publication. (C) Time-dependent relative variance of citations for papers selected in (A). (D) Citation distribution 2 years after publication [\( P(c^2) \)] for papers published by \( \text{Cell, PNAS, and PRB} \). Shaded area highlights papers with \( c^2 < 5 \) that were selected for further study. (E) Citation distributions for papers with \( c^2 < 5 \), selected in (D), after 2, 4, 10, and 20 years. (F) Time-dependent relative variance of citations for papers selected in (D).](image-url)
We measured for each paper published by them the fitness $\lambda$, obtaining their distinct journal-specific $P(\lambda)$ fitness distribution (Fig. 2A). We then selected all papers with comparable fitness $\lambda \approx 1$ and followed their citation histories. As expected, they follow different paths: Cell papers ran slightly ahead and PRB papers stay behind, resulting in distinct $P(c^\omega)$ distributions for years $T = 2 \div 4$. Yet, by year 20 the cumulative number of citations acquired by these papers shows a notable convergence to each other (Fig. 2B), supporting our prediction that given their similar fitness $\lambda$, eventually they will have the same ultimate impact: $c^\infty = 51.5$. To quantify the magnitude of the observed convergence, we measured the coefficient of variation $\sigma_c / c$ for $P(c^\omega)$, finding that this ratio decreases with time (Fig. 2C). This helps us move beyond visual inspection, offering quantitative evidence that in the long run the differences in citation counts between these papers vanishes with time, as predicted by our model. In contrast, if we choose all papers with the same number of citations at year two (i.e., the same $c^2$, Fig. 2D), the citations acquired by them diverge with time, and $\sigma_c / c$ increases (Fig. 2, E and F), supporting our conclusion that these quantities lack predictability. Therefore, $\lambda$ and $c^\infty$ offer a journal independent measure of a publication’s long-term time.

The model (Eqs. 3 to 5) also helps connect the IF, the traditional measure of impact of a scientific journal, to the journal’s $\Lambda$, $M$, and $\Sigma$ parameters (the analogs of $\lambda$, $\mu$, and $\sigma$; supplementary materials S4)

$$IF \approx \frac{m}{2} \left( \exp \left[ \frac{\Phi \left( \frac{M_\lambda - M}{\Sigma} \right)}{\lambda} \right] - \exp \left[ \frac{\Phi \left( \frac{M_\lambda - M}{\Sigma} \right)}{\mu} \right] \right)$$

Knowing $\Lambda$, in analogy with Eq. 6 we can calculate a journal’s ultimate impact as $C^\infty = m(e^\lambda - 1)$, representing the total number of citations a paper in the journal will receive during its lifetime. As we show in the supplementary materials S4, Eq. 8 predicts a journal’s IF in good agreement with the values reported by ISI (Institute for Scientific Information). Equally important, it helps us understand the mechanisms that influence the evolution of the IF, as illustrated by the changes in the impact factor of Cell and New England Journal of Medicine (NEJM). In 1998, the IFs of Cell and NEJM were 38.7 and 28.7, respectively (Fig. 3A). Over the next decade, there was a remarkable reversal: NEJM became the first journal to reach IF = 50, whereas Cell’s IF decreased to around 30. This raises a puzzling question: Has the impact of papers published by the two journals changed so dramatically? To answer this, we determined $\Lambda$, $M$, and $\Sigma$ for both journals from 1996 to 2006 (Fig. 3, D to F). Although $\Sigma$ were indistinguishable (Fig. 3D), we find that the fitness of NEJM increased from $\Lambda = 3.4$ (1996) to $\Lambda = 3.33$ (2005), increasing the journal’s ultimate impact from $C^\infty = 300$ (1996) to 812 (2005) (Fig. 3B). But Cell’s $\Lambda$ also increased in this period (Fig. 3E), moving its ultimate impact from $C^\infty = 366$ (1996) to 573 (2005). If both journals attracted papers with increasing long-term impact, why did Cell’s IF drop and NEJM’s grow? The answer lies in changes in the impact factor $T^* = \exp(M)$. Whereas NEJM’s impact time remained unchanged at $T^* \approx 3$ years, Cell’s $T^*$ increased from $T^* = 2.4$ years to $T^* = 4$ years (Fig. 3C). Therefore, Cell papers have gravitated from short- to long-term impact: A typical Cell paper gets 50% more citations than a decade ago, but fewer of the citations come within the first 2 years (Fig. 3C, inset). In contrast, with a largely unchanged $T^*$, NEJM’s increase in $\Lambda$ translated into a higher IF. These conclusions are fully supported by the $P(\lambda)$ and $P(\mu)$ distributions for individual papers published by Cell and NEJM in 1996 and 2005: Both journals show a shift to higher-fitness papers (Fig. 3G), but whereas $P(\mu)$ is largely unchanged for NEJM, there is a shift to higher-$\mu$ papers in Cell (Fig. 3H).

Can we use the developed framework to predict the future citations of a publication? For this, we adopted a framework borrowed from weather predictions and data mining: We used paper $i$’s citation history up to year $T_{\text{Train}}$ after publication (training period) to estimate $\lambda_i$, $\mu_i$, and $\sigma_i$ and then used the model Eq. 3 to predict its future citations $c_i^\omega$ and Eq. 6 to determine its ultimate impact $C^\infty$. The uncertainties in estimating $\lambda_i$, $\mu_i$, and $\sigma_i$ from the inherently noisy citation histories affect our

---

Fig. 3. Quantifying changes in a journal’s long-term impact. (A) IF of Cell and NEJM reported by Thomson Reuters from 1998 to 2006. (B) Ultimate impact $C^\infty$ (see Eq. 6) of papers published by the two journals from 1996 to 2005. (C) Impact time $T^*$ (Eq. 7) of papers published by the two journals from 1996 to 2005. (Inset) Fraction of citations that contribute to the IF. (D to F) The measured time-dependent longevity ($\Sigma$), fitness ($\Lambda$), and immediacy ($M$) for the two journals. (G) Fitness distribution for individual papers published by Cell (left) and NEJM (right) in 1996 (black) and 2005 (red). (H) Immediacy distributions for individual papers published by Cell (left) and NEJM (right) in 1996 (black) and 2005 (red).
predictive accuracy (supplementary materials S2.6). Hence, instead of simply interpolating Eq. 3 into the future, we assigned a citation envelope to each paper, explicitly quantifying the uncertainty of our predictions (supplementary materials S2.6). We show (Fig. 4A) the predicted most likely citation path (red line) with the uncertainty envelope (gray area) for three papers, based on a 5-year training period. Two of the three papers fall within the envelope; for the third, however, the model overestimated the future citations. Increasing the training period enhanced the predictive accuracy (Fig. 4B).

To quantify the model’s overall predictive accuracy, we measured the fraction of papers that fall within the envelope for all PR papers published in 1960s. That is, we measured the $z_{30}$ score for each paper, capturing the number of standard deviations ($z_{30}$) the real citations $c_{30}$ deviate from the most likely citation 30 years after publication. The obtained $P(z_{30})$ distribution across all papers decayed fast with $z_{30}$ (Fig. 4C), indicating that large $z$ values are extremely rare. With $T_{\text{Train}} = 5$, only 6.5% of the papers left the prediction envelope 30 years later; hence, the model correctly approximated the citation range for 93.5% of papers 25 years into the future.

The observed accuracy prompts us to ask whether the proposed model is unique in its ability to capture future citation histories. We therefore identified several models that either have been used in the past to fit citation histories or have the potential to do so: the logistic (26), Bass (27), and Gompertz (26, 28) models (for formulas, see supplementary materials and table S2).

We fit the predictions of these models to PR papers and used the weighted Kolmogorov-Smirnov (KS) test to evaluate their goodness of fit (see eq. S43 for definition), capturing the maximum deviation between the fitted and the empirical data. The lowest KS distribution across most papers was observed with Eq. 3, indicative of the best fit (Fig. 4D). The reason is illustrated in fig. S18:
The symmetric $r(t)$ predicted by the logistic model cannot capture the asymmetric citation curves. Although the Gompertz and the Bass models predict asymmetric citation patterns, they also predict an exponential (Bass) or double-exponential (Gompertz) decay of citations (table S2) that is much faster than observed in real data. To quantify how these deviations affect the predictive power of each of these models, we used a 5- and a 10-year training period to fit the parameters of each model and computed the predicted most likely citations at year 30 (Fig. 4, E and F). Independent of the training period, the predictions of the logistic, Bass, and Gompertz models always lay outside the 25 to 75% prediction quantiles (red bars), systematically underestimating future citations. In contrast, the prediction of Eq. 3 for both training periods was improving for the 10-year training period (Fig. 4F). In supplementary materials S3.3, we offer additional quantitative assessment of these predictions (fig. S19), demonstrating our model’s predictive power pertaining to both the fraction of papers whose citations it correctly predicts and the magnitude of deviations between predicted and the real citations. The predictive limitations of the current models were also captured by their $P(E_{30})$ distribution, indicating that for the logistic, Bass, and Gompertz models more than half of the papers underestimate with more than two standard deviations the true citations ($z > 2$) at year 30 (Fig. 4C), in contrast with 6.5% for the proposed model (Eq. 3).

Ignoring preferential attachment in Eq. 2 leads to the lognormal model, containing a lognormal temporal decay modulated by a single fitness parameter. As we analytically show in supplementary materials S3.4, for small fitness Eq. 3 converged to the lognormal model, which correctly captured the citation history of small impact papers. The lognormal model failed, however, to predict the citation patterns of medium- to high-impact papers (fig. S20). The proposed model therefore allows us to analytically predict the citation threshold when preferential attachment becomes relevant. The calculations indicate that the lognormal model is indistinguishable from the predictions of Eq. 3 for papers that satisfy the equation

$$
\sum_{n=2}^{\infty} \frac{1}{n!} \Phi^m \lambda^n < 1
$$

Solving this equation predicts $\lambda < 0.25$, equivalent with the citation threshold $c^* < 8.5$, representing the theoretical bound for preferential attachment to turn on. This analytical prediction is in close agreement with the empirical finding that preferential attachment is masked by initial attractiveness for papers with fewer than seven citations (29). Note that the lognormal function has been proposed before to capture the citation distribution of a body of papers (15). However, the lognormals appearing in (15) and in the lognormal model discussed above have different origins and implications (supplementary materials S2.5.2).

The proposed model has obvious limitations: It cannot account for exogenous “second acts,” like the citation bump observed for superconductivity papers after the discovery of high-temperature superconductivity in the 1980s, or delayed impact, like the explosion of citations to Erdős and Rényi’s work 4 decades after their publication, following the emergence of network science (3, 20, 21, 23).

Our findings have policy implications, because current measures of citation-based impact, from IF to Hirsch index (4, 17), are frequently integrated in reward procedures, the assignment of research grants, awards, and even salaries and bonuses (30), despite their well-known lack of predictive power. In contrast with the IF and short-term citations that lack predictive power, we find that $c^*$ offers a journal-independent assessment of a paper’s long-term impact, with a meaningful interpretation: It captures the total number of citations a paper will ever acquire or the discovery of its ultimate impact. Although additional variables combined with data mining could further enhance the demonstrated predictive power, an ultimate understanding of long-term impact will benefit from a mechanistic understanding of the factors that govern the research community’s response to a discovery.

References and Notes

Acknowledgments: The authors thank P. Azoulay, C. Hidalgo, J. Locial, D. Pedreschi, B. Uzzi, M. Vidal, and members of Center for Complex Network Research for insightful discussions and the anonymous referee for suggesting the lognormal model, which led to the analytical prediction of the citation threshold for preferential attachment. The PR data set is available upon request through American Physical Society. Supported by Lockheed Martin Corporation (SRA 11.18.11), the Network Science Collaborative Technology Alliance is sponsored by the U.S. Army Research Laboratory under agreement W911NF-09-2-0053, Defense Advanced Research Projects Agency under agreement 11645021, and the Future and Emerging Technologies Project 317532 “Multiplex” financed by the European Commission.

Supplementary Materials
www.sciencemag.org/content/342/6154/127/suppl/DC1
Materials and Methods
Supplementary Text
Figs. 1 to 525
Tables S1 to S4
References (S1–44)
14 March 2013; accepted 4 September 2013