Quality Research Follows the Power Law

Hanbin Mao[a,b]*, Jixin Chen[c,d]*

[a.] Department of Chemistry and Biochemistry, [b.] Advanced Materials and Liquid Crystal Institute, Kent State University, Kent, Ohio, USA

[c.] Department of Chemistry and Biochemistry, [d.] Nanoscale & Quantum Phenomena Institute, Ohio University, Athens, Ohio, USA

hmao@kent.edu, ORCID: 0000-0002-6720-9429

chenj@ohio.edu, ORCID: 0000-0001-7381-0918

Abstract

Research output can be evaluated with productivity and impact, which are quantified by the numbers of publications and citations, respectively. The H-index unifies both factors. However, as an extensive variable, it grows with quantity of research output and favors senior researchers over juniors. In this report, by analyzing the data of the world top 2% scientists (n = 179,597) from an online database, we found that H-index follows power laws with both impact and productivity in a minimalist fitting with only one free parameter. We propose intensive indices (Q_N and Q_C) to measure quality research by comparing the actual H-index of a researcher with the power-law fitted H-indices from the top 2% scientists with the same numbers of publications and citations respectively. We further calculated a dynamic research quality ($Q_1 = Q_N/Q_C$) and a reduced index ($Q_2 = (Q_N Q_C)^{0.5}$) to evaluate research quality over time. We rationalized that the power law dependency of quality research is due to its heterogeneous production pathways that require extra effort with respect to "regular" research output. A major advantage of these indices is that they are relative to the academic peers with similar accomplishments in publications and citations, and therefore, are independent of career stages. Another advantage is that the average indices are close to 1.0, giving an easily comprehensible physical significance of the indices.

Introduction

Research organizations often use publications and citations as part of their quantitative scholarship assessment. While publications measure the productivity of a researcher, citations reflect the research impact of the scholar.[1] Publications and citations are tracked in various databases such as Web of Science, SciFinder, Scopus, and Google Scholar. There are various reasons to cite a publication, e.g., to inspire and support a statement, to introduce the field and progress, to credit the contributions, among others. Since these reasons are difficult to distinguish, citations are usually treated as a lump sum popularity score received from peers. Both the number of publications (N) and citations (Nc) have played significant roles in many evaluation procedures including hiring and promotion of researchers, recruiting of graduate students and postdocs, and distribution of awards and grants. Therefore, it becomes important to prudently use these publication and citation data by fully understanding their properties and limitations.

In 2005, H-index was suggested by Jorge E. Hirsch as a unified tool to evaluate both productivity and impact of theoretical physicists.[2] It was designed to balance the productivity (i.e., number of publications, N) and the impact of the publications (i.e., number of citations received, Nc) by finding the number of papers (H) each of which is cited no less than H times. Due to its simplicity and inclusiveness for both research productivity and impact, it has been quickly adopted in scientific communities beyond physics.

Data show that H-index is an acceptable global index for a research community but has problems to evaluate research of individuals.[3] As an extensive variable, H-index favors researchers with more publications. **Fig. 1a** shows a thought experiment of four scholars. Scholar C has consistently higher citation per paper than scholar A or B. However, this fact is not reflected by the H-index as C has fewer publications. Even more problematic is for scholar D, who has a much smaller H-index but may have made a breakthrough contribution to the field. If one uses H-index as a threshold to group scholars, D has a higher chance to be excluded from the group. To address these problems, we envisage indices with an intensive property, such as Nc/N, should be used to evaluate research quality.

There have been efforts to modify and use H-index.[4] Based on power-law dependence observed between variables such as H-index, citations, and number of publications for individual scholars, institutes, journals, and countries,[4–19] correction factors of the H-indices for scholars in different fields have been proposed,[13] and new indices are proposed.[19] However, all these corrections still use extensive parameters, which varies significantly for a scholar in different career stages and different research fields. In this manuscript, we found that the H-indices of a large group of scholars in all fields also demonstrated the power law distributions against total citations and/or number of publications. Based on this correlation, we propose ratiometric, intensive indices that have the potential to be used to rank scholars in different fields and different stages of their careers.

Results and Discussion

We analyzed 179,597 researchers (estimated ~2% of total researchers) summarized by John Ioannidis and coworkers.[20] We found the relatively poor correlation between H-index and the research quality represented as *Nc*/*N* (**Fig. 1b**, indicated by the red arrow). As a result, if we randomly choose two scholars from the list, ~31% of time their H-index and average publication quality are oppositely correlated (**Fig. 1c**). Similar result of weak to no correlation between productivity (such as publications) and quality (such as citation per publication) has been observed in the literature.[1,3] Because the absolute value of average citation per publication of each scholar spans four orders of magnitude for the 2% scholars (**Fig. 1b**) in a typical subfield, there is a need to come up with simpler indices to measure the research quality.

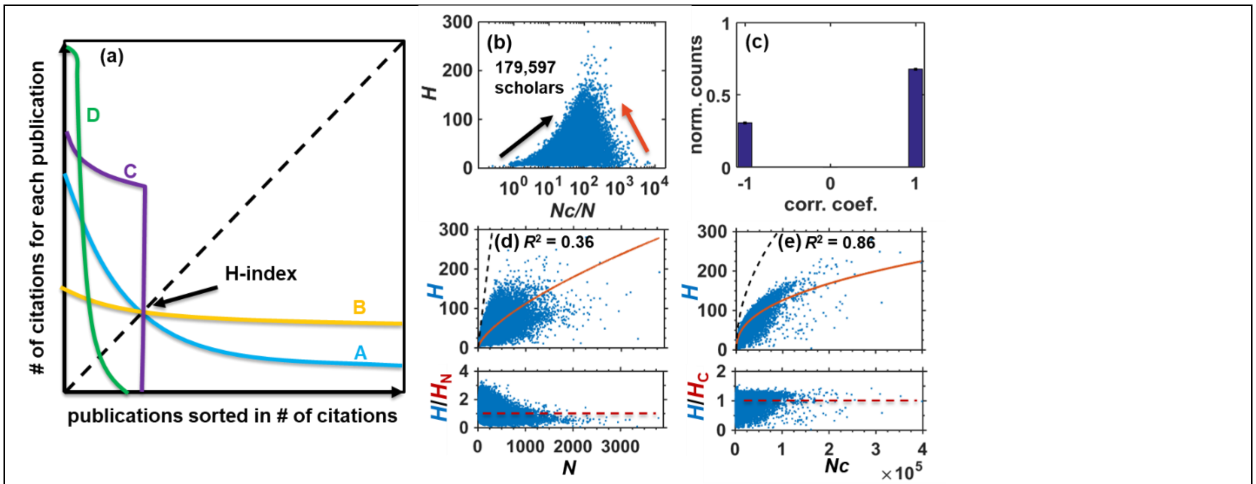


Fig. 1. (a) Thought distributions of four scholars. The *x*-axis is the publications of four scholars (A-D). The dashed line is where the value of the citations equals the number in the sorted publications. The intersection value between the dashed line and the citation line is the H-index for each scholar. (b) *H* vs *Nc*/*N* for 179,597 scholars from the world's top 2% researchers in the Scopus database.[20] (c) Histogram of correlation coefficients between *H* and *Nc*/*N* for 100,000 randomly selected pairs from all the top 2% scholars. (d) *H* vs *N*, a power law fitting (red), and *H* = *N* curve (dashed black line). The lower panel is the relative divergence of the data (H/H_N, see Eqn. 1 for H_N) from the fitting where the red line is the linear fitting. (e) *H* vs *N*c, a power law fitting (red), and the $H = (Nc)^{0.5}$ curve (dashed, which represents the maximum *H* one can obtain for a specific *N*c). The lower panel is the relative divergence (H/H_C, see Eqn. 3 for H_C) where the dashed red line is the fitted results normalized to one.

***H* index follows a power law relationship with the number of publications or citations.** To search for a quality index with a simple and intensive property to measure the research quality, we first investigated the relationship of H with published papers (productivity N) or citation numbers (impact, N_C). H-index is positively correlated with the productivity (**Fig. 1d**) or impact (**Fig. 1e**). The distributions show a power-law like Pareto front[21] following the scattered data points.

After we plotted the H-index vs number of publications (N) for the top 2% researchers[20] from all scientific disciplines, we found that H-index vs N can be fit with a power law using a minimalist fitting without free parameters (**Fig. 1d**),

$$\boldsymbol{H_N = N^n} \qquad (\mathbf{1}).$$

For all top 2% scientists, $n = 0.68$. The goodness of the fit $R^2 = 0.36$, and the relative divergence of the data to the fitting varies from 0 to 3. We can use the divergence as a parameter (Q_N) to measure the distance of a particular scholar to the fitted "average" line of the top 2% scholars with the same N values,

$$\boldsymbol{Q_N = H/H_N} \qquad \mathbf{(2).}$$

Q_N is ratiometric and intensive. The distribution of Q_N of all top 2% scholars can be fitted with three Gaussian distributions (**Fig. 2a**). The main peak is centered at 1.14 with $\sigma = 0.32$.

Interestingly, H index vs number of citations (Nc) can be fit with a better power law expression (**Fig. 1e**),

$$\boldsymbol{H_C = N_C^{\,c}} \qquad (\mathbf{3}).$$

For all top 2% scientists, $c = 0.42$. The goodness of the fit $R^2 = 0.86$, and the relative divergence varies from 0 to 1.5. We can also create a parameter (Q_C) based on the divergence by referencing the "average" H-index of the scholars with the same Nc values,

$$\boldsymbol{Q_C = H/H_C} \qquad (\mathbf{4}).$$

Like Q_N, Q_C is an intensive variable. The distribution of Q_C of all top 2% scholars can be fitted with three Gaussian distributions as well (**Fig. 2b**). The main peak centers at 1.07 with sigma 0.08.

It is significant that the asymmetric distributions in Q_N and Q_C have a mirror pattern (**Fig. 2**). While Q_N shows a tail towards larger values, Q_C depicts a bias for smaller numbers, suggesting that the number of publications should bear less importance (H_N underestimated with respect to H-index) than the number of citations (H_C overestimated relative to H-index) in the H-index calculation. In addition, the asymmetric distribution patterns in Q_N or Q_C make it difficult to quickly evaluate research performance using these two variables.

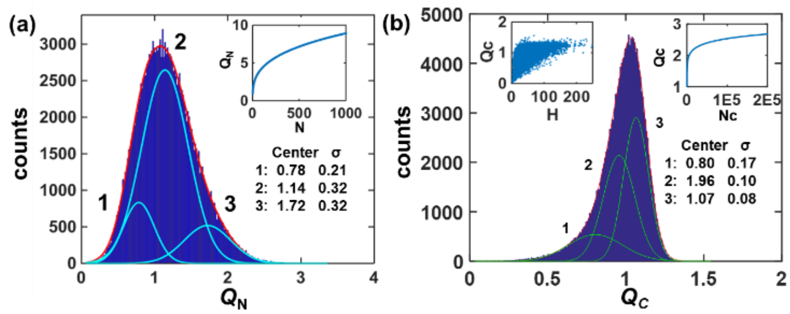


Fig. 2. (a) Histogram of Q_N in top 2% scientists fitted with three Gaussian peaks. Inset shows $H = N$ curve. (b) The distribution of Q_C fitted with three Gaussian peaks. Insets show the scattered plot of Q_C vs H (left) and the $H = (Nc)^{0.5}$ curve (the maximum H value for a particular Nc, right).

Since H-index is defined in such a way that both N and Nc contribute to its value, it is not surprising that the H-index somehow increases with either N or Nc. The fact that the fitting is much better in Nc than N suggests it is the citation, instead of the publication number, that contributes more significantly to the H-index, which is consistent with the asymmetric Q_N and Q_C distributions discussed above. If we consider N represents research output that is a function of total research investment and H represents the output of quality work, then, the power-law relationship between H and N implies that many heterogenous pathways are involved to spend the total research investment (will discuss more in the later sections).[22] By the same token, the power law relationship between H and Nc also suggests heterogenous output for the Nc received. It is conceivable that not all citations positively contribute to the quality work. For example, negative citations may reduce the impact of cited work in the long run by alarming peer scientists. In addition, self-citations may also contribute to the heterogeneous citation distributions that lead to the power law behavior of H vs Nc.

In the simplest form, the average research quality can be represented by citations per published paper (N_C/N) for a given researcher. When we scrutinized Q_N and Q_C, we found Q_N correlates with N_C/N reasonably well (**Fig. 3a**) but bend to anticorrelation for scholars with large N_C/N values (e.g. >500). Q_C carries significant anti-correlation with N_C/N which is inherited from the anti-correlation between H and N_C/N (**Fig. 3b**, indicated by the red circle and arrow). We further propose that a dynamic research quality may be calculated by newly received citations per newly published paper each year for the researcher. This calculation, however, does not distinguish the academic stages of the researcher, nor does it offer a perspective on a comparable basis (e.g. scientists with similar publications and/or citations). As a result, senior researchers gain advantages as new citations received each year are cumulative over all their prior publications.

To address this problem, we introduced a ratiometric research quality index, Q_1 to reflect the reduced quality per paper with respect to that of the "average" from the top 2% scholars,

$$\boldsymbol{Q_1 = Q_N/Q_C = H_C/H_N = N_C^{c}/N^{n}} \qquad (\mathbf{5}).$$

Here, Q_1 depicts the reduced Nc/N with respect to the global power law distribution constants n and c among the top 2% scholars. The distribution of Q_1 among all top 2% scholars can also be fitted with three Gaussian peaks with the average centered ~1.0 for all top 2% scholars. Compared to H, Q_1 has a better linear relationship with Nc/N (compare **Fig. 1b** and **Fig. 3c**). Further correlation analysis confirms that Q_N or Q_1 correlates with Nc/N (each has ~83% correlation coefficients) better than H-index (~70%) (**SI Fig. S1 a-c**).

However, histogram of the Q_1 from all top 2% scientists in **Fig. 3d** also shows that Q_1 does not follow a perfect Gaussian distribution, which is not surprising as the Q_1 calculation (**Eqn. 5**) synergizes the unsymmetric biases in Q_N and Q_C (**Fig. 2**) by taking their ratio. To quest for a quality research index with symmetric Gaussian distribution, we reasoned that a unified equation combining **Eqns. 1** and **3** may give a more accurate representation of H to both N and N_C. To this end, we arranged the N, Nc, and H in a 3D plot with the 3D fitting using the following minimalist power law expression without free parameters,

$$H_{\mathrm{NC}} = N^{a}N_{C}^{b} \qquad (\mathbf{6}).$$

For all top 2% scientists, $a = 0.12$, $b = 0.35$ (**Fig. 3e**). This correlation agrees with Schubert-Glänzel's finding in which the journal H-index (H_J) is related to the number of papers (N) in the journal and its impact factor (IF) ($H_J \propto N^{1/3}\ IF^{2/3}$).[6] Our plot demonstrates the power law relationship among H, Nc, and N with the goodness of the fit $R^2 = 0.89$ for the published data among top 2% scholars. Such a fitting is better than that of the power law shown in **Eqn. 1** or **Eqn. 3**. Close inspection on the fitting reveals that the scholars on the left side of the plot are below the plane (indicated by the red arrow in **Fig. 3e**), indicating their H-indices are smaller than the fitted H_{NC}. This may represent type C and D scholars with high quality publications (**Fig. 1a**), which are not identified by the H-index method. In addition, the relatively narrow Gaussian distributions of the divergence and the improvement in distribution symmetry (**Fig. 3f**) provide a promising direction to generate a more accurate index with more symmetric Gaussian distribution.

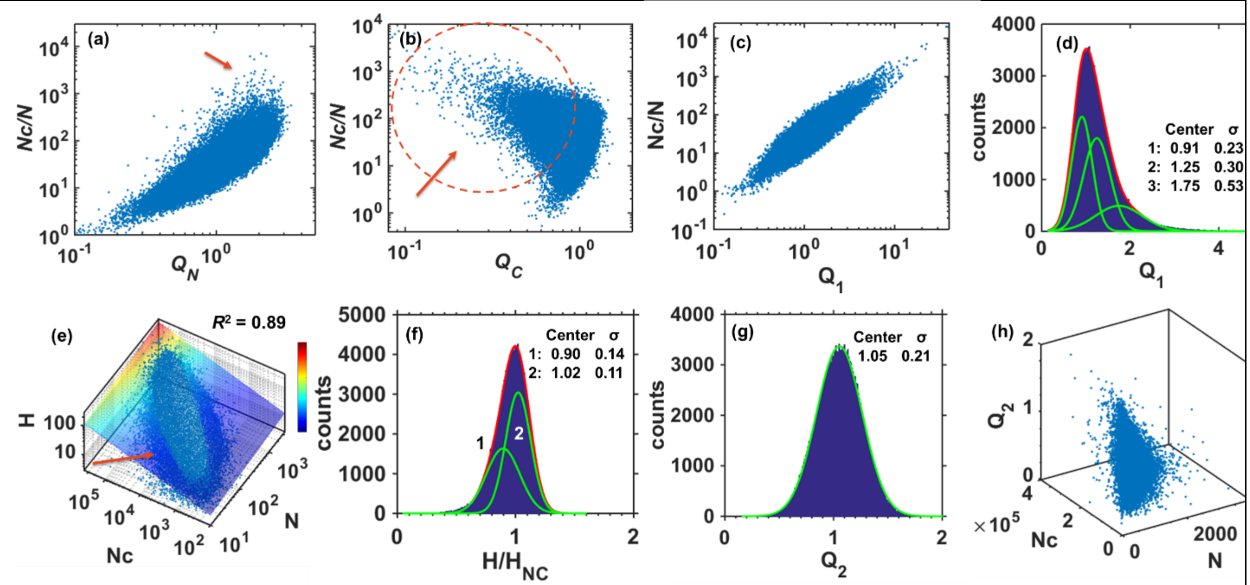


Fig. 3. (a) Correlation between N_C/N and Q_N, and (b) N_C/N vs Q_C of 179,597 scholars among the top 2% scientists.[20] The red arrows indicate the anti-correlated regions. (c) The positive correlation between Q_1 and N_C/N. (d) Distribution of Q_1 among all top 2% scholars. (e) The 3D plot of N, Nc, and H with a fitting using **eqn. 6**. The colored plane surface represents the fitted H_{NC} results. The lighter blue is above the fitting plane and the darker blue is below. The red arrow indicates the group of scholars with anticorrelated N_C and H values. See **SI Fig. S2** for views from other angles. (f) H/H_{NC} distribution fitted by two Gaussian peaks. (g) Distribution of Q_2 and (h) its relationship with N and N_C among all top 2% scholars. See **SI Fig. S3** for views from other angles.

Since H can be represented by the power laws of both N and N_C, it is convenient to express an average H_{NC}' by taking the geometric mean of H_N and H_C to keep the same units, $H_{NC}'=(H_N H_C)^{1/2}$. From Q_N, Q_C, and H_{NC}, we then propose another composite ratiometric variable,

$$\boldsymbol{Q_2 = H/H_{NC}' = H/(H_N H_C)^{1/2} = (Q_N Q_C)^{1/2}} \qquad \mathbf{(7)},$$

to evaluate the quality research of a scientist with respect to the average H_{NC} among top 2% scientists with similar publication and citation records.

It is clear that Q_2 is an almost perfect Gaussian distribution for the top 2% scholars (**Fig. 3g**). The average values among peers are again designed to ~1.0. The distribution of Q_2 over N and N_C is bent comparing to H (**Fig. 3h**). Unlike H that goes up with the increasing N, the peak distribution Q_2 now is seen at $N \sim 100$ (**Fig. 3h**), which is consistent with the finding that quality research is not correlated with the number of publications.[1,3] The anti-correlation between quality and quantity becomes obvious when analyzing the relationship between Q_2 and N (anti-correlation: 56%), as well as between Q_1 and N (anti-correlation: 71%) (**SI Fig S8**), both which are higher than that between N and N_C/N for randomly selected two scholars among the top 2% scholars (anti-correlation: 54%).

Table 1. Number of scholars of the fields and their average parameters.

Fields	**#People (p)**	**<H>**	**<Nc>**	**<N>**	**<Nc/y/p>**	**<N/y/p>**	**n**	**c**	n/c	Q_N	Q_C	Q_1	Q_2
All	**178754**	**38.00**	**7929**	**194**	**246.9**	**6.04**	**0.683**	**0.417**	**1.64**	**1.04**	**0.90**	**1.16**	**0.97**
Agriculture	5634	31.00	4011	138	125.0	4.30	**0.680**	0.419	1.62	1.07	0.97	1.10	1.02
Biology	7293	37.25	6845	133	207.4	4.03	**0.724**	0.416	1.74	1.32	0.94	1.41	1.11
Bio Medical	15247	46.85	11348	188.5	314.9	5.23	**0.721**	0.420	1.72	1.31	0.96	1.37	1.12
Chemistry	12229	38.51	7331	217	225.3	6.67	**0.666**	0.420	1.59	0.98	0.94	1.04	0.96
Clinical	55693	46.00	7929	237.06	230.8	6.90	**0.691**	0.420	1.64	1.10	1.09	1.01	1.09
Earth	6306	39.00	6951	150.9	211.9	4.60	**0.711**	0.418	1.70	1.27	0.98	1.30	1.11
Economics	3364	29.00	5547	91	189.0	3.10	**0.721**	0.395	1.83	1.33	0.80	1.67	1.03
Enabling Tech.	14485	33.00	5756	188	211.3	6.90	**0.655**	0.416	1.57	0.92	0.89	1.03	0.91
Engineering	14229	27.69	3880	171.83	134.4	5.95	**0.635**	0.412	1.54	0.82	0.88	0.93	0.85
History	786	19.00	1940	63.1	55.3	1.80	**0.704**	0.400	1.76	1.12	0.81	1.39	0.95
Information	11309	27.88	5072	179	193.2	6.82	**0.628**	0.399	1.57	0.81	0.79	1.02	0.80
Mathematics	2159	26.00	4055	136	125.2	4.20	**0.645**	0.401	1.61	0.91	0.81	1.12	0.86
Nano	4147	48.53	12686	194	626.5	9.58	**0.699**	0.420	1.66	1.33	0.94	1.41	1.12
Physics	17252	37.00	7896	223	247.1	6.98	**0.650**	0.412	1.58	0.92	0.88	1.05	0.90
Psychology	3550	39.00	7770	139.75	228.9	4.12	**0.723**	0.414	1.75	1.34	0.93	1.44	1.12
Public Health	4005	31.31	5141	148	170.6	4.91	**0.679**	0.415	1.64	1.03	0.89	1.16	0.96
Visual Arts	113	9.40	419	36.49	14.9	1.30	**0.584**	0.362	1.61	0.81	0.76	1.06	0.79
Social	3876	23.23	3139	79.76	105.5	2.68	**0.699**	0.396	1.77	1.17	0.81	1.44	0.98

Notes: Please see the fittings and analysis in **SI Fig. S5-S6**.

Since these analyses have been performed on 179597 (top 2%) scientists, it is questionable whether the power laws persist in the rest of the population or in a population with fewer people. To test whether power laws are valid in scientists ranked with different percentiles, we separated the top 2% into two groups of top 1% and 1-2% populations using either the composite score (C-score),[23] the H-index, the N, or the N_C ranking.[20] We found that while power law relationships are still valid for populations with different percentiles, the power exponents of $\boldsymbol{n}$ (**Eqn. 1**) and $\boldsymbol{c}$ (**Eqn. 2**) do decrease with decreasing percentiles (**SI Fig. S4**). In addition, we also evaluated all 18 disciplines with number of researchers ranging from 113 (Visual Arts and Performance) to 55693 (Clinical Medicine) (**Table 1**), the power laws are valid for H vs N as well as H vs N_C relationships (**SI Figs S5, S6**).

Therefore, with these power-law derived Q indices, it is possible to evaluate the publication quality of any researcher with reference to the top 2% scientists. The same quality indices can also be used to evaluate any scientific community as exemplified by the analyses of the 18 disciplines summarized in **Table 1**.

Dynamic research quality (Qs) for individual researchers over time. According to our definition, the Q indices are integrated functions that reflect the quality research (H-index) with respect to the average H obtained from a population. This value is expected to evolve over time for a group of selected scientists with certain age or academia stage, or for individual research fields. However, it does not measure the dynamic research quality needed to interrogate the Q-indices over time for individual researchers. One way to achieve this task is to plot the Q-indices vs time for a researcher. We analyzed the four Q indices over time for an exemplary researcher, Dr. Roger Y. Tsien, in **Fig. 4**, after referencing him against the top 2% scientists in 2021. We observed an increase in all four Q-indices for Dr. Tsien over the years, indicating a stable growth of impacts with respective to the top 2% scientists. Q_N and Q_1 clearly indicate his breakthrough years (development of Green Fluorescence Protein (GFP[24–26]) when the values climb up to 1.7-1.8, which are significantly above the average values (~1.0) among the world top 2% scientists on the 2021 list (at ~10% percentile of the top 2%). Both Q_N and Q_1 are more stable and indicative than H, N, Nc, or Nc/N, or their respective slopes (**SI Fig. S9**), in distinguishing Dr. Tsien's research stages. As the geometric mean between Q_N and Q_C, the Q_2 values grow from below the average of the top 2% scientists, quickly to slightly above the average when Dr. Tsien became an independent researcher, which continues to grow over the rest of his career. When Dr. Tsien won the Nobel Prize of Chemistry in 2008, his Q_2 score was 1.6, which places him within top 0.3% percentile of the top 2% scholars in 2021 (600/179697). Between the two Q_1 and Q_2 indices defined by **Eqn. 6** and **Eqn. 7**, respectively, Q_1 is more sensitive to the average quality of the publications whereas Q_2 is more sensitive to the total quality of the publications.

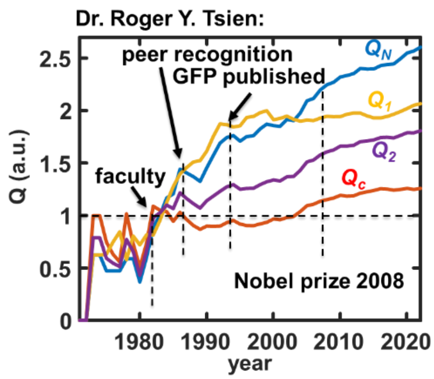


Fig. 4. Exemplary Q_N, Qc, Q_1, and Q_2 analyses of Dr. Roger Y. Tsien's (Nobel laureate of Chemistry 2008) citation data over time.

A model for research quality vs quantity. The power law behavior of H vs Nc and/or N resembles that of the photoluminescent lifetime of individual quantum dots.[22] When a single quantum dot gets excited, it will take one of many possible decay pathways to either emit a visible photon (radiative decays) or phonons (nonradiative decays). As a result, the emission lifetime (proportional to the quantum yield) of each particle will follow a distribution with weakly ergodicity feature that can be described with the power-law distribution.[22] The ensemble behavior of many particles is also power-law dependent rather than a Maxwell-Boltzmann or black-body distribution for well-behaved ergodic particles.[22,27] These observations prompted us to propose a multi-state, multi-pathway model to delineate the relationship between resource input and research output (**Fig 5**).

When a researcher in resting state receives resources (***I***), such as research funding, space, personnel, materials, etc., they rise to the loaded state from the resting state. The relaxation pathway (attrition) follows from the loaded state, which includes training of personnel, preparation for experiments, consumable expenditures (space, instrument, electricity), etc., before the output state is reached. From the output state, at least three heterogenous relaxation pathways exist to produce a total of $\boldsymbol{N}$ publications. In the first pathway, effort is futile, resulting in no publications. In the second pathway, ($\boldsymbol{N}$-$\boldsymbol{H}$) publications are produced for a regular output. In the third pathway, extra effort must be spent to reach the quality state from the output state to generate quality research product ($\boldsymbol{H}$). Overall, $H=N^n$, where N is dependent on the function of resource input as $\boldsymbol{N}=f(\boldsymbol{I})$. Just like the relaxation of excited quantum dots via multiple pathways,[22] heterogeneous pathways are involved in the relaxation of the loaded state back to the resting state of researchers. Therefore, we anticipate $f(\boldsymbol{I})$ follows the power law as well.

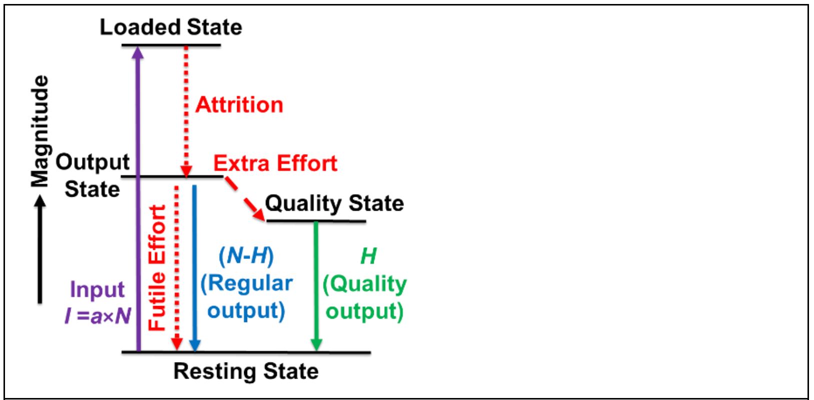


Fig. 5. The M-C diagram of publication states.

Summary and perspective

After analyses on the research record of top 2% scientists, we have revealed that quality research represented by H-index follows a power law relationship with the number of publications and/or citations. Such a behavior is akin to the Pareto distribution originated from the heterogeneity in the system.[22,28–30] We have therefore proposed that high quality publication is a result of the heterogeneous relaxation pathways from the loaded state after a scientist has obtained research resources. Such a power law dependency between H-index and publications/citations allows us to propose indices (Qs) with the intensive property to evaluate the quality research (using H-index as a virtual state indicator) of individual scientists with respect to that predicted from the reference group (e.g., top 2% scientists) with the same citation or publication records. These indices address the problem of H-index, which, due to its extensive nature, favors senior scientists with more publications over juniors.

Like any known quantitative method, predicting or judging young scholars based on these Q indices is dangerous for hiring or promotion purposes because these parameters, although intensive, are still affected by the popularity of a particular field if a reference group is not chosen prudently. They are rather noisy especially at the early stages of a researcher due to the insufficient information for a robust statistical analysis. Thus, it can be biased towards benefitting the scientists who are working in a hot area with its future prospect uncertain. In addition, it is still challenging to predict or to analyze publication records using criteria such as scientific innovation and social impacts.[1,31–34] We believe there is no simple way to quantitatively judge individual scholars.[3,35,36] Therefore, institutions should use caution when adopting these parameters for evaluation purposes.

Methods

Data treatment and analysis method have been included in the manuscript's main text. The analysis can be done using Excel and MATLAB in this manuscript. The Q values for Dr. Tsien are analyzed from the raw citation data downloaded from Scopus using a home-written MATLAB code. Please see SI Appendix 02 for the MATLAB code.

Data, Materials, and Software Availability

Raw data are downloaded from online source cited in the manuscript and Scopus. All other data are included in the manuscript and/or SI appendix 01. MATLAB codes are included in the SI appendix 02.

Conflict of Interest

A copyright disclosure is under preparation for the method described in this manuscript.

Acknowledgments

We thank Dr. Yuening Zhang and Dr. Tao Xu for beneficial discussion; Korie Sell and Albert Green for the disclosure preparation.

Supporting Information

Additional figures are included in the supporting information.

References

(1) Zeng, A.; Fan, Y.; Di, Z.; Wang, Y.; Havlin, S. Impactful Scientists Have Higher Tendency to Involve Collaborators in New Topics. *PNAS* **2022**, *119* (33), e2207436119. https://doi.org/10.1073/pnas.2207436119.

(2) Hirsch, J. E. An Index to Quantify an Individual's Scientific Research Output. *PNAS* **2005**, *102* (46), 16569. https://doi.org/10.1073/PNAS.0507655102.

(3) Kreiner, G. The Slavery of the H-Index—Measuring the Unmeasurable. *Front. Hum. Neurosci.* **2016**, *10*. https://doi.org/10.3389/fnhum.2016.00556.

(4) Glänzel, W.; Moed, H. F.; Schmoch, U.; Thelwall, M. *Springer Handbook of Science and Technology Indicators*; Springer, 2019. https://doi.org/10.1007/978-3-030-02511-3.

(5) Bailón-Moreno, R.; Jurado-Alameda, E.; Ruiz-Baños, R.; Courtial, J. P. Bibliometric Laws: Empirical Flaws of Fit. *Scientometrics* **2005**, *63* (2), 209–229. https://doi.org/10.1007/s11192-005-0211-5.

(6) Schubert, A.; Glänzel, W. A Systematic Analysis of Hirsch-Type Indices for Journals. *J. Informetr.* **2007**, *1* (3), 179–184.

(7) Montazerian, M.; Zanotto, E. D.; Eckert, H. A New Parameter for (Normalized) Evaluation of H-Index: Countries as a Case Study. *Scientometrics* **2019**, *118* (3), 1065–1078. https://doi.org/10.1007/s11192-018-2996-z.

(8) Schubert, A.; Korn, A.; Telcs, A. Hirsch-Type Indices for Characterizing Networks. *Scientometrics* **2009**, *78* (2), 375–382. https://doi.org/10.1007/s11192-008-2218-1.

(9) Barcza, K.; Telcs, A. Paretian Publication Patterns Imply Paretian Hirsch Index. *Scientometrics* **2009**, *81* (2), 513–519. https://doi.org/10.1007/s11192-008-2175-8.

(10) Molinari, A.; Molinari, J.-F. Mathematical Aspects of a New Criterion for Ranking Scientific Institutions Based on the H-Index. *Scientometrics* **2008**, *75* (2), 339–356. https://doi.org/10.1007/s11192-007-1872-z.

(11) Ye, F. Y. A Unification of Three Models for the H-index. *JASIST* **2011**, *62* (1), 205–207. https://doi.org/10.1002/asi.21456.

(12) Egghe, L.; Rousseau, R. An Informetric Model for the Hirsch-Index. *Scientometrics* **2006**, *69* (1), 121–129. https://doi.org/10.1007/s11192-006-0143-8.

(13) Iglesias, J.; Pecharromán, C. Scaling the H-Index for Different Scientific ISI Fields. *Scientometrics* **2007**, *73* (3), 303–320. https://doi.org/10.1007/s11192-007-1805-x.

(14) Fred, Y. Y.; Rousseau, R. The Power Law Model and Total Career H-Index Sequences. *J. Informetri.* **2008**, *2* (4), 288–297. https://doi.org/10.1016/j.joi.2008.09.002.

(15) Spearman, C. M.; Quigley, M. J.; Quigley, M. R.; Wilberger, J. E. Survey of the h Index for All of Academic Neurosurgery: Another Power-Law Phenomenon? *J. Neurosurg.* **2010**, *113* (5), 929–933. https://doi.org/10.3171/2010.4.JNS091842.

(16) Quigley, M. R.; Holliday, E. B.; Fuller, C. D.; Choi, M.; Thomas, C. R. Distribution of the H-Index in Radiation Oncology Conforms to a Variation of Power Law: Implications for Assessing Academic Productivity. *J. Cancer Educ.* **2012**, *27* (3), 463–466. https://doi.org/10.1007/s13187-012-0363-y.

(17) Brzezinski, M. Power Laws in Citation Distributions: Evidence from Scopus. *Scientometrics* **2015**, *103* (1), 213–228. https://doi.org/10.1007/s11192-014-1524-z.

(18) Glänzel, W. On the H-Index - A Mathematical Approach to a New Measure of Publication Activity and Citation Impact. *Scientometrics* **2006**, *67* (2), 315–321. https://doi.org/10.1007/s11192-006-0102-4.

(19) Egghe, L. Theory and Practise of the G-Index. *Scientometrics* **2006**, *69* (1), 131–152. https://doi.org/10.1007/s11192-006-0144-7.

(20) Ioannidis, J. P. A.; Baas, J.; Klavans, R.; Boyack, K. W. A Standardized Citation Metrics Author Database Annotated for Scientific Field. *PLOS Biol.* **2019**, *17* (8), e3000384. https://doi.org/10.1371/JOURNAL.PBIO.3000384.

(21) Pareto, V. *Cours d'économie Politique*; Librairie Droz, 1964; Vol. 1.

(22) Cichos, F.; von Borczyskowski, C.; Orrit, M. Power-Law Intermittency of Single Emitters. *Curr. Opini. Colloid Interface Sci.* **2007**, *12* (6), 272–284. https://doi.org/https://doi.org/10.1016/j.cocis.2007.07.012.

(23) Ioannidis, J. P. A.; Klavans, R.; Boyack, K. W. Multiple Citation Indicators and Their Composite across Scientific Disciplines. *PLOS Biol.* **2016**, *14* (7), e1002501. https://doi.org/10.1371/journal.pbio.1002501.

(24) Heim, R.; Prasher, D. C.; Tsien, R. Y. Wavelength Mutations and Posttranslational Autoxidation of Green Fluorescent Protein. *PNAS* **1994**, *91* (26), 12501–12504. https://doi.org/10.1073/pnas.91.26.12501.

(25) Heim, R.; Cubitt, A. B.; Tsien, R. Y. Improved Green Fluorescence. *Nature* **1995**, *373* (6516), 663–664. https://doi.org/10.1038/373663b0.

(26) Ormö, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. Crystal Structure of the Aequorea Victoria Green Fluorescent Protein. *Science (1979)* **1996**, *273* (5280), 1392–1395. https://doi.org/10.1126/science.273.5280.1392.

(27) Liu, C. S.; Kauffman, J. F. Excitation Power Dependence of Photoluminescence Enhancement from Passivated GaAs. *Appl. Phys. Lett.* **1995**, *66* (25), 3504–3506. https://doi.org/10.1063/1.113778.

(28) Andriani, P.; McKelvey, B. From Skew Distributions to Power-Law Science. *The sage handbook of complexity and management* **2011**, 254–273. https://doi.org/10.4135/9781446201084.n15.

(29) Muchnik, L.; Pei, S.; Parra, L. C.; Reis, S. D. S.; Andrade Jr, J. S.; Havlin, S.; Makse, H. A. Origins of Power-Law Degree Distribution in the Heterogeneity of Human Activity in Social Networks. *Sci. Rep.* **2013**, *3* (1), 1783. https://doi.org/10.1038/srep01783.

(30) Hu, X.; Zhang, Y.; Guzun, D.; Ware, M. E.; Mazur, Y. I.; Lienau, C.; Salamo, G. J. Photoluminescence of InAs/GaAs Quantum Dots under Direct Two-Photon Excitation. *Sci Rep* **2020**, *10* (1), 10930. https://doi.org/10.1038/s41598-020-67961-z.

(31) Aagaard, K.; Kladakis, A.; Nielsen, M. W. Concentration or Dispersal of Research Funding? *Quant. Sci. Stud.* **2020**, *1* (1), 117–149. https://doi.org/10.1162/qss_a_00002.

(32) Xu, F.; Wu, L.; Evans, J. Flat Teams Drive Scientific Innovation. *PNAS* **2022**, *119* (23), e2200927119. https://doi.org/10.1073/pnas.2200927119.

(33) Wu, L.; Wang, D.; Evans, J. A. Large Teams Develop and Small Teams Disrupt Science and Technology. *Nature* **2019**, *566* (7744), 378–382. https://doi.org/10.1038/s41586-019-0941-9.

(34) Perrin, B. How to—and How Not to—Evaluate Innovation. *Evaluation* **2002**, *8* (1), 13–28. https://doi.org/10.1177/1358902002008001.

(35) Wendl, M. C. H-Index: However Ranked, Citations Need Context. *Nature* **2007**, *449* (7161), 403. https://doi.org/10.1038/449403b.

(36) von Bohlen und Halbach, O. How to Judge a Book by Its Cover? How Useful Are Bibliometric Indices for the Evaluation of “Scientific Quality” or “Scientific Productivity”? *Ann. Anat. Anat. Anz.* **2011**, *193* (3), 191–196. https://doi.org/https://doi.org/10.1016/j.aanat.2011.03.011.

Supplementary Information

Quality Research Follows the Power Law

Hanbin Mao[a,b]*, Jixin Chen[c,d]*

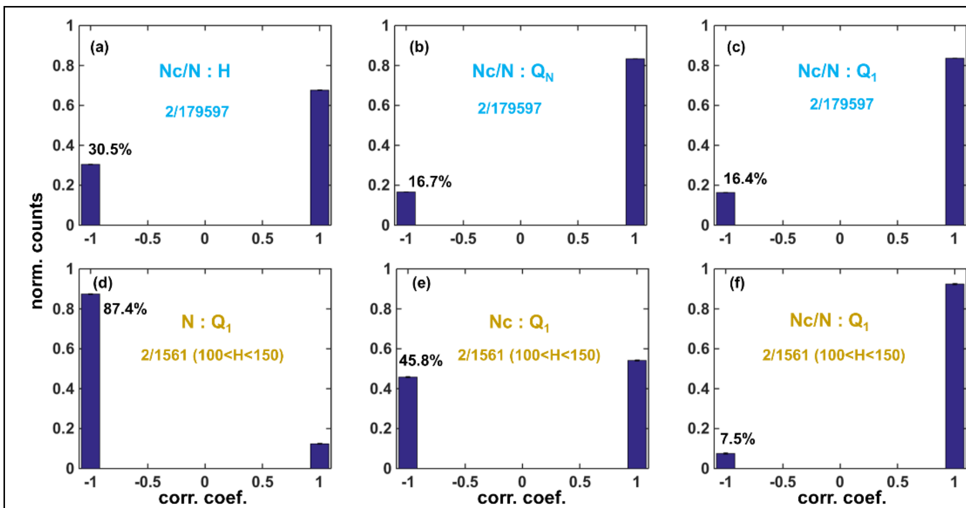
[a.] Department of Chemistry and Biochemistry, [b.] Advanced Materials and Liquid Crystal Institute, Kent State University, Kent, Ohio, USA

[c.] Department of Chemistry and Biochemistry, [d.] Nanoscale & Quantum Phenomena Institute, Ohio University, Athens, Ohio, USA

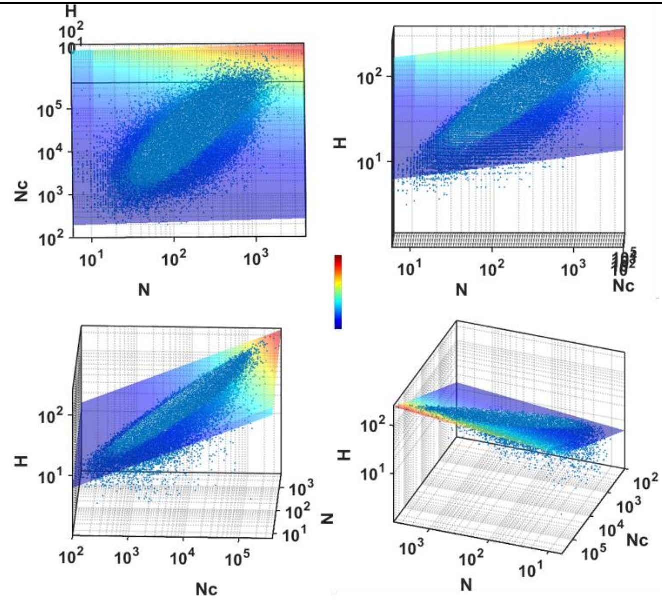
hmao@kent.edu, ORCID: 0000-0002-6720-9429

chenj@ohio.edu, ORCID: 0000-0001-7381-0918

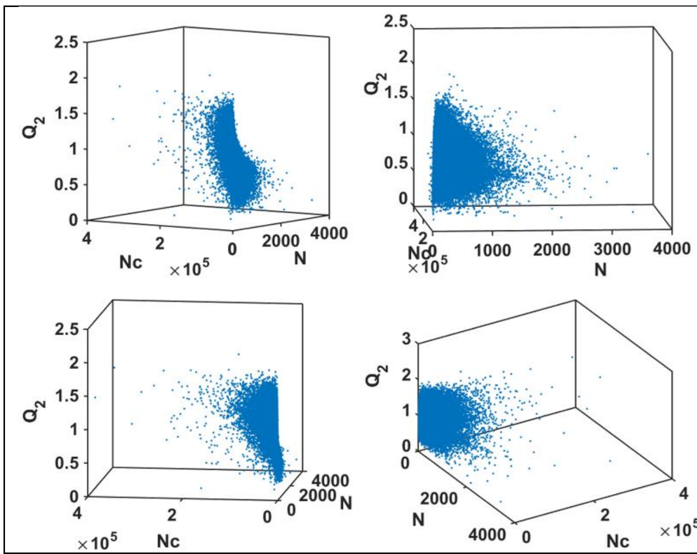
To test whether Q_N or Q_1 is better than H to define the research quality, we evaluated the correlation values of Q_N or Q_1 against N_c/N in randomly selected subgroups of two researchers from top 2% scientists (**SI Fig. S1 a-c**). Q_N and Q_1 both reduce the negative correlation to half with respect to H against N_C/N. So Q_N and Q_1 are better than H to reflect average publication quality. The negative correlation between N_C/N and Q_1 drops to 7.5% if two scholars are randomly selected out of 1561 scholars with H-index between 100 to 150. Q_1 also catches the negative correlation with productivity (N) among scholars with similar H-indices, and it has almost no-correlation with N_C/N, both as expected (**SI Fig. S1 d-e**). Thus, Q_1 is an acceptable index for the research quality depicted by citations per publication (N_C/N) with >83% confidence for all top 2 scholars and >92% confidence for scholars with H-index between 100 to 150.



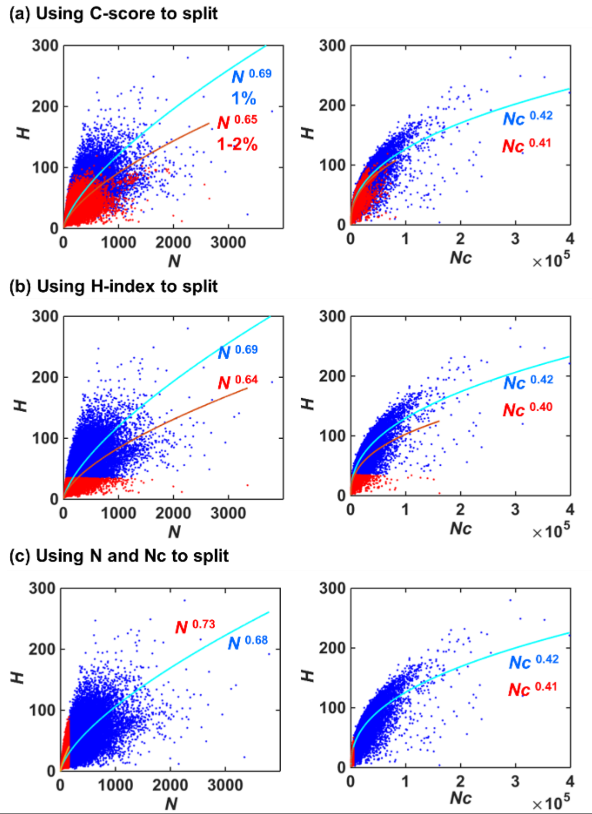
SI Fig. S1. Fractions of correlation (coefficient = +1) and anticorrelation (coefficient = -1) of 100,000 randomly selected pair of scholars from the top 2% scientists,[20] (a)Nc/N vs H, (b) Nc/N vs QN, (c) Nc/N vs Q1. (d-f) 10,000 randomly selected pairs of scholars from 1561 scholars with H-index 100-150.



SI Fig. S2. Random angles of **Fig. 3e** in the main text.

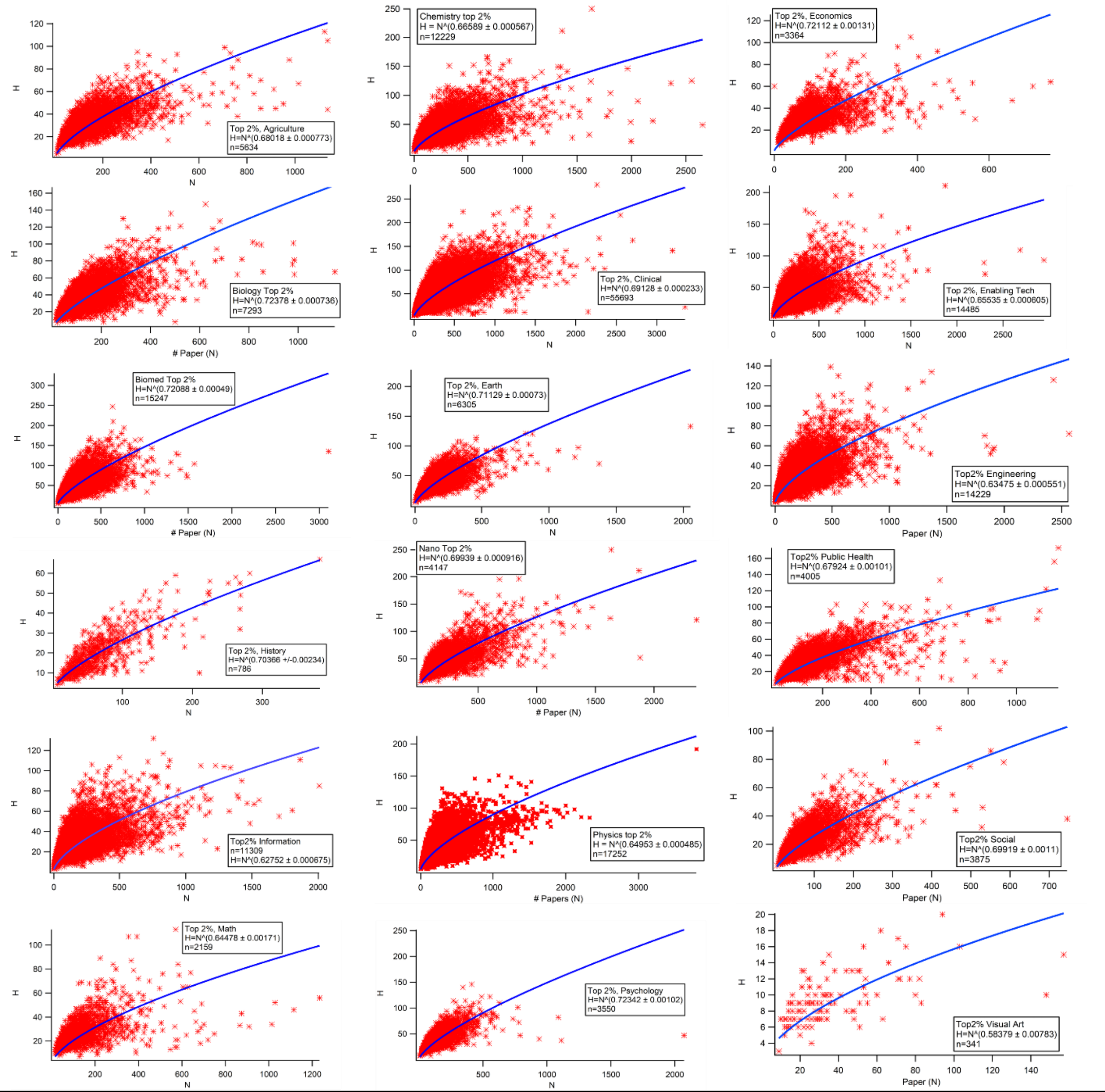


SI Fig. S3. Random angles of **Fig. 3h** in the main text.

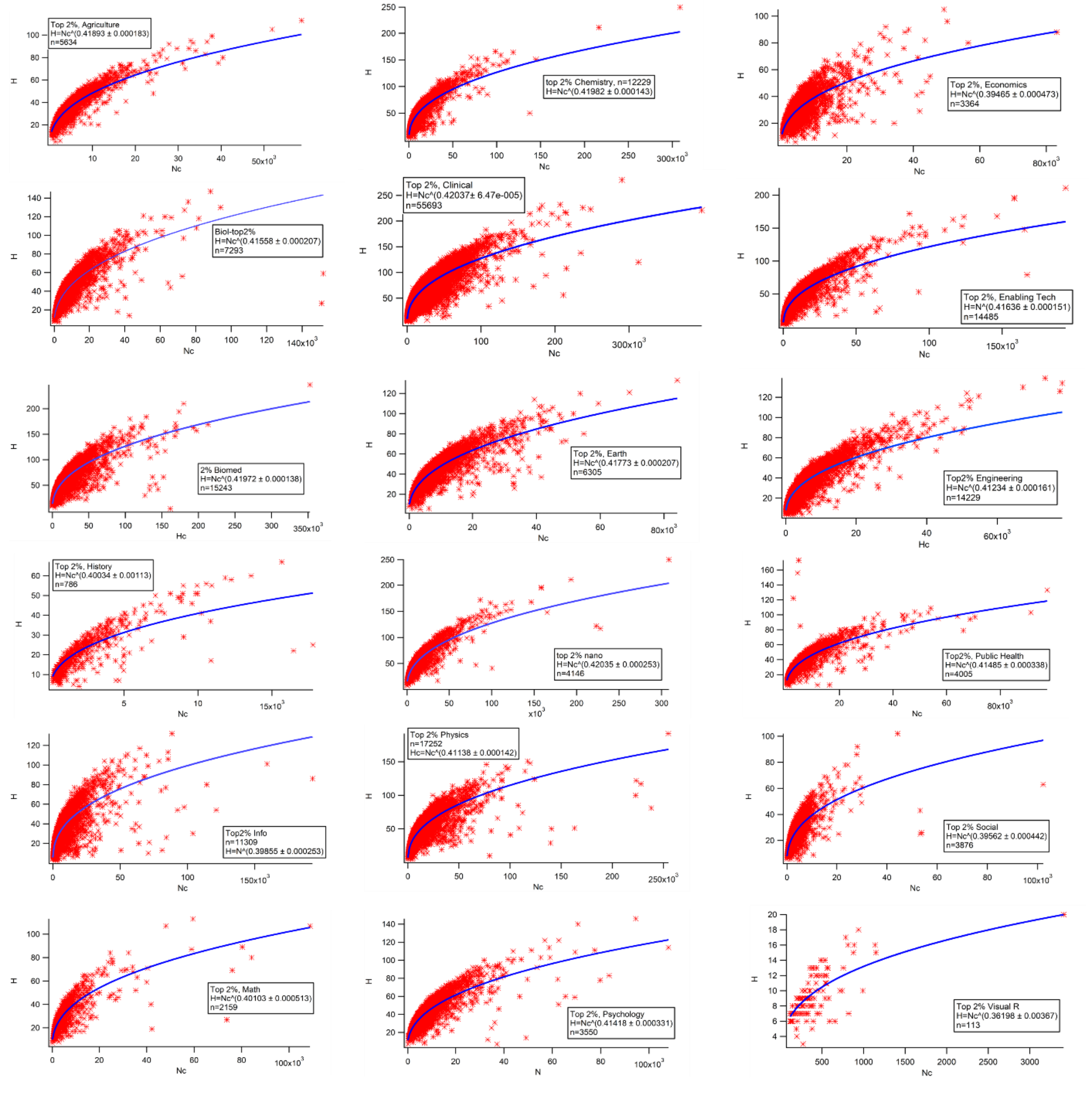


SI Fig. S4. Power law fitting between *H* vs *N* or *H* vs *Nc* in two halves (principal components) of the top 2% scholars. The scholars are split by (a) C-score, (b) H-index, and (c) *N* and *N*c respectively. Red color is top 1-2% and blue color is top 1% scholars.

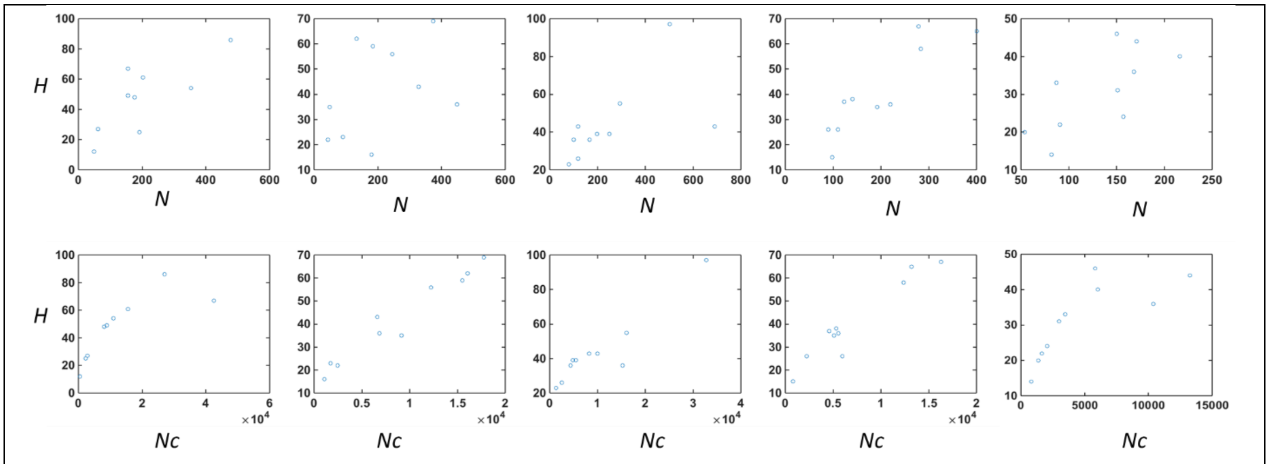
To test the power law relationships in smaller populations, we analyzed top 2% scientists in different disciplines. We found that in all 18 disciplines with number of researchers ranging from 113 (Visual Arts and Performance) to 55693 (Clinical Medicine), the power laws are valid for both *H* vs *N* and *H* vs *Nc* relationships (**SI Figs S5, S6**). The goodness of fitting of *H* vs *Nc* is always better than that of *H* vs *N*. As exponent ***n*** becomes larger, the quality publication as a measure of H-index becomes larger. Therefore, we rationalize that ***n*** can reflect the quality of all publications from the perspective of H-index in a scientific community represented by a discipline, an institute, an age group, etc. (see **Table 1** in the main text for ***n*** values of each discipline). In principle, ***n*** can also be used to indicate the quality of the publication from individual researchers. However, when we plot *H* vs *N* for 10 randomly selected scholars, most time the power law relationships are not observed or not obvious (**SI Fig. S7**). This result is understandable since the power law reflects an ensemble property of a population in which a heterogeneous distribution of different properties exists. For individual researchers, the accruement of *N* or and *H* is a function of time that varies drastically depending on the environment, the stage, as well as other factors experienced by each researcher.



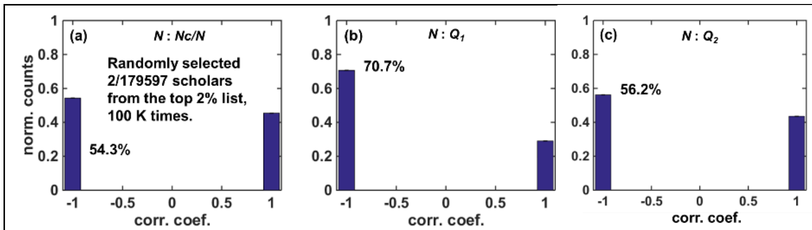
SI Fig. S5. The correlation between *H* and *N* for 18 subfields.



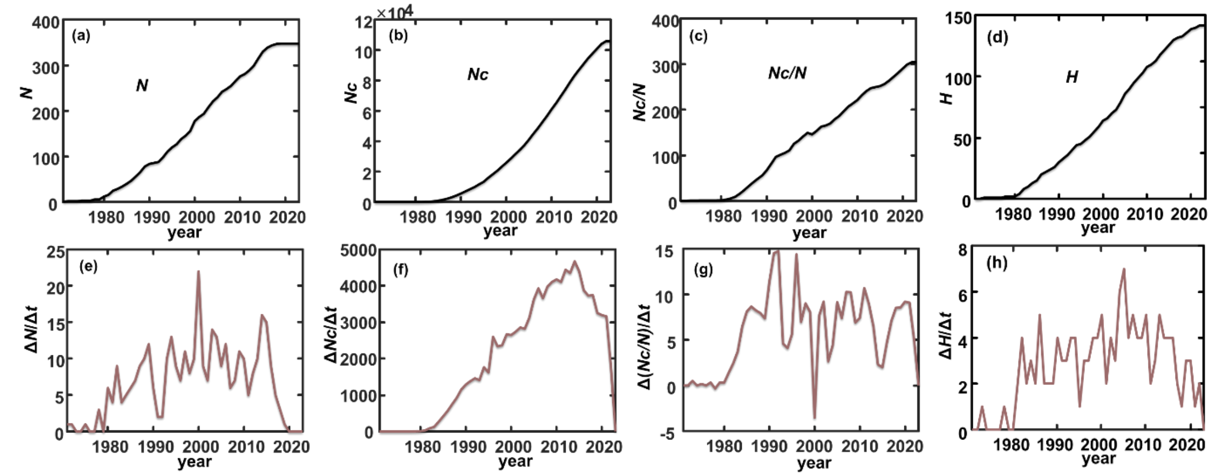
SI Fig. S6. The correlation between *H* and *N*c for 18 subfields.



SI Fig. S7. Five example trails of *H-N* and *H-Nc* plots of 10 randomly selected scholars from the top 2% researcher list.



SI Fig. S8. Histograms of correlation coefficients between N and (a) N_C/N, (b) Q_1 (c) Q_2 of randomly selected 2 scholars from the top 2% scientists averaged with 100,000 times sampling.



SI Fig. S9. Additional citation data analysis for Dr. Roger Y. Tsien. (a-d) *N*, *Nc*, *Nc/N*, and *H* over time; and (e-h) the slopes of these curves.