

# Analytical tools for data-driven management of NIH portfolios

February 4, 2014

George Santangelo

Director, Office of Portfolio Analysis (OPA)  
Division of Program Coordination, Planning,  
and Strategic Initiatives (DPCPSI)

# Mission of the Office of Portfolio Analysis

- **Coordinate portfolio analysis activities at NIH**
  - Conduct analyses for NIH senior leadership
  - Plan and host Poster sessions, Seminars, Workshops, and Symposia
  - Create opportunities for crosstalk within the NIH community
    - ❖ Portfolio Analysis Interest Group (PAIG) and blog (*The Analyst*)
- **Consult**
  - Assist NIH staff in the 27 Institutes and Centers (ICs) with analyses
    - ❖ Has resulted in collaborative development of tools, case studies, etc.
- **Train**
  - Both formal classes and ad hoc sessions
  - OPA web site: user manuals, FAQs, instructional videos (under construction)
- **Develop a science of portfolio analysis**
  - Build new tools / approaches and augment pre-existing ones
    - ❖ Primary focus is biomedical research
  - Build a community of experts: government, academia, private sector

# OPA Training

## Current classes

Portfolio Analysis: Introduction (PA101)

IN-SPIRE: Introduction

IN-SPIRE: Advanced Features

IN-SPIRE: Applied Workshop

## New classes to commence in Spring 2015

Network Analysis

Bibliometrics

U.S. Department of Health & Human Services | National Institutes of Health | Division of Program Coordination, Planning, and Strategic Initiatives (DPCPSI)

OFFICE OF PORTFOLIO ANALYSIS (OPA)  
TRAINING PORTAL VERSION 1.0

Printer Friendly | Text Size [A] [A] [A]

MANAGE CLASSES

MANAGE CLASS SCHEDULES

MANAGE CATEGORIES

ADD NEW CLASS

VIEW PAST CLASSES

VIEW TODAY'S CLASSES

VIEW UPCOMING CLASSES

VIEW REGISTRATIONS

CLASSES FEEDBACK

GENERAL FEEDBACK

View Past Classes

Welcome George Santangelo. You are logged in as Administrator

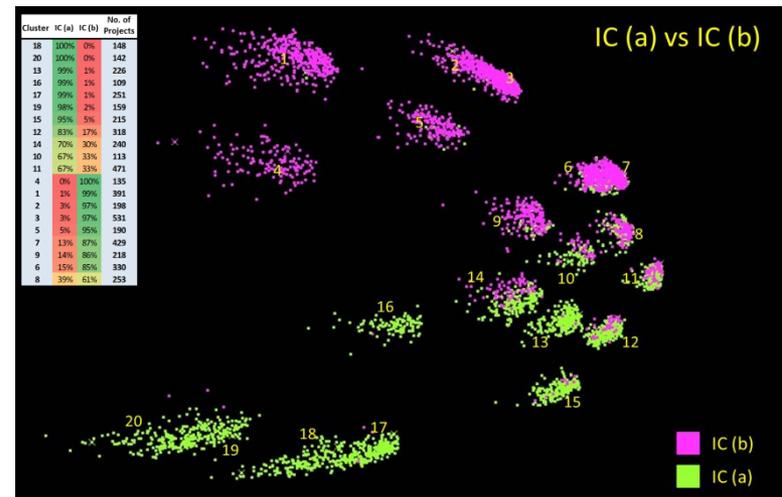
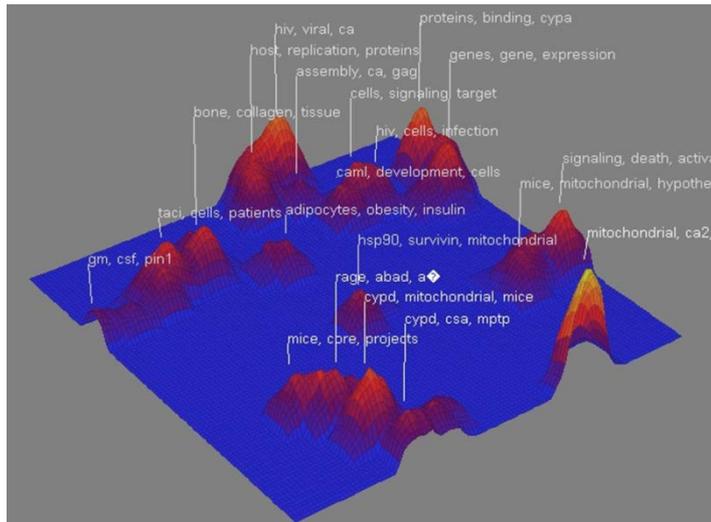
### View Past Classes

102 items in 5 pages

CLASS NAME	CLASS DATE	TIME	LOCATION	SEATS	REGISTRANTS	ACTION
Portfolio Analysis: Introduction	09/25/2014	2:00 PM - 3:30 PM	OPA Tools Lab	9	18	  
Portfolio Analysis: Introduction	10/07/2014	2:00 PM - 3:30 PM	OPA Tools Lab	9	17	  
IN-SPIRE: Advanced Features	10/09/2014	2:00 PM - 4:00 PM	OPA Tools Lab	9	10	  
Portfolio Analysis: Introduction	10/20/2014	2:00 PM - 3:30 PM	OPA Tools Lab	9	13	  
IN-SPIRE: Introduction	10/21/2014	2:00 PM - 4:30 PM	OPA Tools Lab	9	10	  
IN-SPIRE: Advanced Features	11/04/2014	2:00 PM - 4:00 PM	OPA Tools Lab	9	5	  
IN-SPIRE: Introduction	11/06/2014	2:00 PM - 4:30 PM	OPA Tools Lab	9	11	  
Portfolio Analysis: Introduction	11/18/2014	2:00 PM - 3:30 PM	OPA Tools Lab	9	16	  
IN-SPIRE: Advanced Features	11/20/2014	2:00 PM - 3:30 PM	OPA Tools Lab	9	6	  

# Current OPA training and support: content analysis

## Visualization/interactive discovery: Document clustering with IN-SPIRE



# NIH-customization of document clustering

## TBN (To be named later)

Grants

malaria

Search

▼ FY

2015	307
2014	351
2013	391
2012	439
2011	371
2010	424
2009	410
2008	275
2007	280
2006	246
2005	244
2004	218
2003	159

▼ Admin IC

AI	2884
TW	277
HD	198
RR	131
GM	98
NS	87
HL	75
CK	42
OD	38
CI	31
GH	30
MH	23
CA	18
MD	16
BX	13

▼ Activity Code

R01	1347
R21	435
U19	346
U01	197
R43	167
D43	159
R03	157
ZIA	133
P01	101
F32	88
Z01	86
RC1	77
P51	73
K23	62

4,130 Results

Download

- Title
- Abstract
- Specific Aim
- Summary Statement

ID	Grant Number	FY	Admin IC	Title
8814808	R01AI076353-05A1	2015	AI	Human Immunity to PfSEA-1, a Novel Vaccine Candidate for Falciparum Malaria
8895057	U19AI117887-01	2015	AI	Modeling of adaptive immune responses mediating durable protection against malaria
8895061	U19AI117887-01	2015	AI	Immunoprofiling adaptive immune response mediating durable protection against malaria
8892535	R21AI113592-01A1	2015	AI	Evaluation for Novel Strains of Malaria in Haiti
8887282	R44AI055229-11	2015	AI	Plasmodium Falciparum Whole Sporozoite Malaria Vaccine
8813445	K01TW009978-01	2015	TW	Surveillance and response triggers for malaria in Africa
8886652	R01AI110517-01A1	2015	AI	Malaria Prevention by a New Technology: Vectored Delivery of Antibody Genes
8891506	D43TW009074-05	2015	TW	Training of Cameroonian Scientists in Research on Malaria
8805027	R21AI115063-01	2015	AI	Function and mechanism of O-fucosylation of malaria parasite TSR-domain proteins
8882241	R21AI105619-02	2015	AI	Midgut Transcriptome and Proteome Analyses: Non-model Anopheline Malaria Vectors
8911438	R13AI118217-01	2015	AI	2015 Malaria Gordon Research Conference and Gordon Research Seminar
8836260	F32AI116038-01	2015	AI	Profiling Innate Immune Activation During Liver Stage Malaria Infection
8837779	R01AI114292-01A1	2015	AI	Multivalent chimeric subunit malaria vaccines
8909575	F32AI114197-01A1	2015	AI	Mosquito-associated fungi enhance vector competence for malaria
8905351	F32AI114116-01A1	2015	AI	Characterization of a novel falciparum malaria vaccine candidate: PfSEA-1
8817134	R01AI075045-06	2015	AI	Resistance of Malaria Parasites to Artemisinin-Based Therapies
8833531	F32AI115965-01	2015	AI	Elucidating the role of heat shock protein 101 in malaria effector protein export
8895246	U19AI089702-06	2015	AI	Epidemiology of Malaria Transmission in Low to Moderate Settings of Latin America
8898960	D71TW010070-01	2015	TW	Genital Schistosomiasis and Co-infection with Malaria and or HIV (SCHIVMA)
8906017	R43AI118185-01	2015	AI	Development of a whole parasite Plasmodium falciparum sexual and mosquito stages Vaccine to Interrupt Malaria Transmission



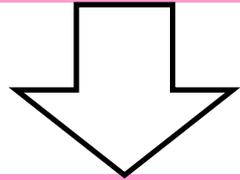
# Developing a science of portfolio analysis

- Use existing data-driven approaches to characterize research investments and the resulting impact
- Develop and deliver effective approaches and methodologies
  - Tools in development:

Functionality	Tool
Content analysis	<i>IN-SPIRE + TBN (To be named later)</i>
Efficient disambiguation	<i>iClean</i>
Effective bibliometrics	<i>iCite</i>
Map translational science	<i>iTrans</i>
Track patent, licensing, start-up activity	<i>iTech</i>

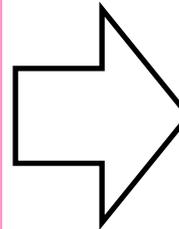
# Data integration requires an effective disambiguation tool

Nicholson & Ioannidis gave us a list of 1172 author names that we needed to disambiguate

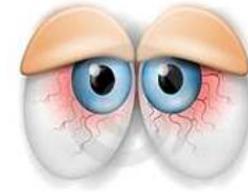


Top 20 difficult author names to disambiguate

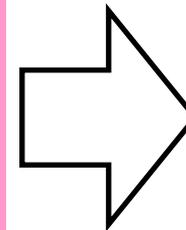
Allen, R.P.  
Anderson, K.C.  
Berger, C.  
Cannon, C.P.  
Cohen, J.D.  
Collins, P.G.  
Johnson, D.H.  
Kahn, R.  
Kim, R.Y.  
Lee, T.I.  
Liu, E.A.  
Miller, K.  
Park, J.  
Russell, M.E.  
Scott, E.W.  
Singh, A.K.  
Stern, M.  
Stone, G.W.  
Ward, E.  
Zhou, G.



~22 PIs in IMPAC II per name (554 total)



Manual inspection



60.4% of the authors have never applied for an NIH award

Refine Results

Search within results for...



Web of Science Categories

GENETICS HEREDITY (325)

Refine

Document Types

ARTICLE (325)

Refine

Research Areas

Authors

Group Authors

Editors

Source Titles

Book Series Titles

Conference Titles

Publication Years

Organizations-Enhanced

Funding Agencies

Languages

Countries/Territories

Open Access

For advanced refine options, use

Analyze Results

Funding Agencies

Refine

Exclude

Cancel

Sort these by:

Record Count

The first 100 Funding Agencies (by record count) are shown. For advanced refine options, use [Analyze results](#).

NATIONAL INSTITUTES OF HEALTH (83)

NATIONAL SCIENCE FOUNDATION (61)

NATIONAL INSTITUTES OF HEALTH NIH (29)

NIH (10)

SWISS NATIONAL SCIENCE FOUNDATION (7)

CANADIAN INSTITUTES OF HEALTH RESEARCH (6)

HOWARD HUGHES MEDICAL INSTITUTE (5)

NATIONAL SCIENCE FOUNDATION NSF (5)

SWEDISH RESEARCH COUNCIL (5)

WELLCOME TRUST (5)

AMERICAN CANCER SOCIETY (4)

NATURAL SCIENCES AND ENGINEERING RESEARCH COUNCIL OF CANADA (4)

ACADEMY OF FINLAND (3)

BIOTECHNOLOGY AND BIOLOGICAL SCIENCES RESEARCH COUNCIL (3)

BIOTECHNOLOGY AND BIOLOGICAL SCIENCES RESEARCH COUNCIL BBSRC (3)

BURROUGHS WELLCOME FUND (3)

CANADA RESEARCH CHAIR (3)

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (3)

DEUTSCHE FORSCHUNGSGEMEINSCHAFT (3)

ENGINEERING AND PHYSICAL SCIENCES RESEARCH COUNCIL (3)

EUROPEAN UNION (3)

ISRAEL SCIENCE FOUNDATION (3)

JAPAN SOCIETY FOR THE PROMOTION OF SCIENCE (3)

MINISTRY OF EDUCATION CULTURE SPORTS SCIENCE AND TECHNOLOGY OF JAPAN (3)

NATIONAL CENTER FOR RESEARCH RESOURCES (3)

NATIONAL INSTITUTE OF GENERAL MEDICAL SCIENCES (3)

NATIONAL INSTITUTE OF NEUROLOGICAL DISORDERS AND STROKE (3)

NATIONAL INSTITUTES OF HEALTH NATIONAL CENTER FOR RESEARCH RESOURCES (3)

NATIONAL NATURAL SCIENCE FOUNDATION OF CHINA (3)

ALFRED P SLOAN FOUNDATION (2)

ALFRED P SLOAN RESEARCH FELLOWSHIP (2)

CANADIAN FOUNDATION FOR INNOVATION (2)

CANADIAN INSTITUTES FOR HEALTH RESEARCH (2)

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS (2)

EUROPEAN MOLECULAR BIOLOGY ORGANIZATION (2)

EUROPEAN MOLECULAR BIOLOGY ORGANIZATION EMBO (2)

EUROPEAN REGIONAL FUND (2)

EUROPEAN RESEARCH COUNCIL (2)

FRENCH MINISTRY OF RESEARCH (2)

HUMAN FRONTIER SCIENCE PROGRAM (2)

JAMES S MCDONNELL FOUNDATION (2)

JAPAN SCIENCE AND TECHNOLOGY AGENCY (2)

MEDICAL RESEARCH COUNCIL (2)

NATIONAL CENTER FOR RESEARCH RESOURCES OF THE NATIONAL INSTITUTES OF HEALTH (2)

NATIONAL INSTITUTES OF HEALTH NIH NATIONAL CENTER FOR RESEARCH RESOURCES (2)

NATURAL SCIENCES AND ENGINEERING COUNCIL OF CANADA (2)

NETHERLANDS ORGANIZATION FOR SCIENTIFIC RESEARCH (2)

NSF (2)

ONTARIO GRADUATE SCHOLARSHIP (2)

PACKARD FELLOWSHIP FOR SCIENCE AND ENGINEERING (2)

ROYAL SOCIETY (2)

SPANISH MINISTRY OF SCIENCE AND EDUCATION (2)

U S DEPARTMENT OF AGRICULTURE NATIONAL RESEARCH INITIATIVE (2)

UNIVERSITY OF HELSINKI (2)

UNIVERSITY OF TEXAS AT ARLINGTON (2)

VETENSKAPSRADET (2)

NATIONAL INSTITUTES OF HEALTH NATIONAL HUMAN GENOME RESEARCH INSTITUTE (1)

NATIONAL INSTITUTES OF HEALTH NATIONAL INSTITUTE OF ALLERGY AND INFECTIOUS DISEASES (1)

NATIONAL INSTITUTES OF HEALTH NATIONAL INSTITUTE OF CHILD HEALTH AND HUMAN DEVELOPMENT (1)

NATIONAL INSTITUTES OF HEALTH NIH NATIONAL CENTER FOR RESEARCH RESOURCES NCRR (1)

NATIONAL INSTITUTES OF HEALTH NIH NATIONAL CENTER FOR RESEARCH RESOURCES NCRR (1)

NATIONAL INSTITUTES OF HEALTH NIH NATIONAL INSTITUTE OF CHILD HEALTH AND HUMAN DEVELOPMENT NICHD (1)

NATIONAL INSTITUTES OF HEALTH NIH NATIONAL INSTITUTE OF DIABETES AND DIGESTIVE AND KIDNEY DISEASES (1)

NATIONAL INSTITUTES OF HEALTH NIH NATIONAL RESEARCH SERVICE (1)

NATIONAL INSTITUTES OF HEALTH UNIVERSITY OF CALIFORNIA SAN DIEGO (1)

NATIONAL INSTUTES OF HEALTH (1)

NATIONAL KEY BASIC RESEARCH PROJECT (1)

NATIONAL LIBRARY OF MEDICINE (1)

NIH NATIONAL CENTER FOR RESEARCH RESOURCES (1)

NIH NATIONAL HUMAN GENOME RESEARCH INSTITUTE (1)

NIH NATIONAL INSTITUTE OF DIABETES AND DIGESTIVE AND KIDNEY DISEASES (1)

NIH ROADMAP INITIATIVE ON NANOMEDICINE (1)

NIH RUTH L KIRSCHSTEIN NATIONAL RESEARCH SERVICE (1)

NORTHWESTERN UNIVERSITY ALUMNI ASSOCIATION (1)

NORWEGIAN RESEARCH COUNCIL (1)

NOVARTIS RESEARCH FOUNDATION (1)

NSERC (1)

NSF PLANT GENOME (1)

OFFICE OF SCIENCE U S DEPARTMENT OF ENERGY (1)

OKLAHOMA CENTER FOR THE ADVANCEMENT OF SCIENCE AND TECHNOLOGY OCAST (1)

ONTARIO GOVERNMENT (1)

PACKARD FOUNDATION (1)

PEROT FAMILY FOUNDATION (1)

PEW LATIN AMERICA FELLOWSHIP (1)

PIGBIODIV2 (1)

PLANT SCIENCES INSTITUTE (1)

POLYCYSTIC KIDNEY DISEASE FOUNDATION (1)

PRESTO PRECURSORY RESEARCH FOR EMBRYONIC SCIENCE AND TECHNOLOGY (1)

PROGRAM OF WUHAN SUBJECT CHIEF SCIENTIST (1)

PROGRAMME D ACTIONS UNIVERSITAIRES INTEGRES LUSO FRANCAISES (1)

PUBLIC HEALTH SERVICE GRANT (1)

RAMON ARECES FOUNDATION (1)

RAMON Y CAJAL PROGRAM (1)

REGION BRETAGNE (1)

REGIONAL GOVERNMENT (1)

RESEARCH COUNCILS UNITED KINGDOM (1)

RESEARCH FOUNDATION FLANDERS (1)

RESEARCH SCHOOL IN FOREST GENETICS AND BREEDING (1)

RESEARCH SCHOOL OF TREE BREEDING AND FOREST GENETICS (1)

RETINA RESEARCH FOUNDATION (1)

# *iClean*: A web-based tool for automated disambiguation

- Multimodal information sources
  - Author name
  - Publication title
  - Publication keywords
  - Automatically extracted terms (title and abstract)
  - Abstract similarity
  - Author institution(s)
  - Co-author network
  - Citation network
- Uses a machine learning approach
- Will prompt users to manually curate hard cases
  - This information can feedback to improve performance
- Permits accurate, less time-consuming network analysis, which is increasingly popular among NIH program staff
- Can be used to link any grants database to the literature with minimal ambiguity

# Developing a science of portfolio analysis

- Use existing data-driven approaches to characterize research investments and the resulting impact
- Develop and deliver effective approaches and methodologies
  - Tools in development:

Functionality	Tool
Content analysis	IN-SPIRE + TBN ( <i>To be named later</i> )
Efficient disambiguation	<i>iClean</i>
Effective bibliometrics	<i>iCite</i>
Map translational science	<i>iTrans</i>
Track patent, licensing, start-up activity	<i>iTech</i>

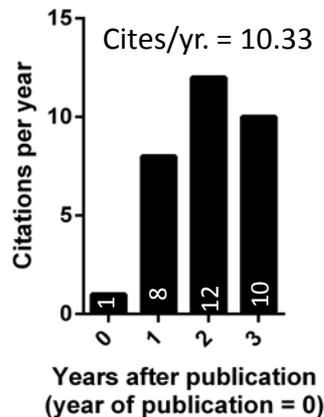
## Web-based use of the Relative Citation Rate (RCR) metric

- **Limitations of commonly used bibliometrics in measuring/comparing the value of a publication or group of publications:**
  - ✓ *Publication Counts*: field-dependent, use-independent
  - ✓ *Impact Factor*: journal-level metric
  - ✓ *Citation Rates*: field- and journal-dependent
  - ✓ *h-index*: field-dependent, time-dependent
- **RCR**
  - ✓ Need: An article-level metric that is independent of field, journal, and time
  - ✓ Assumption: Citation of a publication reveals value to or influence on the citer
  - ✓ RCR normalizes citations to the publications co-citation network

# Calculating Relative Citation Rate

$$\text{RCR} = \frac{\text{Article Citation Rate}}{\text{Expected Citation Rate}}$$

Article Citation Rate  
(denominator excludes year of publication)

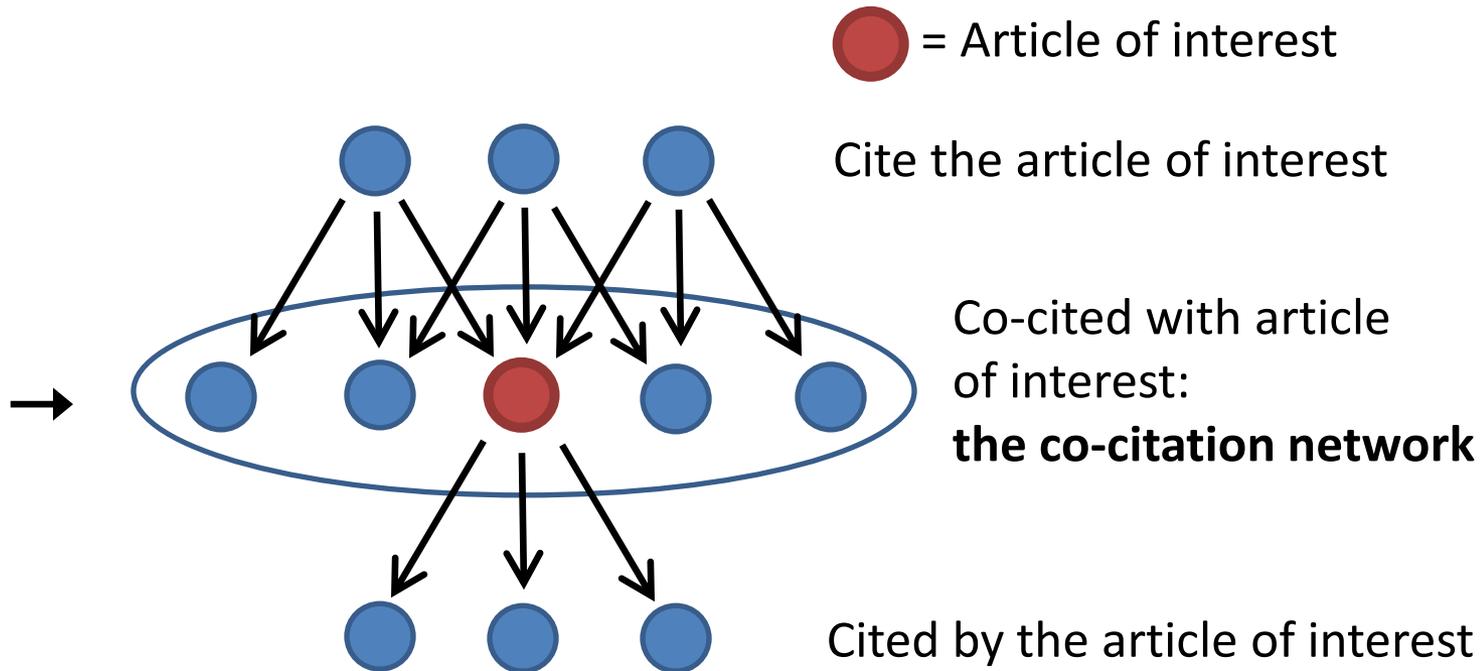


- Article-level metric
- Number of cites per year
- RCR changes over time with the accrual of new citations
- Scalable to large portfolios containing tens of thousands of articles

# Calculating Relative Citation Rate

## Expected Citation Rate Part 1: Calculate the Field Citation Rate

**Field Citation Rate**  
= Average the *journal citation rates* for these co-cited articles (includes article of interest)

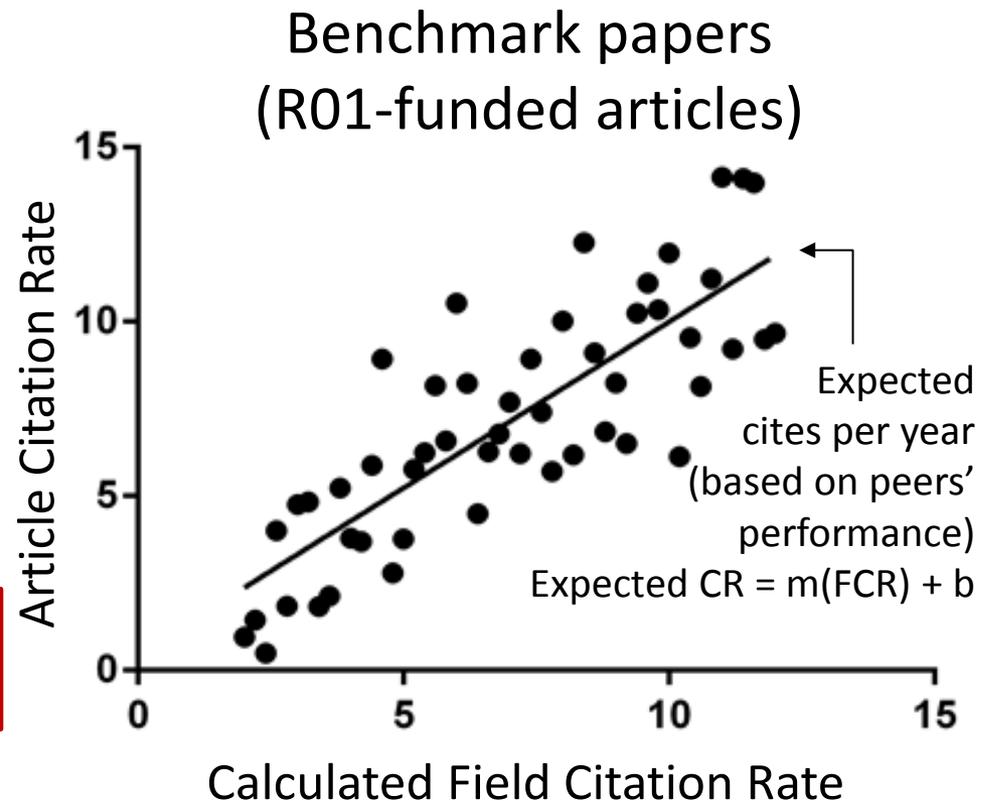


# Calculating Relative Citation Rate

## Expected Citation Rate Part 2: Benchmark the Field Citation Rate

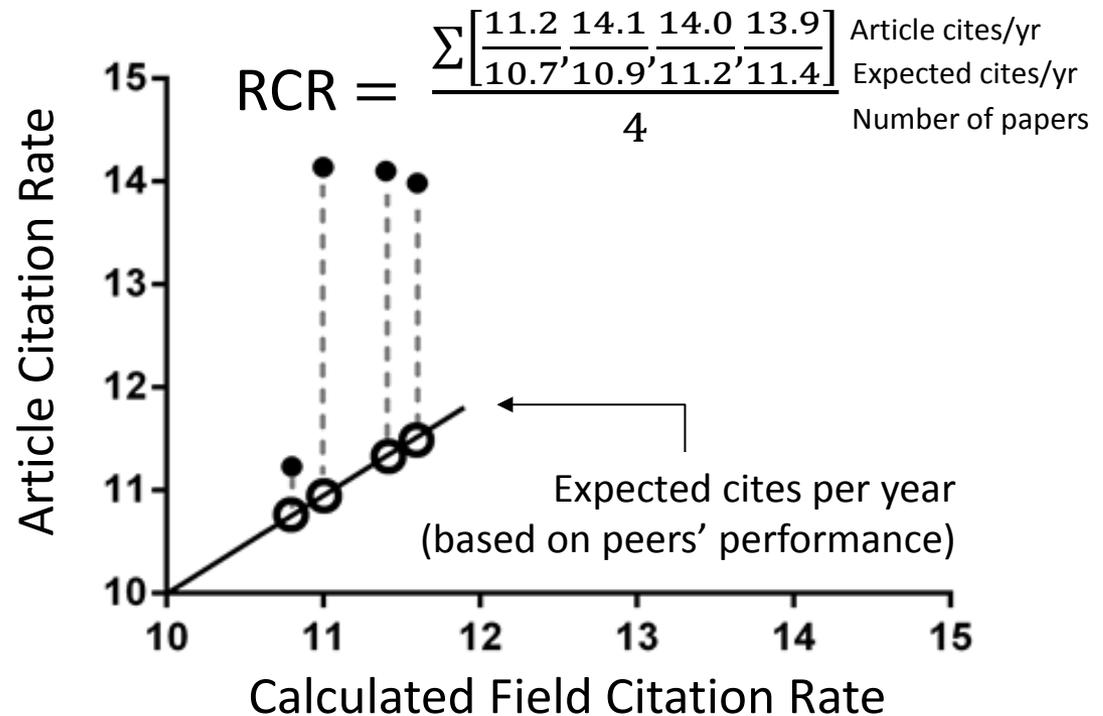
- Use this linear regression equation to transform the Field Citation Rate
- This benchmarks expectations to R01-funded papers for any Field (avg. = 1.0)

$$RCR = \frac{\text{Article Citation Rate}}{\text{Expected Citation Rate}}$$



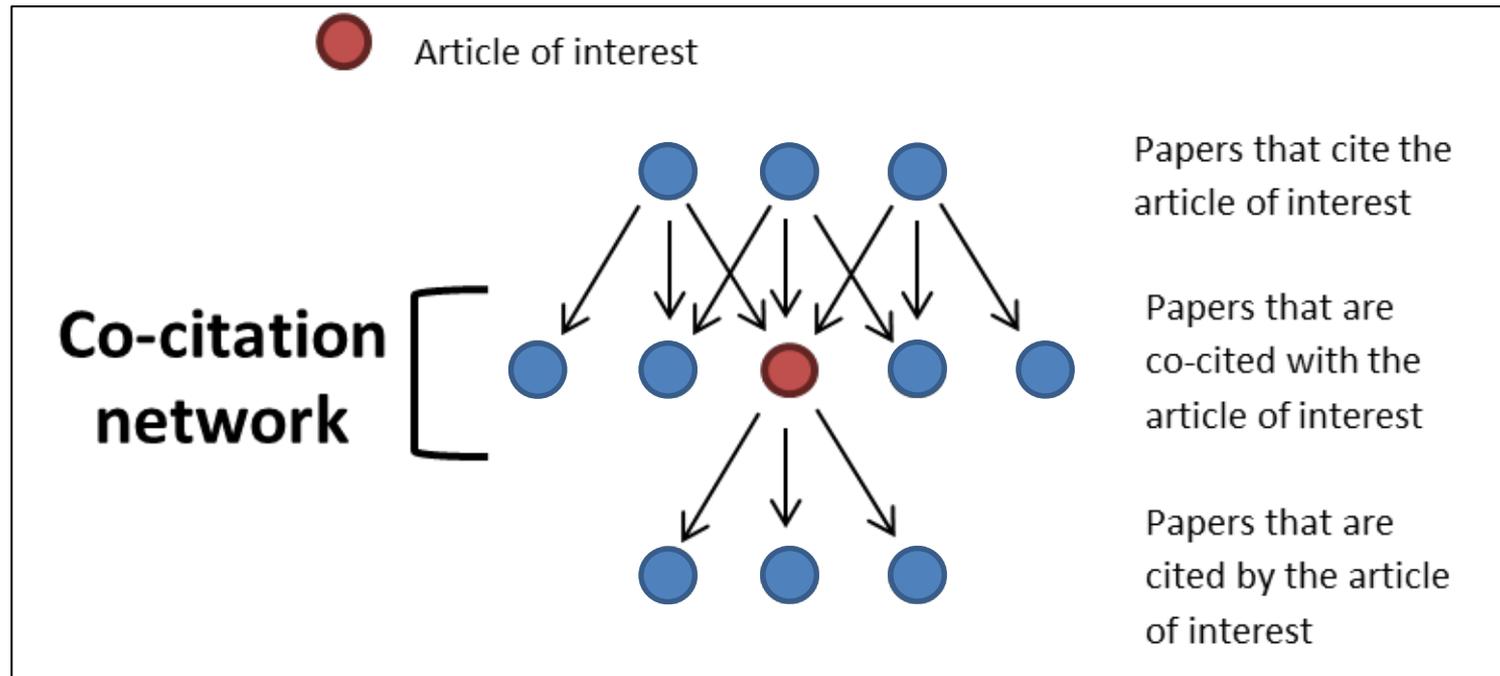
# RCR extends easily to a portfolio of papers

- Easily extended to a portfolio by averaging the RCRs of the individual articles



# RCR

How is the paper of interest cited relative to other papers in its co-citation network?



- 0 = never cited
- 1 = average
- 2 = twice the average
- >20 = exceptionally highly cited

Thomson Reuters Science Citation Index Expanded, 2002-2012

# RCR denominator: sample co-citation networks

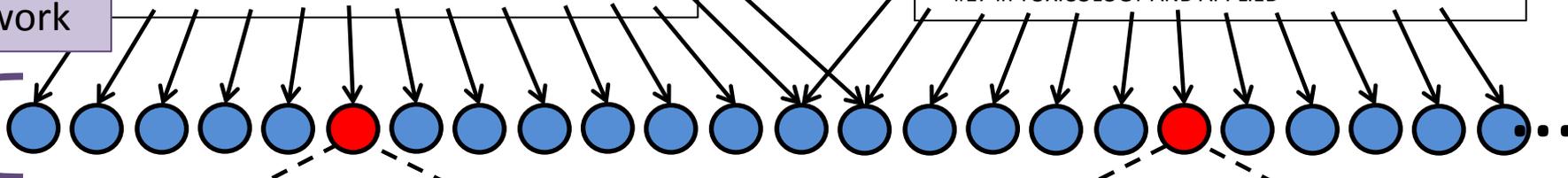
2012 publications citing  
a Nature 2011 paper on Hepatitis C:

- #1 in CELL HOST & MICROBE
- #2 in CELL REPORTS
- #3 in COLD SPRING HARBOR PERSPECTIVES IN BIOLOGY
- #4 in CURRENT OPINION IN STRUCTURAL BIOLOGY
- #5 in CURRENT OPINION IN VIROLOGY
- #6 in CURRENT OPINION IN VIROLOGY
- #7 in CURRENT OPINION IN VIROLOGY
- #8 in EMBO JOURNAL
- #9 in IMMUNITY
- #10 in JOURNAL OF APPLIED CRYSTALLOGRAPHY
- #11 in JOURNAL OF BIOLOGICAL CHEMISTRY
- #12 in JOURNAL OF VIROLOGY
- #13 in MOLECULAR THERAPY
- #14 in NATURE
- #15 in PLOS ONE
- #16 in PLOS PATHOGENS
- #17 in PLOS PATHOGENS
- #18 in PROCEEDINGS NATL ACAD OF SCI USA
- #19 in PROCEEDINGS NATL ACAD OF SCI USA
- #20 in STRUCTURE

2012 publications citing  
a PLoS ONE 2011 paper on Hepatitis C:

- #1 in AMERICAN JOURNAL OF PATHOLOGY
- #2 in CNS NEUROSCIENCE & THERAPEUTICS
- #3 in EXPERIMENTAL AND MOLECULAR PATHOLOGY
- #4 in HEPATOLOGY INTERNATIONAL
- #5 in JOURNAL OF TRANSLATIONAL MEDICINE
- #6 in JOURNAL OF TRAUMA AND ACUTE CARE
- #7 in SURGERY
- #8 in JOURNAL OF VIROLOGY
- #9 in LEUKEMIA
- #10 in PLOS ONE
- #11 in PLOS ONE
- #12 in PLOS ONE
- #13 in PLOS ONE
- #14 in PROCEEDINGS NATL ACAD OF SCI USA
- #15 in SHOCK
- #16 in SHOCK
- #17 in TOXICOLOGY AND APPLIED

Co-citation  
network



**RCR = 3.0**

**Nature 2011  
Hepatitis C paper**

**PLoS ONE 2011  
Hepatitis C paper**

**RCR = 4.2**



# RCR validation study using subject matter experts

## Is there a correlation between RCR and expert assessment of value/quality/impact?

- 684 papers published in 2009 supported by R01s and representing a range of RCRs
- Papers assigned based on content of the reviewer's published work
- 537 IRP Investigators were recruited with approval of their SDs
- 3-5 PIs received the same set of 5 publications
  - ✓ Pubs and responses via a secure intranet site
- 6 questions asked with a 5 point response scale
- Response as of Oct 14, 2014
  - ✓ 44.3% Investigators responded
  - ✓ 1028 responses to 562 papers, 290 with  $\geq 2$  responses/paper

# Relative Citation Rate – Review Criteria

- **IMPORTANCE**

Rate whether the question being addressed is important to answer.

(1 = Not Important, 2 = Slightly Important, 3 = Important, 4 = Highly Important, 5 = Extremely Important)

- **METHODS**

Rate whether you agree that the methods are appropriate and the scope of the experiments adequate.

(1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree)

- **ROBUSTNESS**

Rate how robust the study is based on the strength of the evidence presented.

(1 = Not Robust, 2 = Slightly Robust, 3 = Moderately Robust, 4 = Highly Robust, 5 = Extremely Robust)

- **HUMAN HEALTH RELEVANCE**

Rate the likelihood that the results could ultimately have a substantial positive impact on human health outcomes.

(1 = Very unlikely, 2 = Unlikely, 3 = Foreseeable but uncertain, 4 = Probable, 5 = Almost Certainly)

- **LIKELY IMPACT**

Rate the impact that the research is likely to have or has already had.

(1 = Minimal Impact, 2 = Some Impact, 3 = Moderate Impact, 4 = High Impact, 5 = Extremely High Impact)

- **OVERALL EVALUATION**

Provide your overall evaluation of the value and impact of this publication.

(1 = minimal or no value, 2 = Moderate value, 3 = Average value, 4 = High value, 5= Extremely high value)

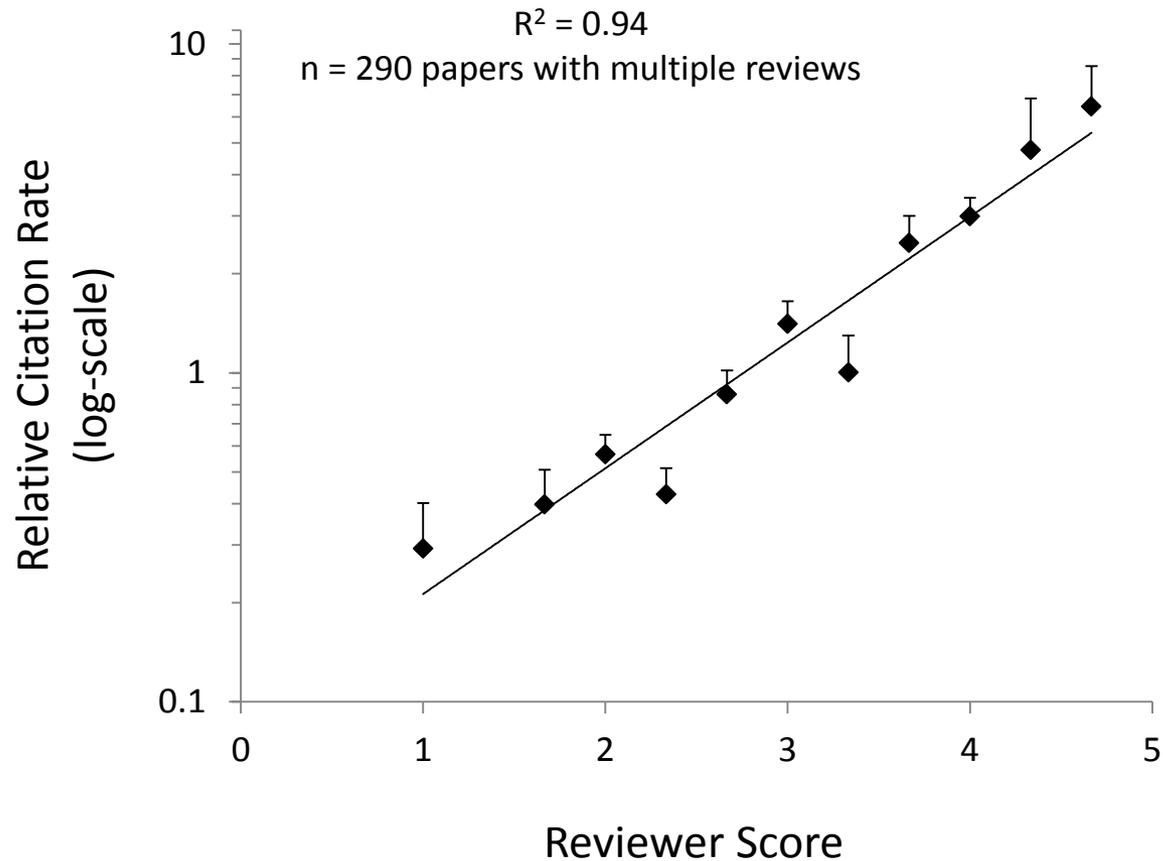
# Fields of science represented in the RCR validation study

Cell Biology  
Genetics and Genomics  
Neuroscience  
Chromosome Biology  
Developmental Biology  
Immunology  
Molecular Biology and Biochemistry  
Cancer Biology  
Stem Cell Research  
Molecular Pharmacology  
Systems Biology

Epidemiology  
Clinical Research  
Virology  
Computational Biology  
Biomedical Engineering & Biophysics  
Chemical Biology  
Microbiology and Infectious Diseases  
Structural Biology  
Health Disparities  
Social and Behavioral Science

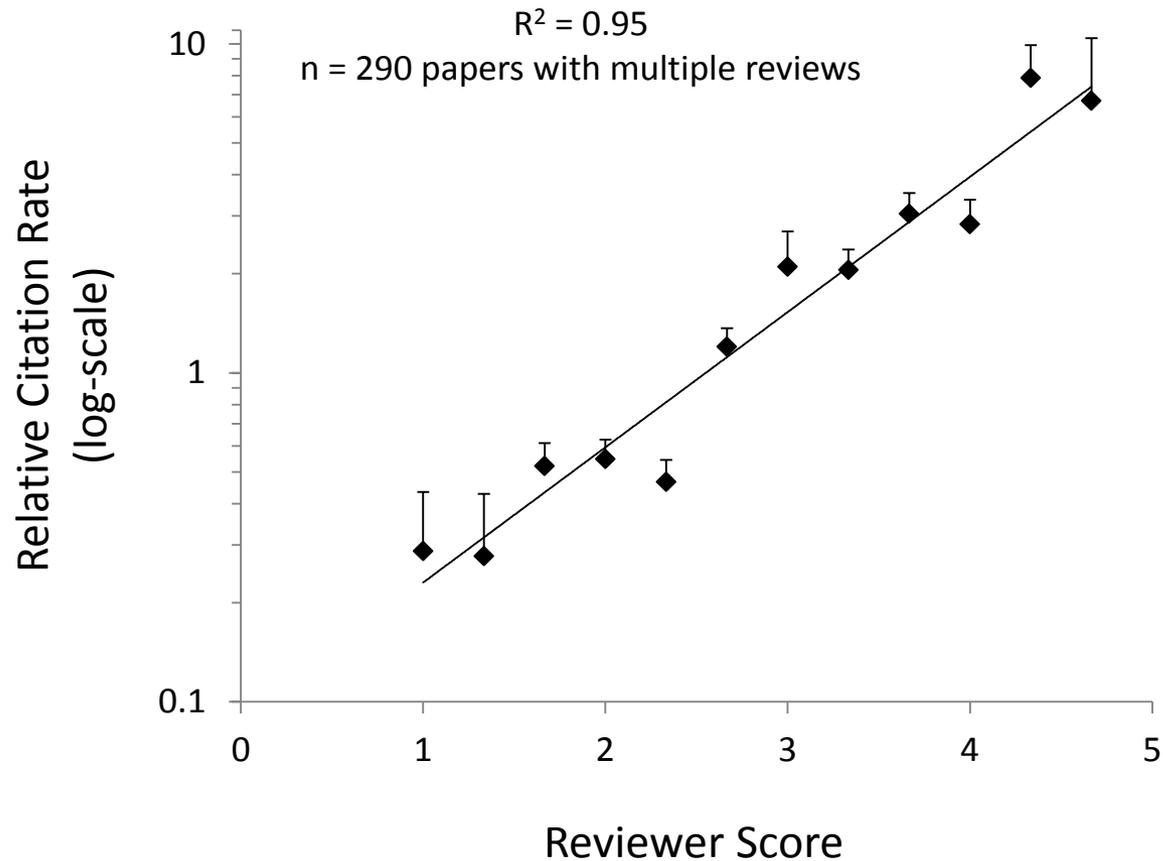
# Relative Citation Rate (RCR) validation study

## OVERALL EVALUATION



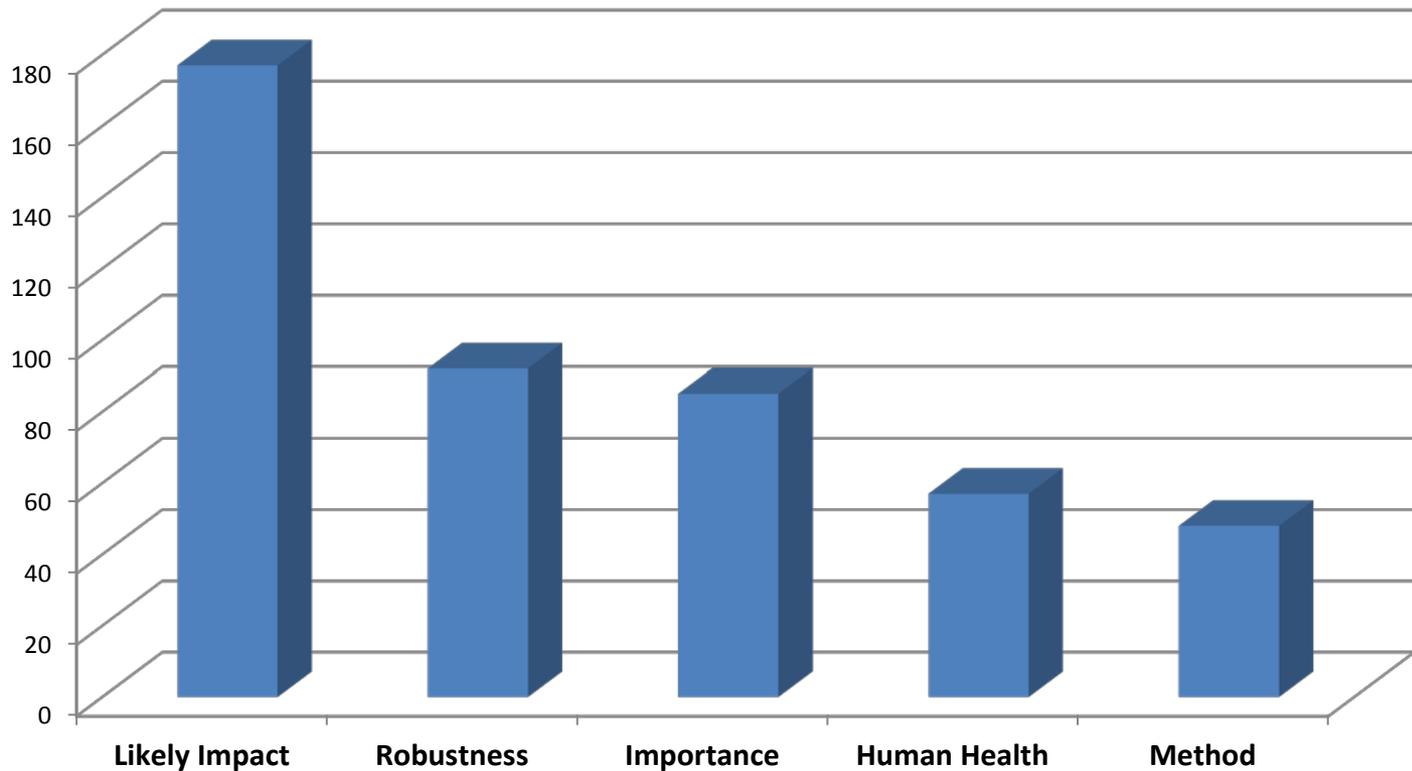
# Relative Citation Rate (RCR) validation study

## LIKELY IMPACT



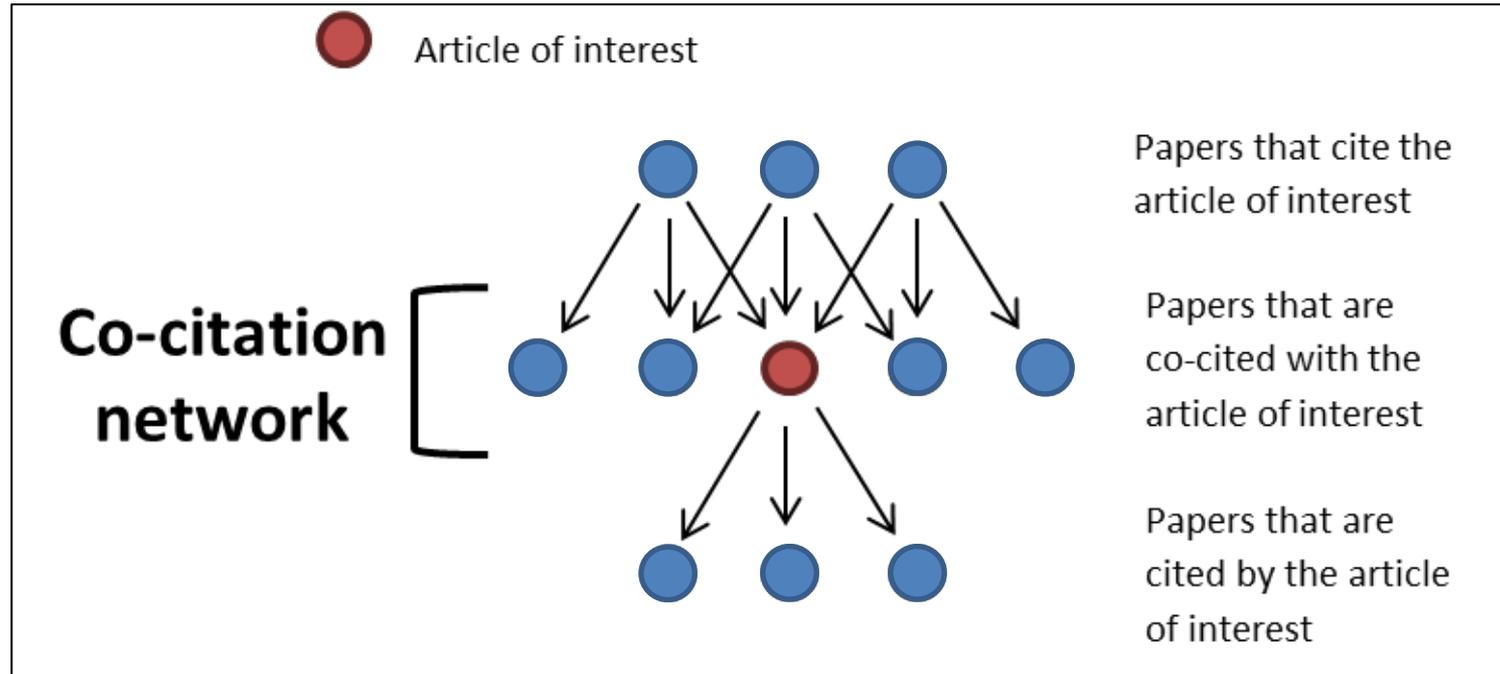
# How do reviewers weigh these factors when deciding on a paper's overall value?

Relative Importance on Article Value  
(Random Forest)



# RCR

How is the paper of interest cited relative to other papers in its co-citation network?



- 0 = never cited
- 1 = average
- 2 = twice the average
- >20 = exceptionally highly cited

Thomson Reuters Science Citation Index Expanded, 2002-2012

# Accurately gauging expected citations (the denominator in normalized bibliometrics)

## Thomson Reuters

### Journal categories used to calculate denominator

- R01-funded pubs with Expected CPY < 2.0 in 2009
- TR Ratio ranged from 0 to 23.4
- Average TR Ratio = **1.6**

## RCR

### Co-citation networks used to calculate denominator

- R01-funded pubs with Expected CPY < 2.0 in 2009
- RCR ranged from 0 to 3.7
- Average RCR = **0.24**

# Accurately gauging expected citations (the denominator in normalized bibliometrics)

Social networkers' attitudes toward direct-to-consumer personal genome testing  
*The American Journal of Bioethics* 9 : 3-10 (2009)

**26.3 cites per year**

## Thomson Reuters

Journal categories used to calculate denominator

- Expected CPY: **1.36**
- TR ratio: **19.4**

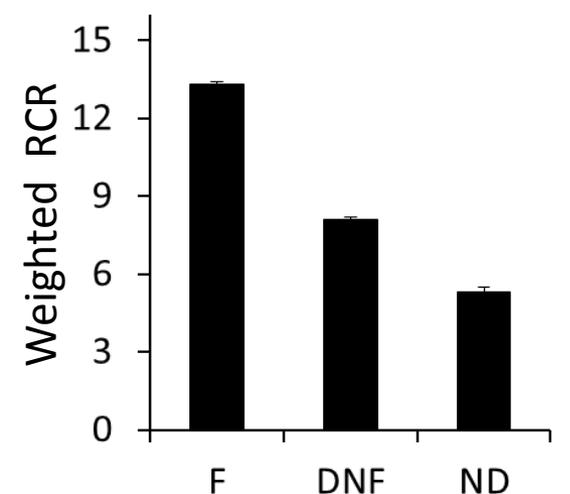
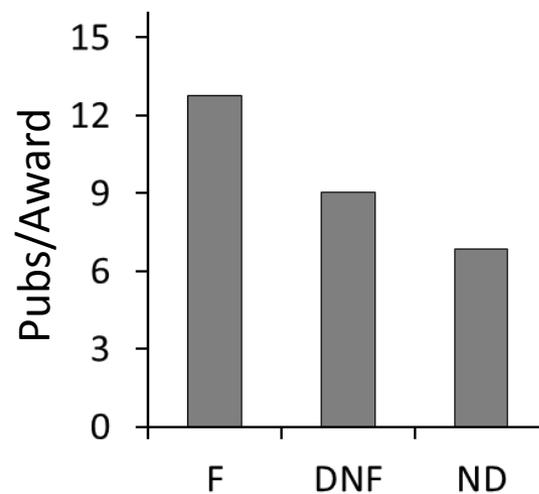
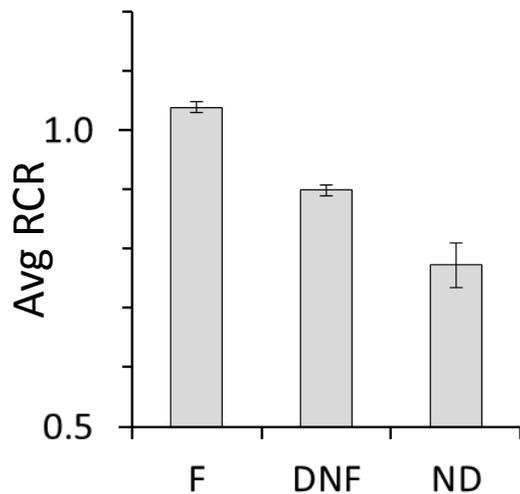
## RCR

Co-citation networks used to calculate denominator

- Expected CPY: **10.1**
- RCR: **2.6**

# Bibliometric analysis of renewed vs. declined T2 applications (FY08-FY12)

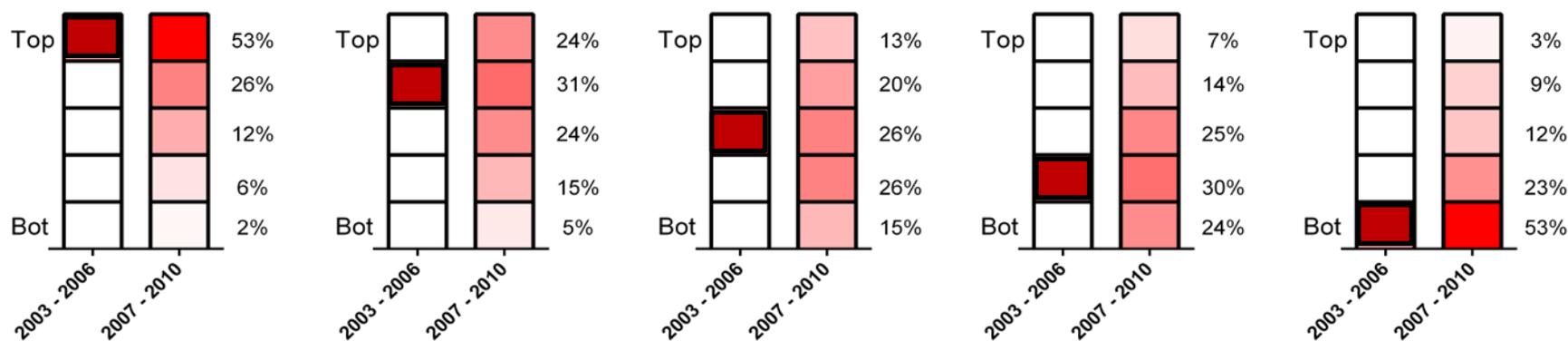
T2 Outcome	# Awards	# Discussed*	# Funded	# Pubs	Pubs/Award	RCR		
						Avg	SEM	Weighted**
Funded	10816	10816	10816	138212	12.8	1.04	0.01	13.3
Discussed*	6973	6973	0	62836	9.0	0.90	0.01	8.1
Not discussed	7015	0	0	47896	6.8	0.77	0.04	5.3



\*DNF = Discussed but not funded

\*\*Weighted RCR = Avg RCR \* Pubs/Award

# Scientific mobility of PIs continuously funded through two consecutive 4-year periods



Continuously funded (R01) NIH Investigators, ranked by RCR, 3089 PIs

# Uses of RCR to compare impact

- Validated use:
  - ✓ compare individual publications within a network
- Additional uses:
  - ✓ compare output of individual investigators
  - ✓ compare output of cohorts (portfolios, programs, mechanisms...)
  - ✓ flagging low impact for inspection
  - ✓ to follow trends
- Possible misuses of RCR or any bibliometric assessment:
  - ✓ determining importance of the endeavor
  - ✓ predicting long-term impact of the work
  - ✓ evaluating effectiveness of “downstream” application (e.g. patentable)

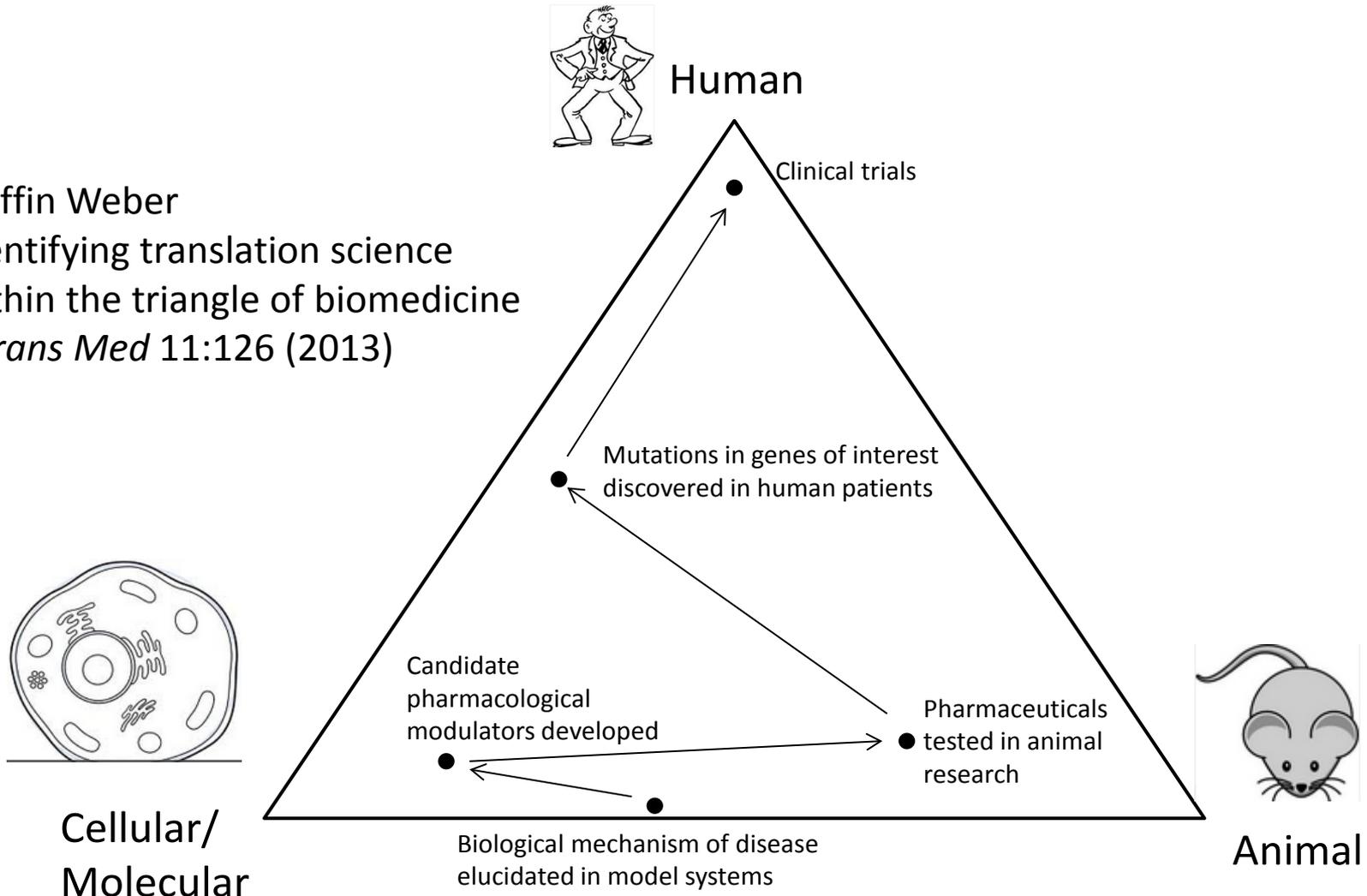
# Developing a science of portfolio analysis

- Use existing data-driven approaches to characterize research investments and the resulting impact
- Develop and deliver effective approaches and methodologies
  - Tools in development:

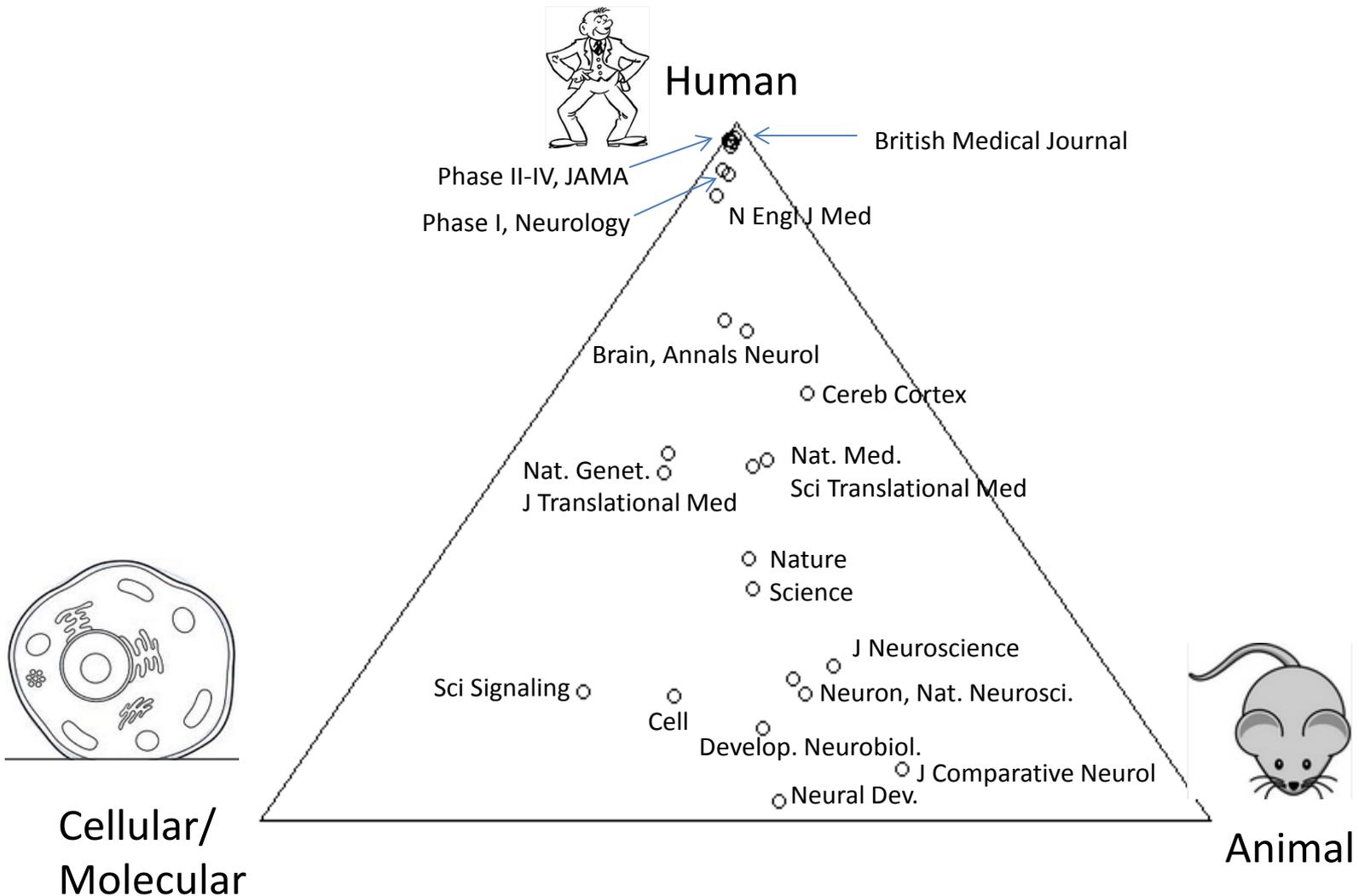
Functionality	Tool
Content analysis	IN-SPIRE + TBN ( <i>To be named later</i> )
Efficient disambiguation	<i>iClean</i>
Effective bibliometrics	<i>iCite</i>
Map translational science	<i>iTrans</i>
Track patent, licensing, start-up activity	<i>iTech</i>

# *iTrans*: Tracking bench to bedside trends in scientific knowledge

Griffin Weber  
Identifying translation science  
within the triangle of biomedicine  
*J Trans Med* 11:126 (2013)



# *iTrans*: Tracking bench to bedside trends in scientific knowledge



# NIH-funded publications in 2012

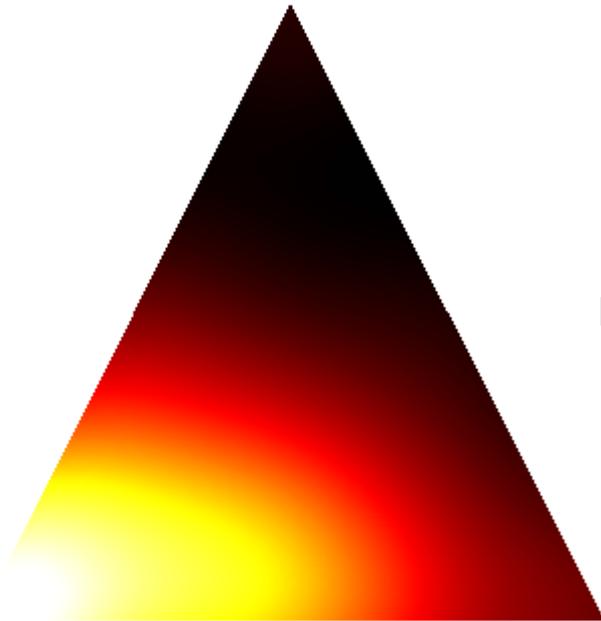
High article density



Low article density

National Institute of  
General Medical Sciences

Human

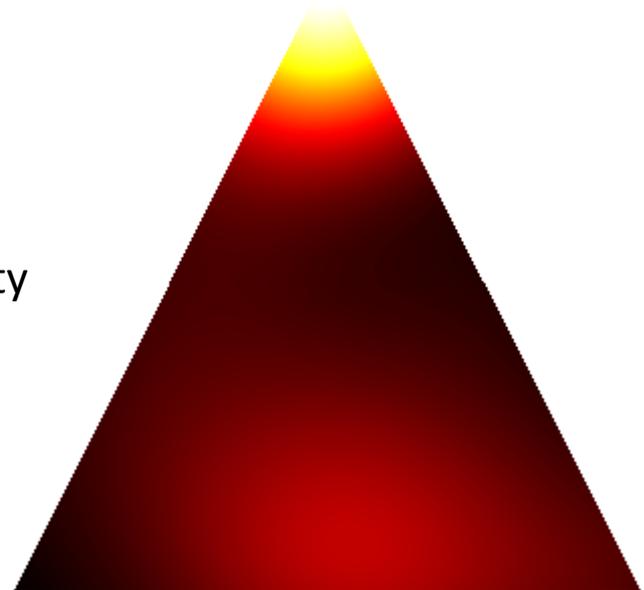


Cellular/  
Molecular

Animal

National Institute of Aging

Human

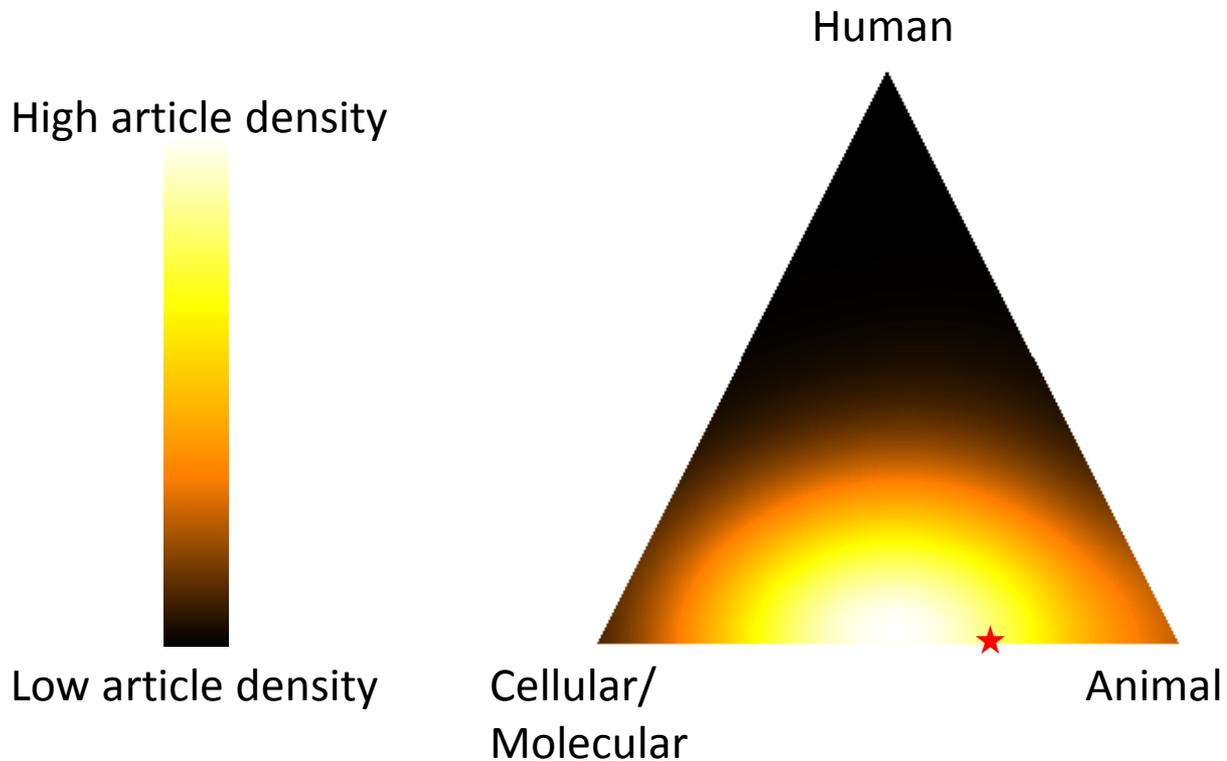


Cellular/  
Molecular

Animal

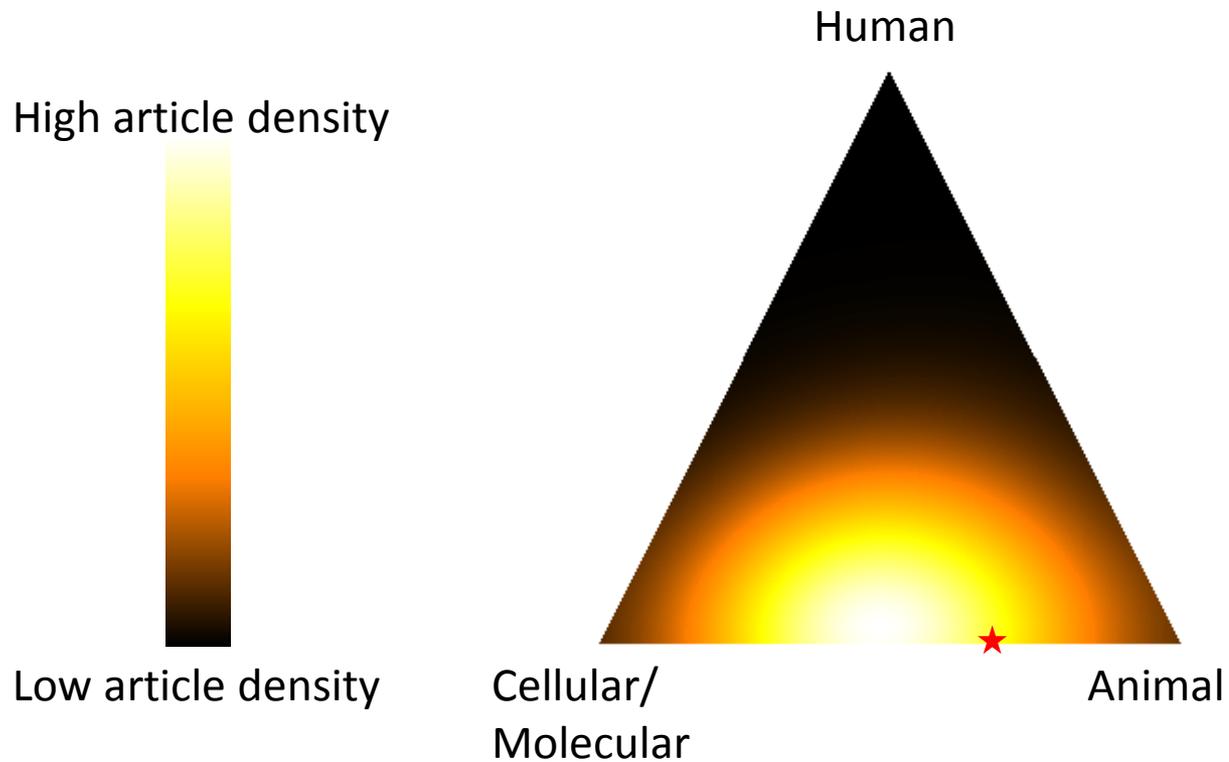
# 1998

Articles citing Fire et al., Potent and specific genetic interference by double-stranded RNA in *C. elegans* (1998) *Nature* 391:806-11 (★)



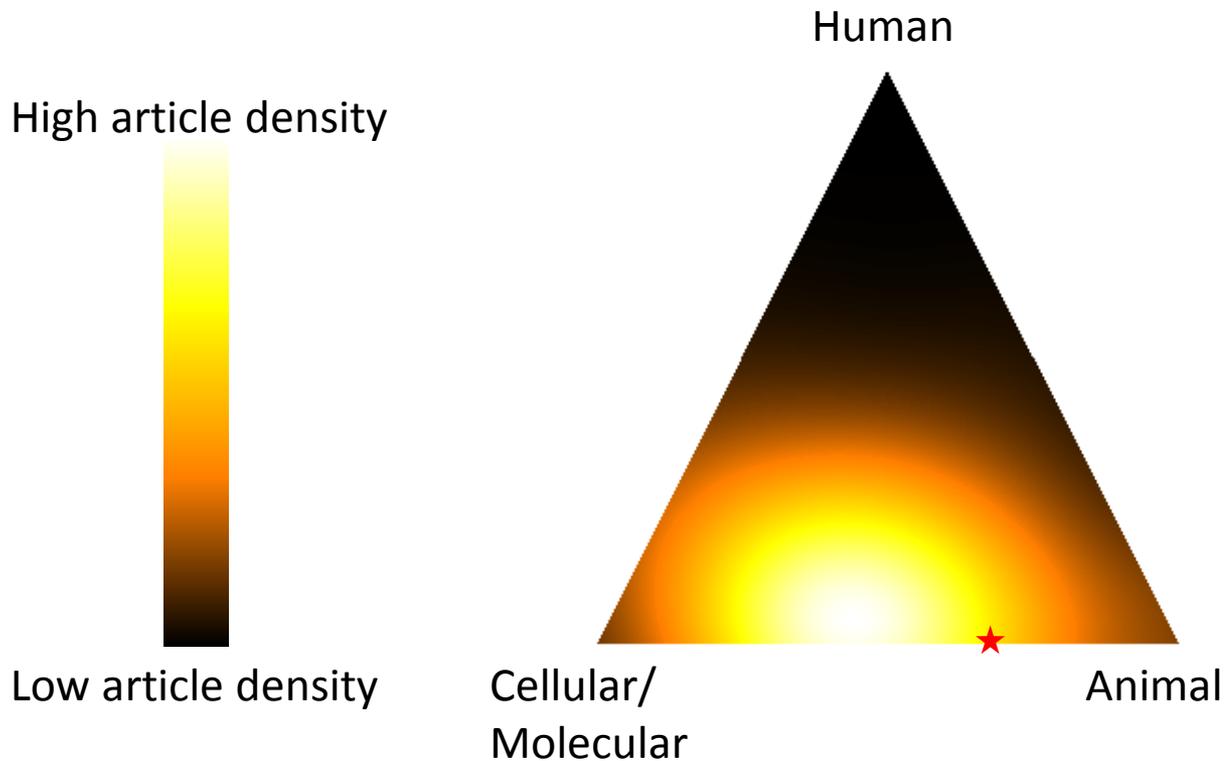
# 2000

Articles citing Fire et al., Potent and specific genetic interference by double-stranded RNA in *C. elegans* (1998) *Nature* 391:806-11 (★)



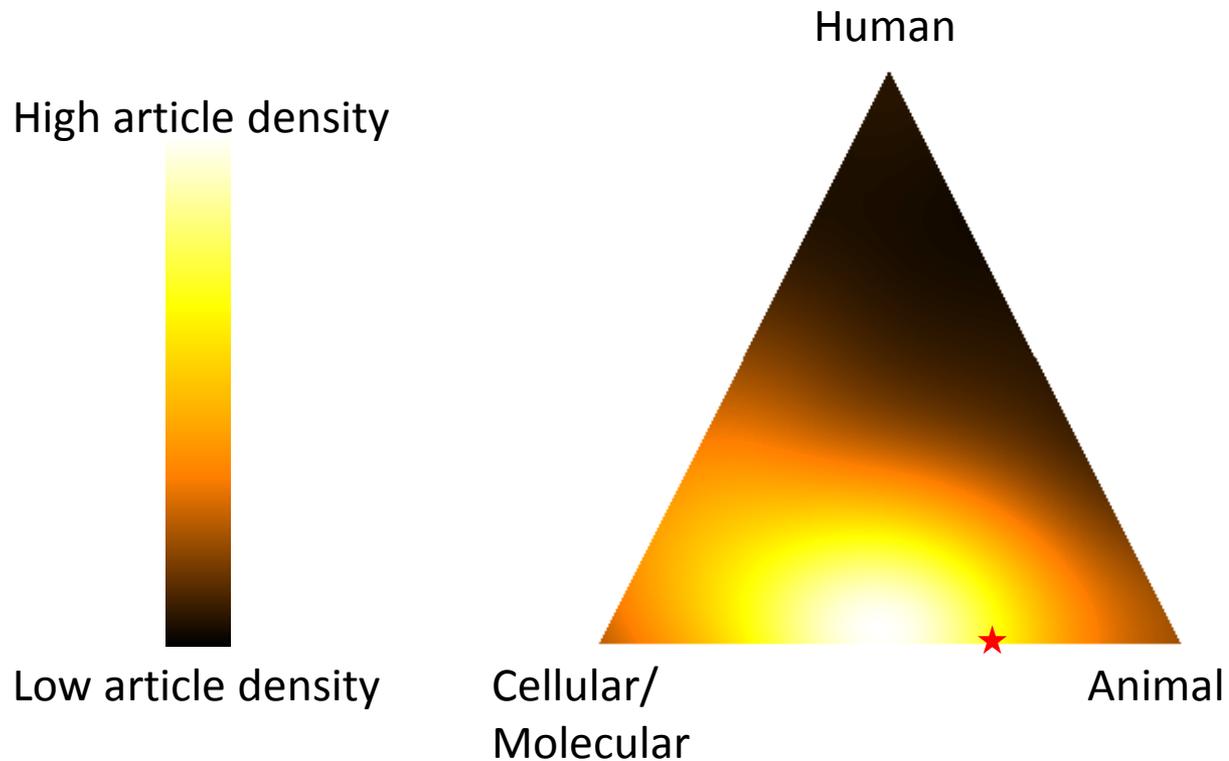
# 2002

Articles citing Fire et al., Potent and specific genetic interference by double-stranded RNA in *C. elegans* (1998) *Nature* 391:806-11 (★)



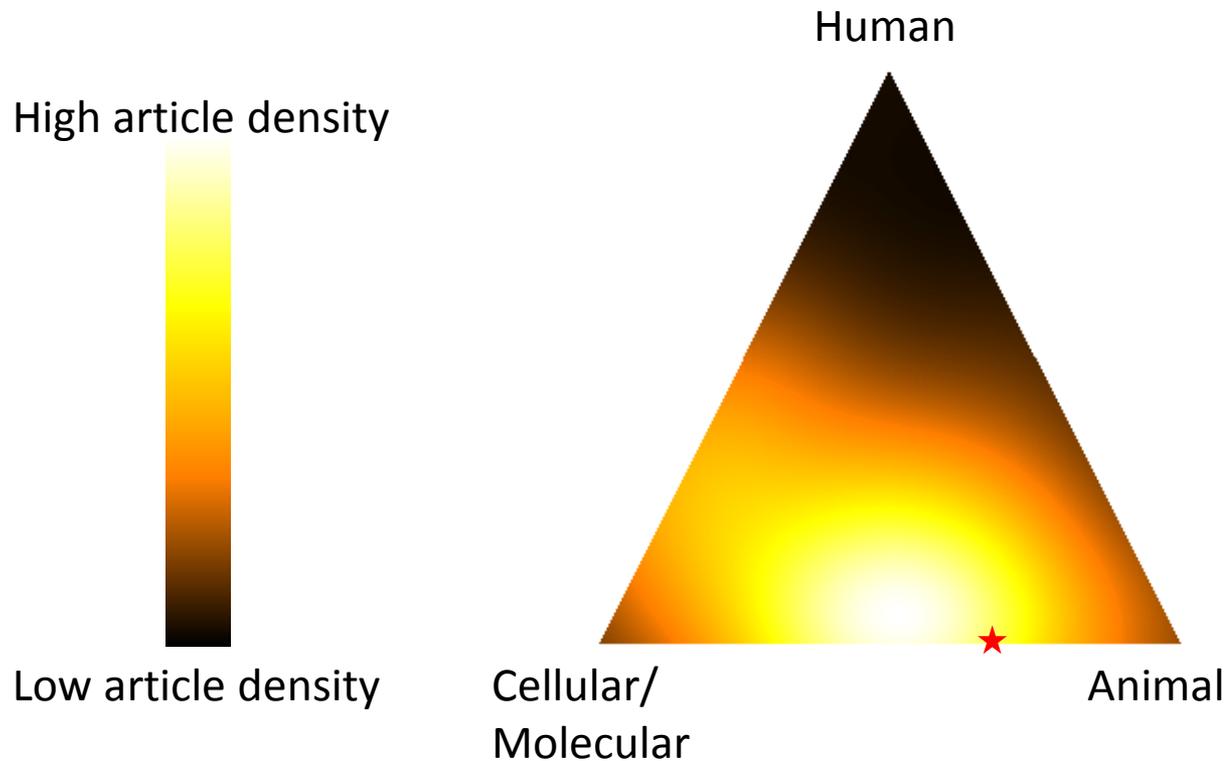
# 2004

Articles citing Fire et al., Potent and specific genetic interference by double-stranded RNA in *C. elegans* (1998) *Nature* 391:806-11 (★)



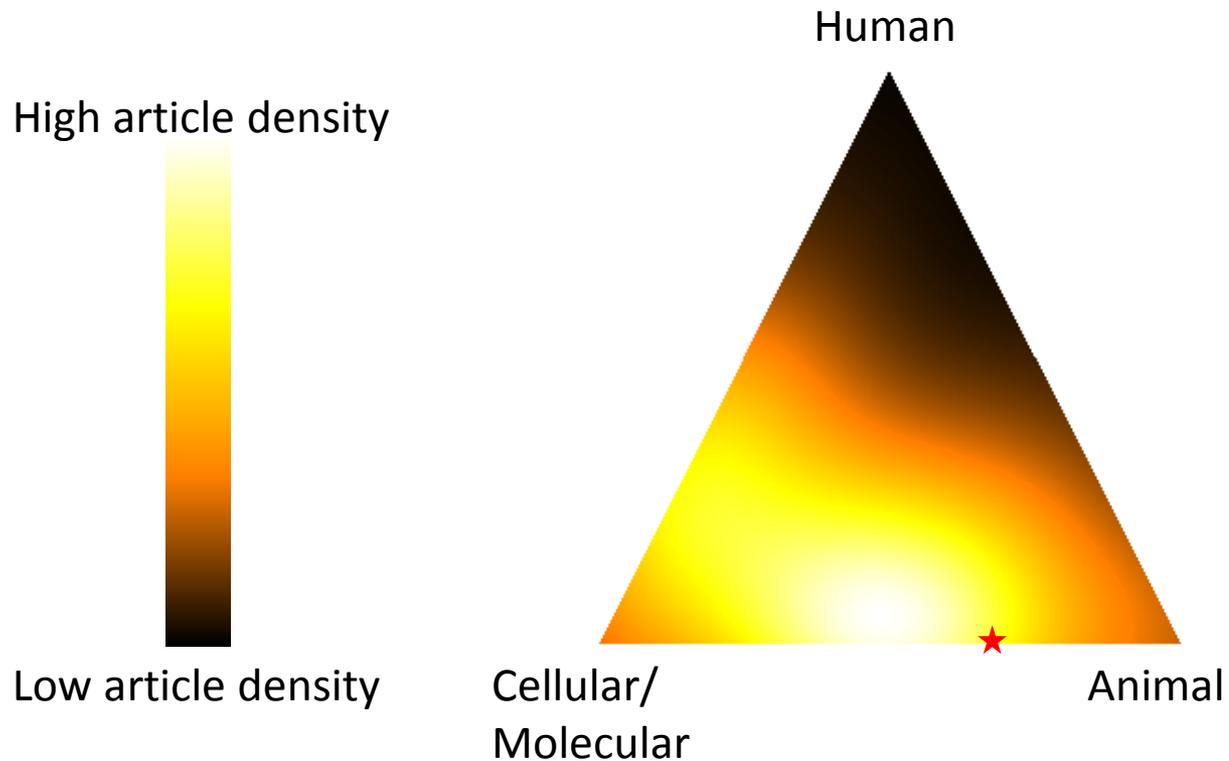
# 2006

Articles citing Fire et al., Potent and specific genetic interference by double-stranded RNA in *C. elegans* (1998) *Nature* 391:806-11 (★)



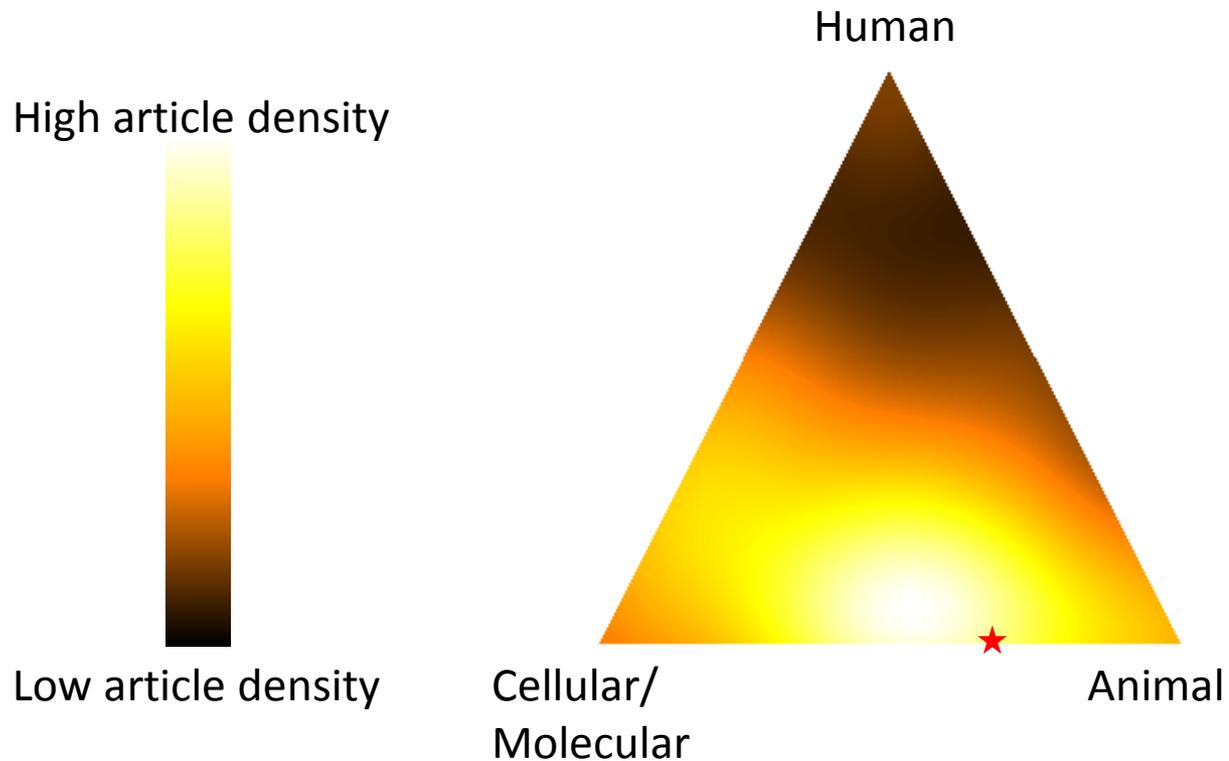
# 2008

Articles citing Fire et al., Potent and specific genetic interference by double-stranded RNA in *C. elegans* (1998) *Nature* 391:806-11 (★)



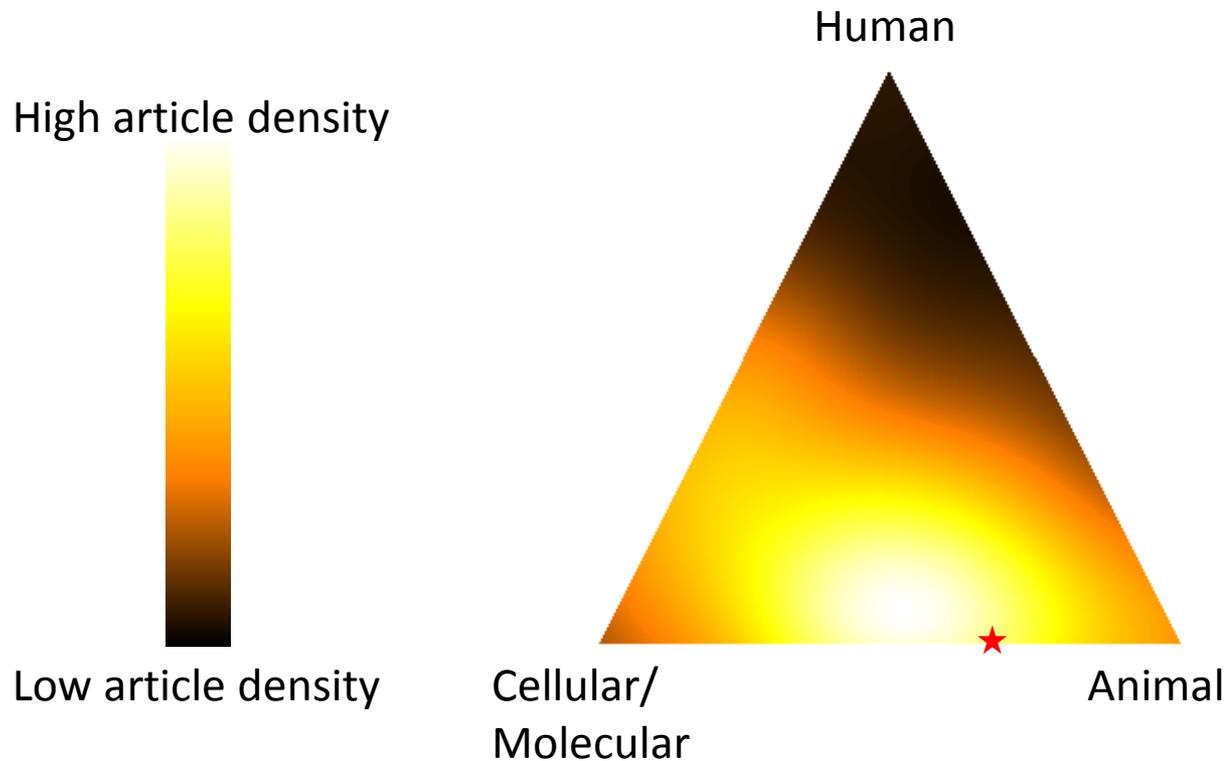
# 2010

Articles citing Fire et al., Potent and specific genetic interference by double-stranded RNA in *C. elegans* (1998) *Nature* 391:806-11 (★)



# 2012

Articles citing Fire et al., Potent and specific genetic interference by double-stranded RNA in *C. elegans* (1998) *Nature* 391:806-11 (★)



1998

2010

HOME PAGE TODAY'S PAPER VIDEO MOST POPULAR TIMES TOPICS

The New York Times **Science**

WORLD U.S. N.Y. / REGION BUSINESS TECHNOLOGY SCIENCE HEALTH SPORTS OPINION

ENVIRONMENT SPACE & COSMOS

February 7, 2011

**Drugmakers' Fever for the Power of RNA Interference Has Cooled**

By ANDREW POLLACK  
Published: February 7, 2011

When RNA interference first electrified biologists several years ago, pharmaceutical companies rushed to harness what looked like a swift and surefire way to develop new drugs.

Billions of dollars later, however, some of those same companies are now losing their enthusiasm for RNAi, as it is called. And that is raising doubts about how quickly, if at all, the [Nobel Prize-winning](#) technique for turning

**Multimedia**

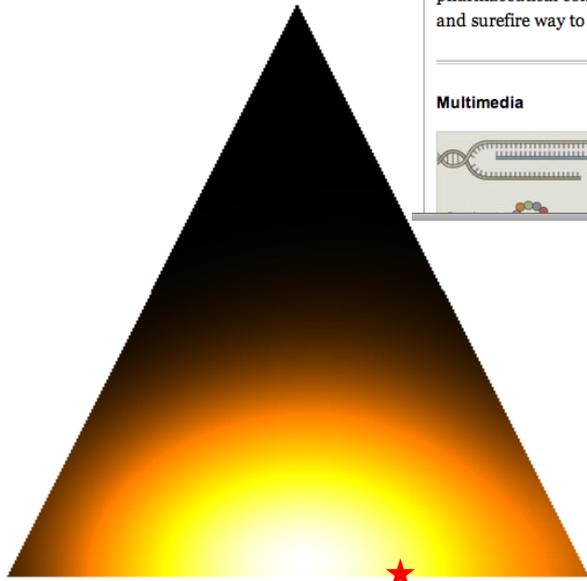


DNA  
mRNA transcript

TWITTER LINKEDIN PRINT REPRINTS SHARE

Human

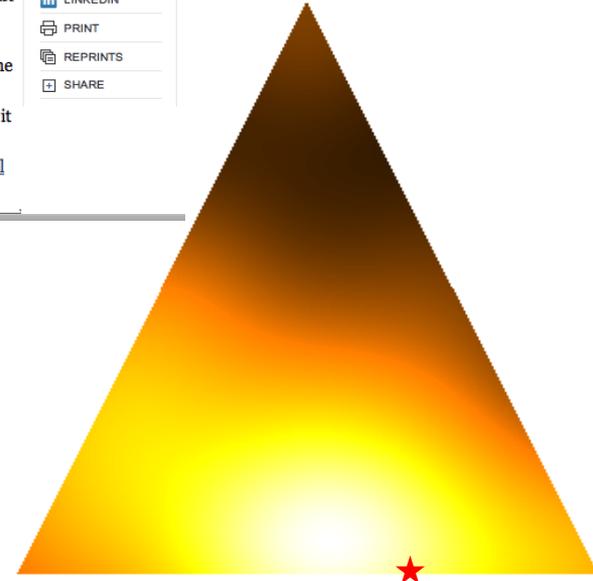
Human



Cellular/  
Molecular

Animal

Low article density



Cellular/  
Molecular

Animal

Articles citing Fire et al., Potent and specific genetic interference by double-stranded RNA in *C. elegans* (1998) *Nature* 391:806-11 (★)

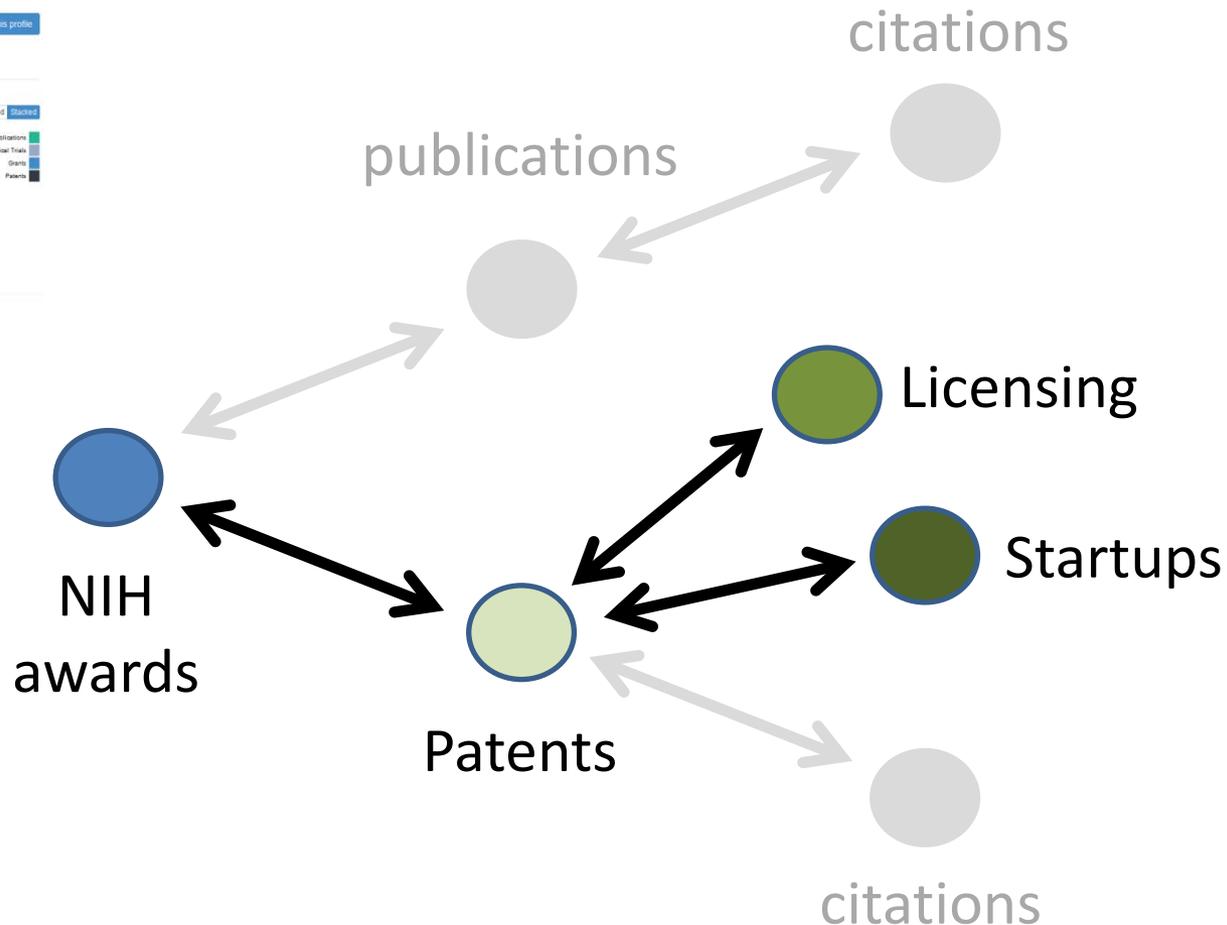
# Developing a science of portfolio analysis

- Use existing data-driven approaches to characterize research investments and the resulting impact
- Develop and deliver effective approaches and methodologies
  - Tools in development:

Functionality	Tool
Content analysis	IN-SPIRE + TBN ( <i>To be named later</i> )
Efficient disambiguation	<i>iClean</i>
Effective bibliometrics	<i>iCite</i>
Map translational science	<i>iTrans</i>
Track patent, licensing, start-up activity	<i>iTech</i>

# iTech: Patents, licensing and start-ups

## Connecting the dots (bidirectionally)



# Case study: the druggable genome



Journal home > Archive > Perspective > Opinion > Full Text

## JOURNAL CONTENT

- Journal home
- Advance online publication
- Current issue
- Archive
- Web Focuses

## Perspective

*Nature Reviews Drug Discovery* 1, 727-730 (September 2002)

### OPINION

## The druggable genome

Andrew L. Hopkins<sup>1</sup> & Colin R. Groom<sup>1</sup> [About the authors](#)

## Drug Discovery Today

Volume 10, Issues 23–24, December 2005, Pages 1607–1610



Update

## The druggable genome: an update

Andreas P. Russ Stefan Lampel

[Show more](#)

DOI: 10.1016/S1359-6446(05)03666-4

## nature methods

Techniques for life scientists and chemists

[Home](#) | [Current issue](#) | [Comment](#) | [Research](#) | [Archive](#) | [Authors & referees](#) | [About the journal](#)

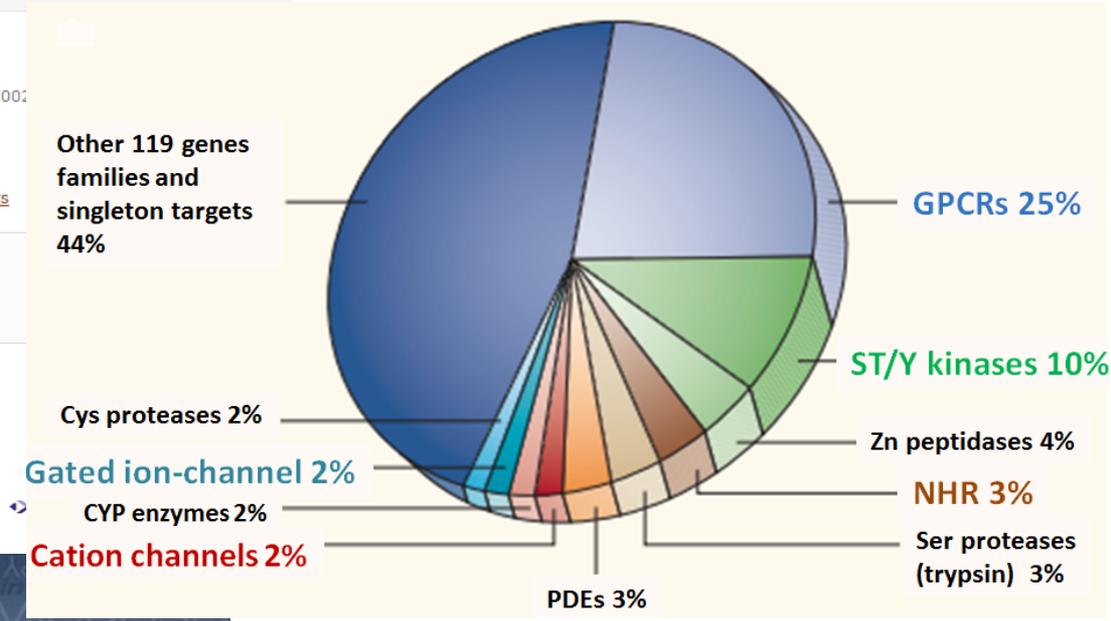
[home](#) > [archive](#) > [issue](#) > [brief communication](#) > [full text](#)

NATURE METHODS | BRIEF COMMUNICATION



## DGIdb: mining the druggable genome

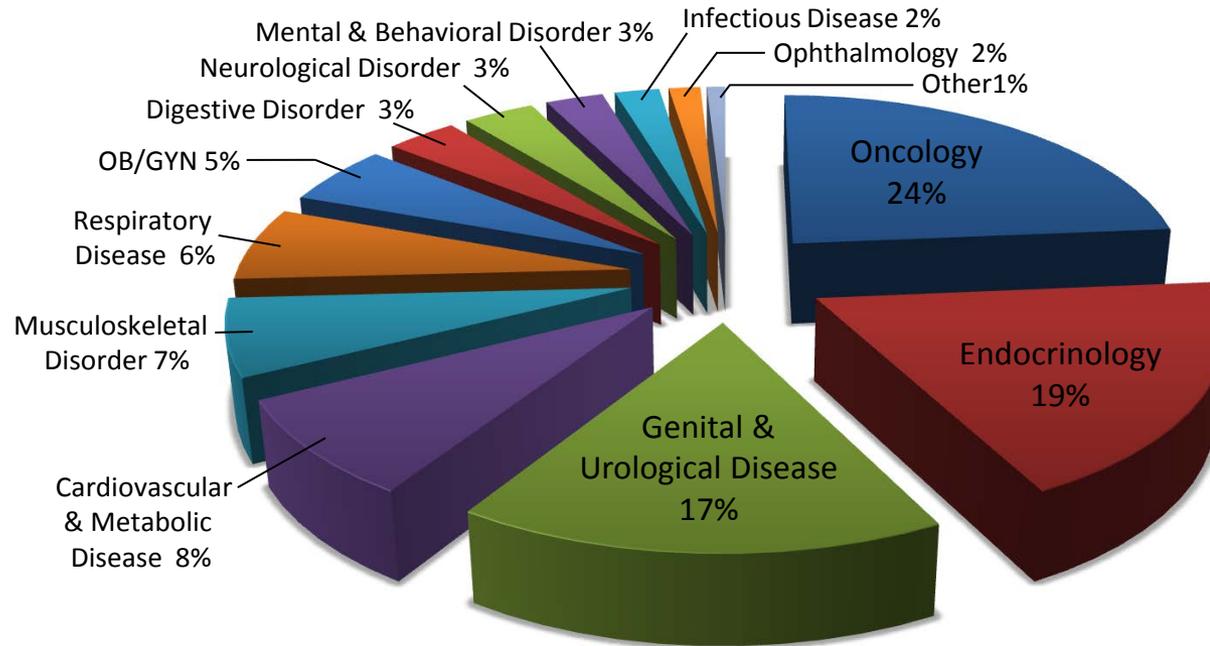
Malachi Griffith, Obi L Griffith, Adam C Coffman, James V Weible, Josh F McMichael, Nicholas C Spies, James Koval, Indraniel Das, Matthew B Callaway, James M Eldred, Christopher A Miller, Janakiraman Subramanian, Ramaswamy Govindan, Runjun D Kumar, Ron Bose, Li Ding, Jason R Walker, David E Larson, David J Dooling, Scott M Smith, Timothy J Ley, Elaine R Mardis & Richard K Wilson



Adapted from: Hopkins AL, Groom CR. The druggable genome. *Nat Rev Drug Discov.* 2002,1(9):727-30.

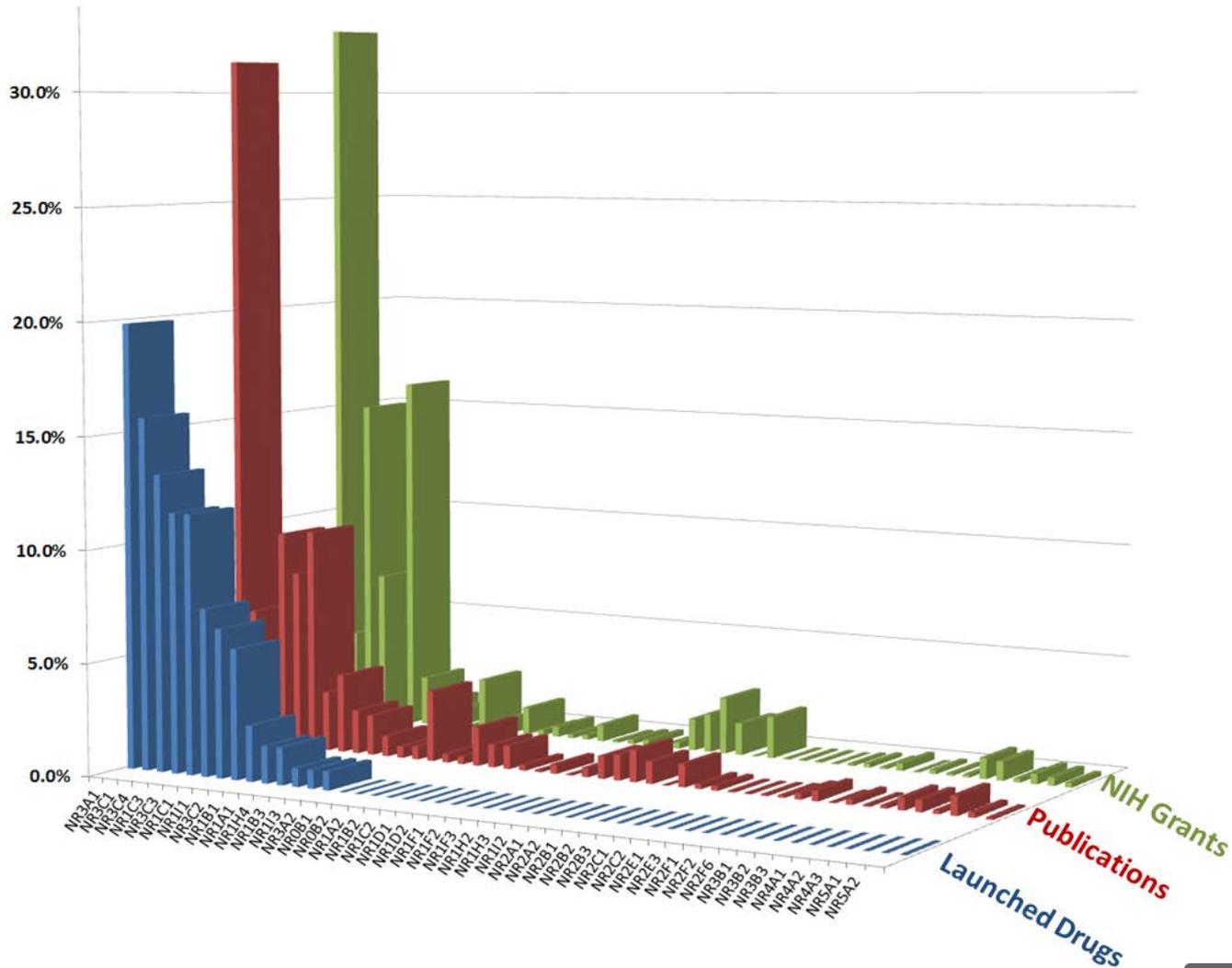
# The druggable genome: nuclear receptors

## Drugs Launched against Nuclear Receptors



# The druggable genome: nuclear receptors

Profile of NIH funding, publications, and launched drugs since 1940

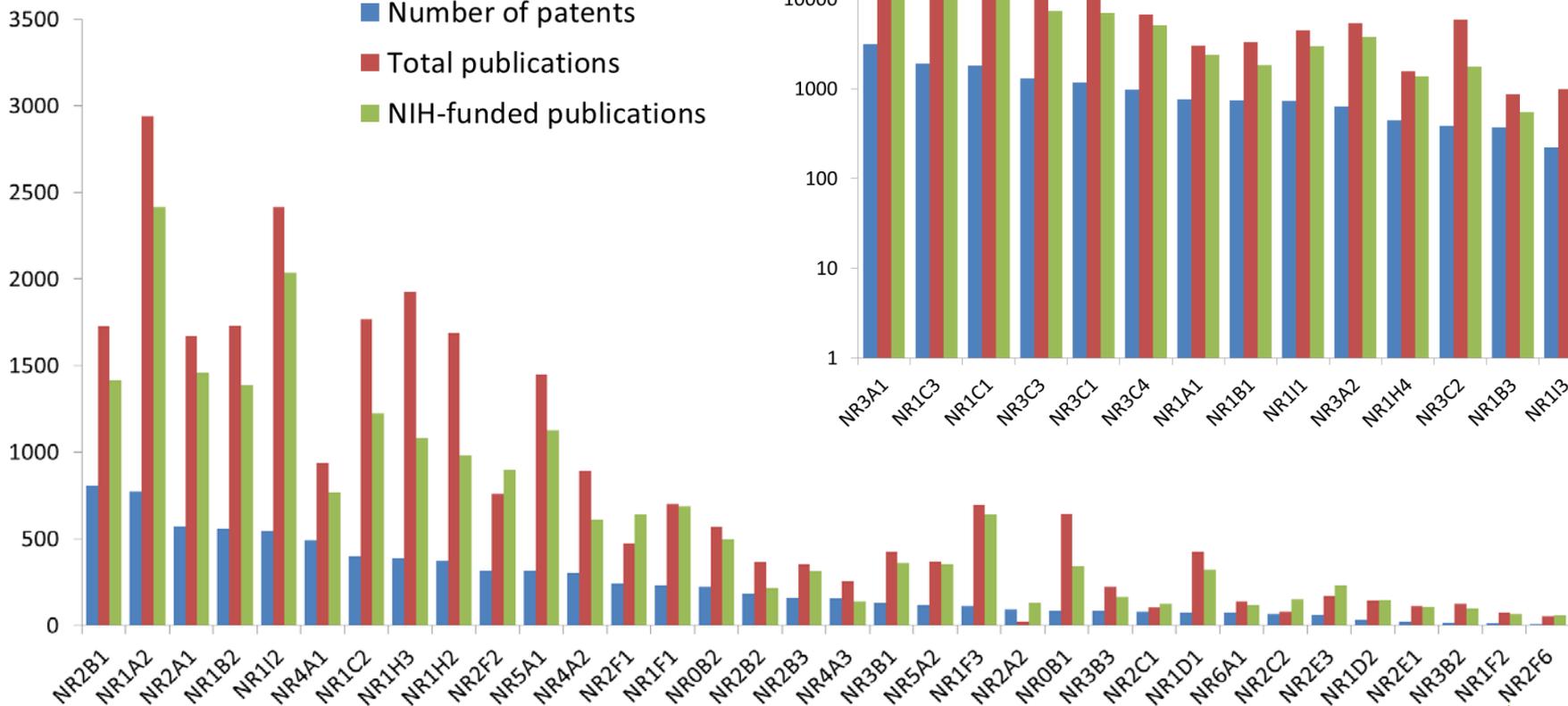
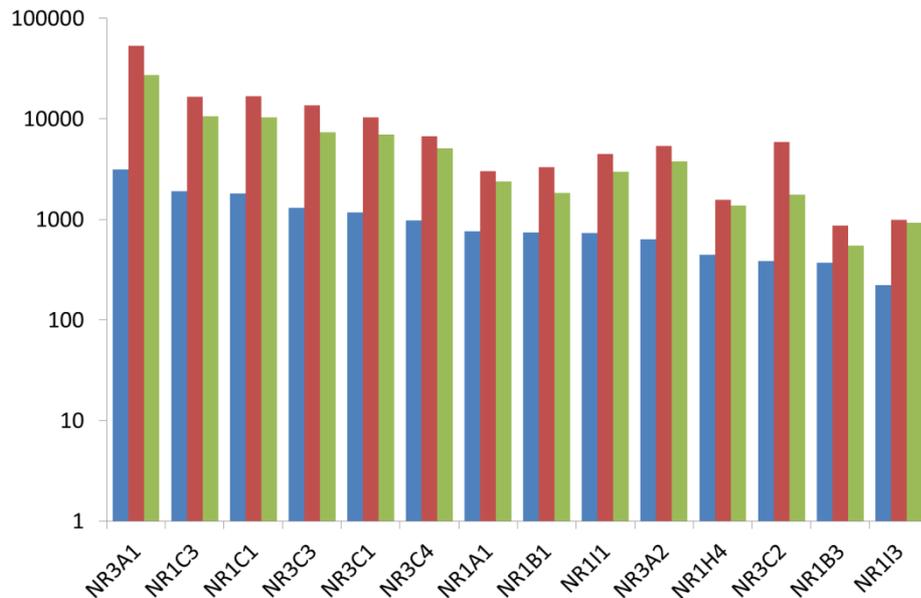


# The druggable genome: nuclear receptors

Do have drugs

Don't have drugs

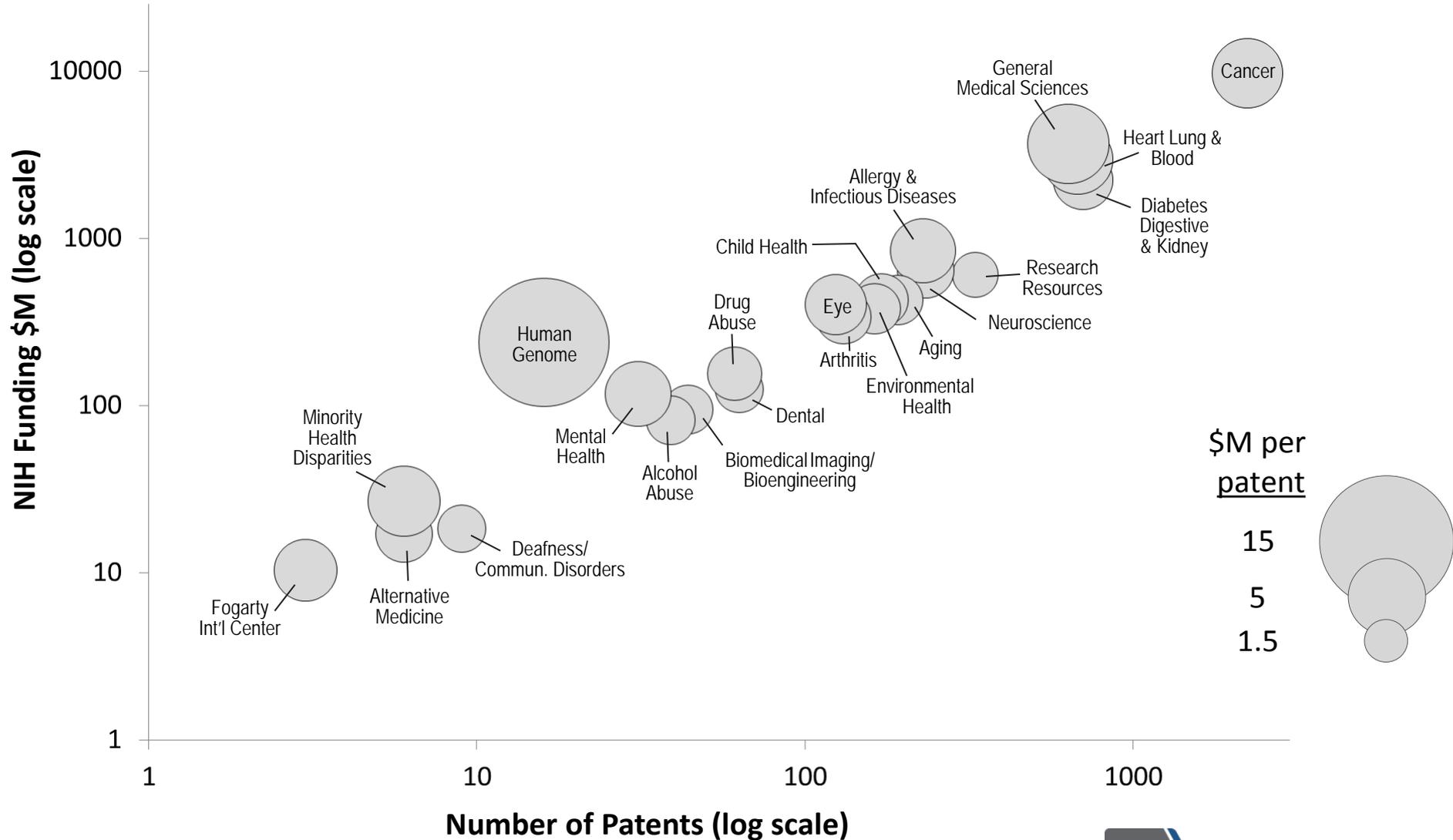
■ Number of patents  
 ■ Total publications  
 ■ NIH-funded publications



Bottom ~15% of the publication space

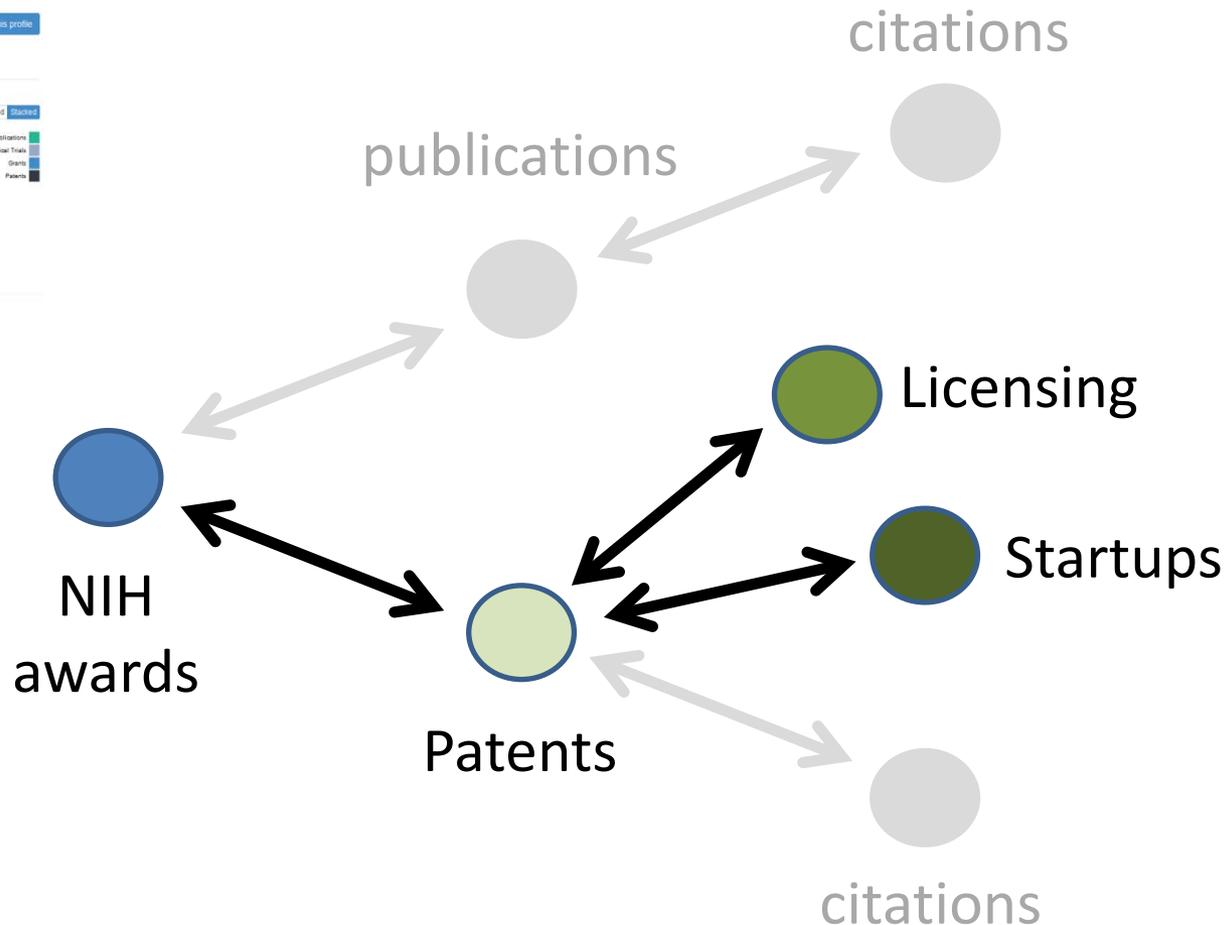
# The druggable genome: nuclear receptors

Awarded patents citing NIH-supported research on nuclear receptors, by Institute



# iTech: Patents, licensing and start-ups

## Connecting the dots (bidirectionally)



[OPA HOME](#)[TRAINING](#)[THE ANALYST](#)[TOOLS AND RESOURCES](#)[MEETINGS](#)[ABOUT US](#)

The OPA Tools Lab is located in B301 in building 1. For access please contact us.

For updates on training and other OPA activities, please sign up for our listserv.

▶ [OPA analysis of Nicholson and Ioannidis dataset](#)



[Home](#) » [OPA Home](#) » [Tools and Resources](#)

## Resource Center

Are you interested in details of what tools are available to NIH staff and how to get hold of them? Would you like to see examples of how tools can be used in analyses? Do you need a refresher on a tool that you took a class for?

...Welcome to the OPA Resource Center.

OPA is currently developing computational tools that you can access from your desktop. Tools currently in development include:

Tool	Functionality	Beta Release Date
<i>iClean</i>	Efficient disambiguation	Spring, 2015
<i>iCite</i>	Effective bibliometrics	Spring, 2015
<i>iTrans</i>	Map translational science	Fall, 2015
<i>iTech</i>	Track patent, licensing, and start-up activity	Summer, 2016

Let us know if you are interested in beta-testing one or more of these tools in the coming months. In your email, please indicate which tool(s) you are interested in testing.

# Acknowledgments

## OPA Analysts

Jean Yuan

Ian Hutchins

Paula Fearon

Carole Christian

Rob Harriman

Patricia Forcinito

Ling Bai

Terry Bishop

Kristina McLinden

Rayna Carter

## OPA Software Developers

Fai Chan

Kirk Baker

Marius Oster

Ehsan Haque

Kevin Small

## OPA IT Specialist

Chuck Lynch

## NIH National Library of Medicine

Tom Rindfleisch

## NIH Center for Information Technology

Calvin Johnson

Krishna Collie

## ÜberResearch

Steve Leicht

Ashlea Higgs

## Pacific Northwest National Labs

Dennis McQuerry

The screenshot shows the NIH Office of Portfolio Analysis (OPA) website. At the top, there are navigation links for the U.S. Department of Health & Human Services, National Institutes of Health, and Division of Program Coordination, Planning, and Strategic Initiatives (DPCPSI). The NIH logo and 'Office of Portfolio Analysis' are prominently displayed. A search bar and printer-friendly options are in the top right. A left sidebar contains a navigation menu with links to OPA HOME, TRAINING, THE ANALYST, TOOLS AND RESOURCES, MEETINGS, and ABOUT US. Below the menu is a text box about the OPA Tools Lab location and a link to sign up for updates. The main content area features three circular network diagrams with highlighted clusters in red, blue, and green. Below these are sections for 'Consultations' and 'Office Hours'. The 'Consultations' section describes the service and provides a link to schedule. The 'Office Hours' section states that analysts are available every other Wednesday from 10:00 AM to 12:00 PM. Below this is the 'Portfolio Analysis Interest Group (PAIG)' section, which describes the group's purpose and provides the next meeting date: January 27th, 2015. At the bottom, there is a link to 'The Analyst - the OPA blog' and a 'View More' button.

**Questions?**  
**Comments?**