The Nature of scientific communication in the 21st century

9th October, 2013

James Butcher PhD
Associate Director
Open Publishing
Nature Publishing Group

- ~120 titles
- 34 ‘Nature’ titles
- 60 academic and society journals
- Scientific American magazine

- 1500 employees in 17 offices on 5 continents
- Over 25 academic and society partners
- Online services, databases, conferences/events, multimedia
- We aim to be the best – not the biggest
  - 36% of the top 50 science journals by Impact Factor; (2012 JCR*)
  - 18 NPG journals #1 in their fields by Impact Factor (2012 JCR*)

*Source: 2012 Journal Citation Report, (Thomson Reuters, 2013)
Nature's mission statement – written in 1869 – still guides us today…

“First, to serve scientists through prompt publication of significant advances in any branch of science, and to provide a forum for the reporting and discussion of news and issues concerning science.

Second, to ensure that the results of science are rapidly disseminated to the public throughout the world, in a fashion that conveys their significance for knowledge, culture and daily life.”
“Study the past if you would define the future”

Confucius
Role of journals in the 20\textsuperscript{th} century

- Selectivity
- Certification of quality (peer review)
- Provide readership (edit, print and distribute paper copies)
- Archiving (version of record)
- Registration of precedence
- Reward researchers

Question: how will this change in the 21\textsuperscript{st} century?
The evolution of *Nature*
**Evolution of Nature: 1953**

**DNA Structure 1953**

1 page
2 authors
1 figure
no data
Evolution of Nature: 2001

Human Genome 2001
62 pages, 150 authors, 49 figures, 27 tables
Evolution of Nature: 2010

1000 Genomes Project 2010

12 pages, 76 institutions, 5 related publications in Science and Genome Research
Evolution of research: 2012

Encode Project 2012
30 papers, 3 journals, 13 threads
Threads in ENCODE

• At the heart of the ENCODE explorer lie the threads.

• These are a new way in which to explore the wealth of information collectively described by the 30 papers published across three different journals.

• They complement the papers by highlighting topics that are otherwise covered only in subsections of individual papers.

• Each thread consists of relevant paragraphs, figures and tables from across the papers, united around a specific theme.
An expansive human regulatory lexicon encoded in transcription factor footprints

Nature
doi: 10.1038/nature11212
(2012)

Leave explorer & read full paper
Most major histocompatibility complex (MHC) peptide ligands that can be bound with high affinity are destined for presentation by MHC class I molecules during degradation and cytosolic aminopeptidase. The proteasome and cytosolic aminopeptidase are responsible for MHC class I antigen processing. However, in many cases, epitopes require an additional step, where the transporter associated with antigen processing (TAP), is responsible for that activity identified an ER-associated endopeptidase that was termed ER-associated aminopeptidase 1 (ERAP1) or ER aminopeptidase associated with antigen processing (ERAP1). Studies of ERAP1-deficient cells have shown that ERAP1 generates more larger precursors, but it can destroy other epitopes by trimming them below class I binding. The peptide repertoire carried by MHC molecules of wild-type and ERAP1-deficient cells differ in wild-type mice infected with LCMV, and T-cell responses against many viral antigens were reduced in ERAP1-deficient mice. Furthermore, cross-presentation of cell-associated epitopes was lower in ERAP1-deficient mice. ERAP1-dependent epitopes were shown to be important in resistance to Toxoplasma gondii.
The article of the near future...

Molecular entities:
- Compounds
- Genes
- Proteins
- Techniques

Article level metrics

Authors

Data & Accession codes

Updates (corrections)

Comments

Social media

Related content & editorial perspective

Figures

Sections

videos

visualisation tools:
- 3D structure
- Genome browser
- Open microscopy
- Environment

blogs
Text, video podcast

Protocol exchange
The Fourth Paradigm (the importance of data)
Four science paradigms
Jim Gray (1944 - 2007)

• Experimental science
  – Test hypotheses

• Theoretical science
  – Creation of models, equations, generalisations

• Computational science
  – Simulating complex phenomena

• Data exploration
  – Unification of experimental, theoretical, and computational
Jim Gray’s vision

“A world of scholarly resources—text, databases, and any other associated materials—that are seamlessly navigable and interoperable.”
...which is a vision of...

An edifice of text, numbers, equations, data, software, images, graphics, videos, tables, annotations and metadata that’s:

– Seamless and hyperdimensional
– Readable, searchable and computable by humans and machines
– Checked and ranked by editors and crowds
– Sourced from journals, labs, community databases, institutional databases, meta-analyses, grey literatures,...
– Underpinned by community standards of nomenclature, annotation, sharing and integrity (technical and ethical)
– Free to all users
– Financially sustainable
Unique identifiers are key!

- DOI (articles)
- ORCID (people)
- FundRef (funders)
- Ringgold (institutions)
- Datacite (data)
The growth of open access
Open access is growing

Figure 2 Annual volumes of articles in full immediate open access journals, split by type of open access journal.

Laakso and Bjork BMC Medicine 2012; 10:124
Open access output by region

Laakso and Bjork BMC Medicine 2012; 10:124
Open access at Nature Publishing Group

- 2009-2011: all non-Nature journals offer OA option
- 2011: *Scientific Reports*
- 2013: Frontiers
Open access at Nature Publishing Group

- In 2013 Nature Publishing Group will publish ~13000 OA papers
  - 8000 in Frontiers
  - 2500 in Scientific Reports
  - 500 in Nature Communications
  - 2000 in other NPG journals
Nature Communications

- Launched in April 2010
- Scope: all areas of the natural sciences
- Authors can choose subscription or OA at acceptance
- ~20% accept rate
- In-house editorial team
- Offers three Creative Commons licenses
  - CC BY ($5200)
  - CC BY-NC-ND ($4800)
  - CC BY-NC-SA ($4800)
• Received ~20,000 submissions since launch

• The journal received ~1200 submissions in September 2013; (Nature receives ~900 / month)

• ~33% of submissions were previously considered at another Nature journal
The editorial team has grown...

18 months ago...
The editorial team has grown...
Publications

- Published 701 papers in 2012
- Expect to publish ~1600 papers in 2013, of which ~500 will be OA
- In 2012, the 16 Nature Research Journals published ~2100 papers

- 53% biology
- 33% physics
- 11% chemistry
- 3% earth and environmental sciences
• In 2012, 41% of authors chose OA
• In 2013, 31% of authors chose OA

• OA uptake rate varies by subject
• In 2013:
  – 39% of biologists chose OA
  – 34% of physicists chose OA
  – 23% of earth scientists chose OA
  – 22% of chemists chose OA
Metrics

Web traffic

- 5.5m page views in 2012
- 7.2m page views in Jan to Aug 2013

Impact factor

- The 2012 impact factor is 10.015
- ~150 journals (out of 8500) have an IF >10
Open access and peer review
The sting in this week’s *Science*

- *Science* journalist submitted bogus paper to 255 OA journals
  - 157 journals accepted it
  - 98 rejected it

- 153/255 of journals (60%) did not peer review the paper
  - 86/153 journals accepted it

- 102/255 journals (40%) peer reviewed the paper
  - 71/102 journals accepted it
Location of the journals

Follow the money

- Accepted
- Rejected
- Bank
- Editor
- Publisher

Tangled web. The location of a journal's publisher, editor, and bank account are often continents apart. Explore an interactive version of this map at http://scim.org/DA-String.
One major fault in the study

“It’s nuts to construe this as a problem unique to open access publishing, if for no other reason than the study didn’t do the control of submitting the same paper to subscription-based publishers.”

“The real story is that a fair number of journals who actually carried out peer review still accepted the paper, and the lesson people should take home from this story not that open access is bad, but that peer review is a joke.”

Mike Eisen, co-founder of PLoS
What’s the evidence that peer review works?
What’s the purpose of peer review?

• Selecting what should be published

• Improving what is published
  – Detecting errors
  – Improving the arguments

• Detecting misconduct
  – Plagiarism
  – Fabrication
  – Falisification
“Editorial peer review, although widely used, is largely untested and its effects are uncertain”

Peters and Cecia (1982)

Methods
• They changed the names and institutions on 12 published articles, one from each of 12 American psychology journals, to fictitious names.
• Resubmitted the articles to the journals that published them

Results
• 3 of the 12 journals noticed that the papers had been published before and rejected them
• 9 of the 12 journals didn’t spot the deliberate error
• 8 of those 9 journals rejected the papers, even though they had previously published them (with different authors and institutions)!!!
• The grounds for rejection were often because of “serious methodological flaws”
“The mistake, of course, is to have thought that peer review was any more than a crude means of discovering the acceptability — not the validity — of a new finding. Editors and scientists alike insist on the pivotal importance of peer review. We portray peer review to the public as a quasi-sacred process that helps to make science our most objective truth teller. But we know that the system of peer review is biased, unjust, unaccountable, incomplete, easily fixed, often insulting, usually ignorant, occasionally foolish, and frequently wrong.”

“There seems to be no study too fragmented, no hypothesis too trivial, no literature too biased or too egotistical, no design too warped, no methodology too bungled, no presentation of results too inaccurate, too obscure, and too contradictory, no analysis too self-serving, no argument too circular, no conclusions too trifling or too unjustified, and no grammar and syntax too offensive for a paper to end up in print.”

“If peer review was a drug it would never be allowed onto the market.”

That’s it? That’s peer review?
How much does peer review cost

• The Research Information Network calculated that the global cost of peer review was £1.9 billion

http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/activities-costs-and-funding-flows-scholarly-commu
Important!

Absence of evidence and Evidence of absence

...are not the same thing!
Authors like peer review

• All *Scientific Reports* authors are asked to complete a survey
• “The content of my paper was improved as a result of the peer-review process”
  – 60% “strongly agree”
  – 33% “somewhat agree”
The paper is not sacred

Peer review continues long after a paper is published, and that analysis should become part of the scientific record, say Adam Marcus and Ivan Oransky.

The Borstel board found her “ultimately responsible” for the misconduct in her lab and for failing to deal with it in a timely and open manner. (Bulfone-Paus has made few public statements about the case, but she has noted that her results were confirmed by other researchers.) In 2011, journals retracted 13 of her published articles, the stated reasons varying from detailed explanations such as “evidence of data manipulation in Fig. 2C, 4B, and 9, a clear violation of ASM’s ethical standards,” to the wholly unhelpful “This article has been withdrawn by the authors.”

Lines like the latter make us want to pull out whatever hair we have left on our heads. Journal readers should find them similarly frustrating. But we singled out this particular notice for concern not because it said too little, but because, in our view, it allowed the authors to say too much.

Too Laissez-faire?
The EMBO Journal’s notice also included this: “The authors declare that key experiments presented in the majority of these figures were recently reproduced and that the results confirmed the experimental data and the conclusions drawn from them.”

The statement from Bulfone-Paus and her colleagues described new data and signalled to readers that they could still rely on the original paper, even though it had been retracted. It suggested that the journal stood behind the statement. But when we asked the editor whether that was the case, we were told: “We did not formally investigate this case at the journal and we have not seen this data as
Two options…

• Peer review then publish
• Publish then peer review
  – Letters to the editor (slow)
  – Online comments (fast, but hardly used)
  – Blogs (fast, but not indexed or linked to paper)
  – Faculty of 1000 (comments of and flags papers)

• To be useful, post-publication peer review needs to be discoverable by readers of the original paper.
CrossMark

• Allows readers to easily find corrections, retractions and other revisions

• Potentially allows publishers to link through to blog posts, media coverage etc. that aren’t hosted by them
Readers’ comments

• Most papers have no comments posted on them
• How should we judge such papers?
  – No comments = robust paper?
  – No comments = no one’s checked?
• Why are there so few comments?
  – Because there is no reward system in place?
  – Because researchers don’t like to criticise in public?
Peer review quality is independent of a journal’s business model

- UK parliamentary report: Peer review in scientific publications (2011)

“We encourage increased recognition that peer-review quality is independent of journal business model. For example, there is a ‘misconception that open access does not use peer review’”
OA and peer review: conclusions

• Open access does not equal poor quality

• Peer review is imperfect, but better than the other systems we have

• Let the buyer beware!

• Go for an open access brand you can trust
  – Check Jeffrey Beall’s list of predatory publishers
  – Scientific Reports (Nature Publishing Group)
  – Frontiers (Nature Publishing Group)
  – PLoS One
  – BioMed Central
  – Hindawi
The future of science publishing
Role of journals in 21st century

- Selectivity
- Certification of quality (peer review)
  - Post-publication or pre-publication peer review?
- Provide readership
  - Mark up metadata to enhance discoverability
  - Create a web of navigable content
  - Provide interactive datasets
- Archiving (version of record)
- Registration of precedence
- Reward researchers
  - For data collection and curation
  - For peer review
My predictions: medium term (not necessarily NPG’s)

• Open access will become more prevalent
  – But degree of implementation will vary by field
• Open data will become more prevalent
  – Common metadata standards will be essential
• Funders will start to ‘aggressively’ mandate open access
• Fewer journals will publish more papers
  – The rise of the megajournal
• Economies of scale will become increasingly important
  – Small society journals will suffer
A Crisis in Communication?

“A day will come when journals will be superseded as a means of publishing new research”

Theodore “Robbie” Fox (1965), Editor of *The Lancet* 1944-1964
My predictions: long term (not necessarily NPG’s)

• Filtering and quality control will always be valuable

• Top journals won’t necessarily publish research papers
  – Controversial!

• Research papers will be published in central repositories (journals?)
  – Linked to underlying data
  – International metadata standards

• Branded publications will dip into the pool and select papers to highlight
“Prediction is very difficult, especially if it's about the future.”

Niels Bohr
Thank you!