

## 第 1 回 SPARC Japan セミナー2016

「オープンアクセスへの道」

# 生命科学分野における 研究者の投稿先雑誌選択趣向と OA への意味づけ

坊農 秀雅

(情報・システム研究機構ライフサイエンス統合データベースセンター)

### 講演要旨



オープンアクセス (OA) 誌に掲載された論文は、インターネット検索エンジンの発達に伴い、目に触れる機会が格段に増えることとなり、その結果 PLOS (Public Library of Science) や BMC (BioMedCentral) といったゴールド OA 誌に投稿される論文が増え、古くからある論文誌も OA オプションが選べるようになってきた。その一方、OA 紙で必要となる論文掲載料は、研究費から捻出する状況となっており、実験するのに必要不可欠な試薬代や消耗品代を圧迫し、研究活動を行う上で無視できない存在となっている。OA 紙に論文を載せて広く自分の研究を知ってもらいたいが、それにはお金がかかる、そういった状況となっているのが現状である。本講演では、生命科学研究者の周辺で起こってきた投稿先雑誌選択趣向の変化について実経験に基づいて紹介する。



### 坊農 秀雅

理化学研究所において FANTOM (Functional annotation of mouse) プロジェクトの立ち上げに関わった後、埼玉医科大学ゲノム医学研究センターを経て、2007年7月より情報・システム研究機構ライフサイエンス統合データベースセンター (DBCLS) にて統合データベースプロジェクト (統合DB) に従事。統合DBの広報・普及活動として統合TVや統合データベース講習会AJACSの立ち上げに関わり、現在は国立遺伝学研究所にて日本DNAデータバンク (DDBJ) と連携して大規模塩基配列データ利用環境の構築とそれを利用した生命科学研究を行っている。京都大学博士 (理学)。

私は生命科学の研究者で、普段はライフサイエンス統合データベースセンター (DBCLS) にいます (図1)。本務は生命科学分野における公共データベース統合利用環境の構築ならびにその利用普及で、利用普及のためのウェブツールを作成したり、講習会を行ったりしています。また、遺伝子配列データを使ったデータ解析もしています。このようなことをずっと 20 年ぐらい続けています。もう一つ、「BMC Genomics」のアソシエイトエディターも務めています。

私は生命科学の分野にいますが、それが全てではなく、researchmap で私のエントリーを見ていただくと

図 2 に書いてあるように、研究分野はゲノム科学、分

坊農 秀雅 (ほうのう ひでまさ)

- ・ 本務(特任): 生命科学分野における公共データベース (DB) 統合利用環境の構築ならびにその利用普及
- ・ ライフサイエンス統合DBセンター (DBCLS) @三島
- ・ (一部のエフォートで) 研究活動: 遺伝子データ解析
- ・ 理研時代に FANTOM プロジェクトの立ち上げ
- ・ 低酸素・酸化ストレスの分子生物学
  - ・ パーキンソン病モデルカイコ
  - ・ ハダカデバネズミ (がん・老化モデル)
- ・ BMC Genomics Associate Editor

DBCLS For Life Science  
TOGO TV All of Gene Expression RefEx  
Licensed under CC-BY 4.0 ©2016 HIDEKASA BONO/DBCLS  
Licensed under CC-BY 4.0 ©Togo picture gallery by DBCLS

(図 1)

子生物学などなので、こんなことを言っている人もいと参考程度に受け止めてください。

### 生命科学研究者から見たオープンアクセス

私は生命科学の研究者と普段よく話をしたり、学会へ行ったりしますが、「オープンアクセス」という言葉は知っていても、ゴールド、グリーンという話は誰も知りません。そのような話になったことは一度もありません。生命科学の研究者がオープンアクセスと言くと、それはゴールドのことを指しています。

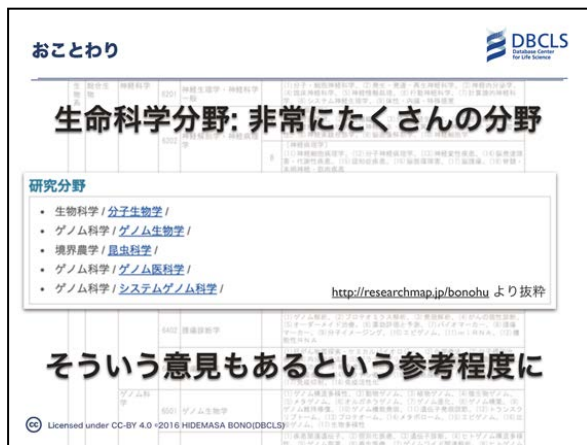
グリーンオープンアクセスへ出すということは、生命科学者から言ってみればインパクトファクターの付いていない雑誌に出すようなもので、よほど特殊な事情がない限り、まず出さないです。また、特に生命科学がそうなのかもしれませんが、ピアレビューを経たものでないと論文の業績として認めないという伝統があります。ピアレビューを経ずに、誰からでも見られるようにすることは強い抵抗があります。

私が思う研究者から見たオープンアクセスの良い点は、誰にでも自分の論文を場所を問わず見てもらえることです。インターネットでキーワードを検索されれば自分の論文がヒットします。生命科学でもいろいろな分野があって、自分の専門誌の人たちが見て、その人たちだけで分かっていたらいいという時代は終わっています。他の人に見てもらって広めていかなければ駄目で、専門誌に出しても見られないのが現実ではないかと私は思います。

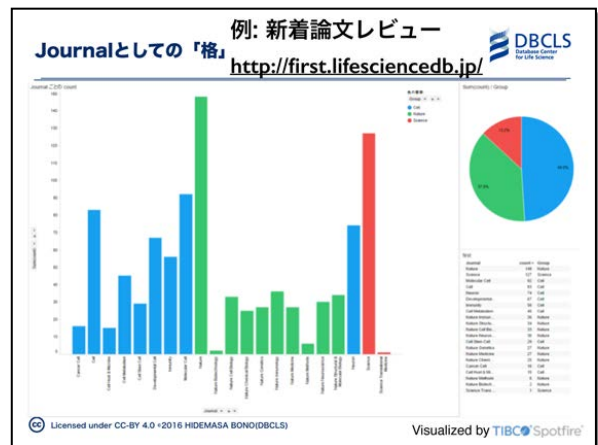
ネット検索に引っ掛かることは重要です。別に PubMed に入っていればそれでいいではないかといわれますが、PubMed は抄録部分、アブストラクトだけです。一方、オープンアクセスに出していると全文があるので、検索ワードがどこかにヒットすれば、興味を持って論文を見てもらえるので、やはり全文がインターネットで検索されるのが良い点ではないかと思います。

悪い点は APC の問題です。今は APC を自分の研究費から出していることがほとんどです。もちろん大学によっては手当てがあるのかもしれませんが、基本的には自分が取ってきた少ない研究費の中から出すのが普通です。また、オープンアクセス誌は長くても 10 年ぐらいですから、ジャーナルとして格がないのです。格がある雑誌とは何かというと、インパクトファクターで測る場合もありますが、「いい雑誌」とみんなが言うものです。

DBCLS では、新着論文レビューを行っています。これは、一般に生命科学者がプレステージジャーナル、いいジャーナルと言っている約 20 誌に掲載された論文について、著者自身にレビューを書いてもらい、誰でも自由に閲覧・利用できるよう Web 上にいち早く無料で公開するサービスです (図 3)。既に 6 年間やっていて、レビューは 1000 ぐらいあります。この対象になっている雑誌は飯田啓介という編集長が決めたものですが、これが生命科学で言うところのいい雑誌だという意見もあります。



(図 2)



(図 3)

## 論文の探し方の変化

そのような良い雑誌とは昔からある雑誌がほとんどですが、論文と一口に言っても、昔からある雑誌、最近になって出てきた雑誌というように、論文の探し方は私の中でもだいぶ変化しました。私は研究を始めて20年ぐらいですが、研究を始めた当時はまだPubMedが試用段階でした。そのころは図書館に来ている雑誌を見に行っていて、そのうちに電子メールが普及して電子メールでタイトルだけは送られてくるようになりました。

今でも私は古い人間なので、電子メールで来るサービスを利用してどんな新しい雑誌が出ているか確認しますが、最近の若い研究者はこんなことをせずに、自分の研究に関係ありそうなものをPubMedで定点観測して、差分を見てそれで合ったものを見ている、また、さまざまなサービスを使って、注目している論文を引用した論文を見ているのではないかと思います。

今、PubMedはエントリー数がものすごく多くなり、全文が入っているPubMed Centralも、PubMedの15%ぐらいの割合であるので、昔よりはだいぶ信頼できるリソースになったのではないかと考えています(図4)。

PubMedのデータは全て研究目的で使えるので、DBCLSはそれを使ってサービスを幾つかつくっています。一つは、PubMedの中に出てくる略語が何のことを意味しているのか検索できるAllieというサービスです(図5)。「APC」で検索すると、PubMedでは

antigen-presenting cells (抗原提示細胞) や adenomatous polyposis coli (大腸ポリポーシス) の略語としてよく使われていることが分かります。ちなみに、article processing charge の略語としての APC は19番目でした。だいぶ下の方です。

もう一つは、Colil という、論文の引用情報に関する検索サービスです(図6)。ある論文が他の論文から引用されているとき、本文中ではどのような文脈で引用されているかについて、効率良く検索します。PubMed ID またはキーワードで検索すると、論文がコンテンツスト、イントロダクションなどどこに引用されたかが出てきます。これは全文が使えるようになってきます。限定ですが、PubMed CentralにあるデータだけでこういうものをつくることもDBCLSで行っています。

The screenshot shows the 'Allie A Search Service for Abbreviation / Long Form' interface. The search term 'APC' has been entered, and the results are displayed in a table format. The table has columns for 'Category No.', 'Category Name', 'Main Research Field', 'Associated Terms', and 'PubMed/MEDLINE Title (Year, Volume)'. The results include: 1. antigen-presenting cells (Antigen Presentation) with terms like DC, MHC, TCR and a PubMed reference; 2. adenomatous polyposis coli (Neoplasms) with terms like CRC, LOH and a PubMed reference; 3. activated protein C (Hematology) with terms like EPCR, PC and a PubMed reference; 4. anaphase-promoting complex (Cell Biology) with terms like CDK, CBF, CDC27 and a PubMed reference.

(図5)

The screenshot shows the PubMed search interface. The search term 'ALL[filter]' has been entered, and the results are displayed in a table format. The table has columns for 'Database', 'Count', and 'Description'. The results include: Books (525,545), MeSH (265,382), NLM Catalog (1,551,414), PubMed (26,385,322), and PubMed Central (4,051,228).

(図4)

The screenshot shows the 'Colil by DBCLS' interface. The search term 'colil' has been entered, and the results are displayed in a table format. The table has columns for 'Category No.', 'Category Name', 'Main Research Field', 'Associated Terms', and 'PubMed/MEDLINE Title (Year, Volume)'. The results include: 1. Reconstruction of amino acid biosynthesis pathways from the complete genome sequence; 2. Metabolic-scale prediction of intermediate compounds in multiple metabolic pathways; 3. Modular architecture of metabolic pathways revealed by conserved sequences of reactions; 4. Quantified de novo reconstruction of metabolic pathways from metabolite-scale compound sets; 5. Endophytes in Trappesporales; 6. Dark genomes of Bifidobacterium sensu lato.

(図6)

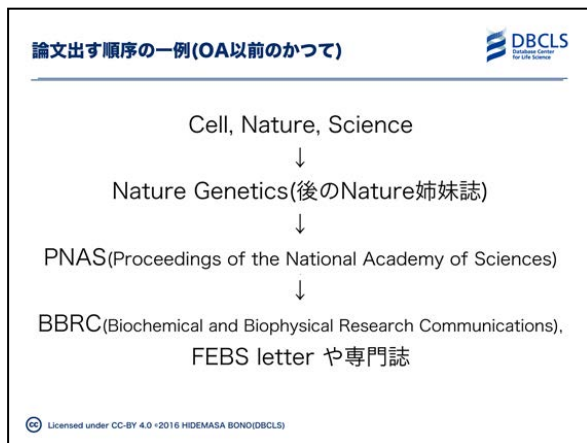
## 論文の投稿先、投稿順序

投稿先雑誌の選択趣向については、生命科学者はインパクトファクターの高いジャーナルからなっています。先ほどから議論を聞いていて「あれ？」と思ったのですが、研究者はジャーナルを選ぶ自由を持っていると思っています。投稿先は、インパクトファクターの高さと、自分がその雑誌に出したいという思い入れ、自分の研究している分野でよく見られている雑誌かどうかで決まると考えています。

論文はピアレビューがあるので、落とされることがままあります。ほとんどの場合は1回では通りません。何回も挑戦しなければいけないので、何種類か論文誌のリストを持っていて、それで投稿していきます。インパクトファクターが高いもので PubMed に入っていないものはないのですが、PubMed に収録されていることは必須です。PubMed に登録されていないジャーナルがありますが、それは制約条件になると思います。皆さんは就職するときなどに論文リストを書きますが、PubMed で検索して真偽を確かめることが多いので、ここは必須です。

分野によっては所属学会誌に投稿します。高エネルギー物理学だとそうなると思うのですが、私がやっている分子生物学などの分野は所属学会誌が一応ありますが、皆さんあまり出したがりません。こういう分野もあります。大体、分野の雑誌にはあまり出さないのではないかと考えています。

オープンアクセスをはやる前の論文を出す順序は、



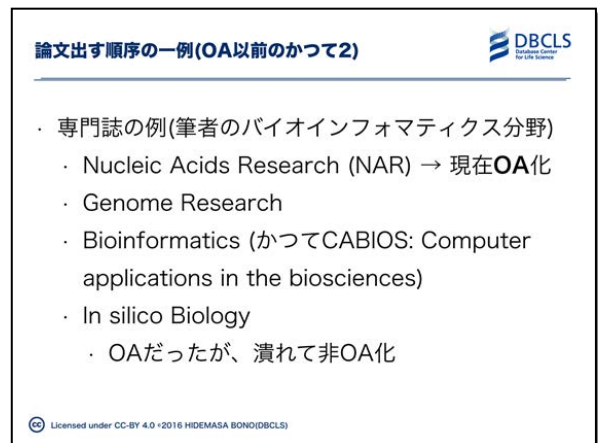
(図 7)

私が周りを見たとこ図7だったと思います。Cell、Nature、Science に出して駄目だったら、だんだん下に行き、最後の受け皿は速報誌や専門誌になるのではないかと思います。

私のバイオインフォマティクスという分野でも、専門誌、ゲノムの研究に特化した雑誌などに出していました(図8)。いろいろな道筋をたどっていて、Bioinformatics という雑誌は、昔は Computer applications in the biosciences (CABIOS) という変わった名前だったのですが、バイオインフォマティクスという言葉が1998年ぐらいからはやりだした途端に名前が変わって、いい雑誌になりました。In Silico Biology という雑誌は一度、オープンアクセスで誰でも全部読めるということを出していたのですがつぶれて、どこかの雑誌社に買われた途端に非OA化になってしまったという非常に悲しい歴史があります。私もここに2本ぐらい論文があるのですが、今は自分も読めなくなっています。

生命科学には Nature ラブな人が多いのですが、そういう人の場合、Nature に出して通らなくても、いいレスポンスがあれば Nature 姉妹誌に出します(図9)。太字になっているのはオープンアクセスの雑誌です。それでも駄目だったら、Nature Communications、Scientific Reports です。

私自身は Public Library of Science (PLOS) に出しています。マイクロアレイという、遺伝子の発現を測定する実験ツールを開発した人たちが、実験するときに、



(図 8)

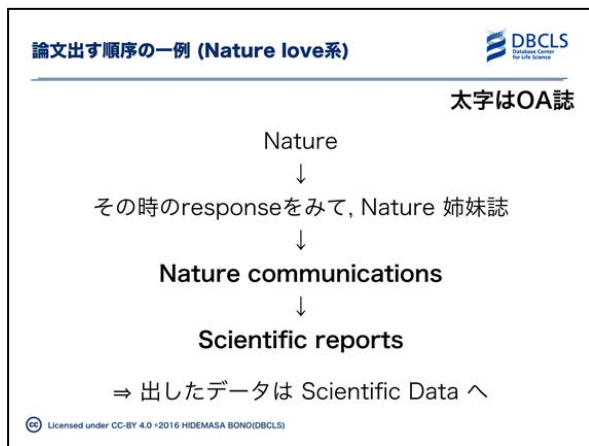
ゲノム全体から新しい遺伝子が発現した、ある遺伝子がすごく発現していたといったことを探し当てるのですが、その遺伝子に関して書いてある論文が読めないことがままありました。そこで、誰でも読めるものを始めようと言って PLOS を始めたのです。

私は PLOS の方が好きで、実際に論文を出すときは図 10 の道筋です。PLOS が一番推している古い機関誌は PLOS Biology で、それ以外の PLOS Genetics や PLOS Medicine など、割と中間のものもたどって PLOS ONE に行きます。いきなり PLOS ONE という場合もちろんありますが、やはりいい仕事をしたという場合は、皆さんはより上流からトライするようです。

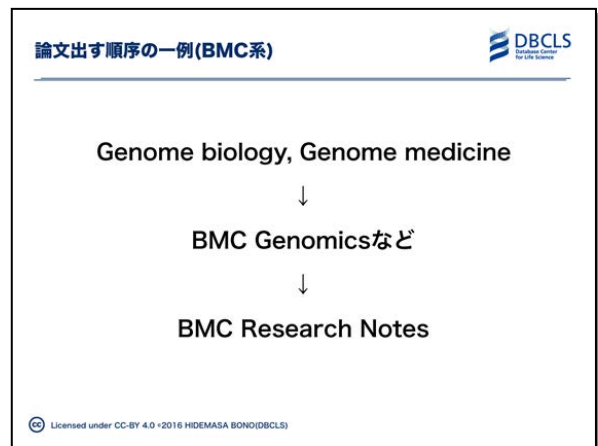
私は BMC Genomics のアソシエイトエディターをしているので BMC 系の場合も出しておきます (図 11)。BMC Genomics は多分中間層で、その前には Genome

Biology、Genome Medicine といういい雑誌があります。BMC Genomics など中間層で駄目だったら、最終的には BMC Research Notes に出してくださいと言われていますが、これも今後変わるそうです。どうやって変わるか見ものです。

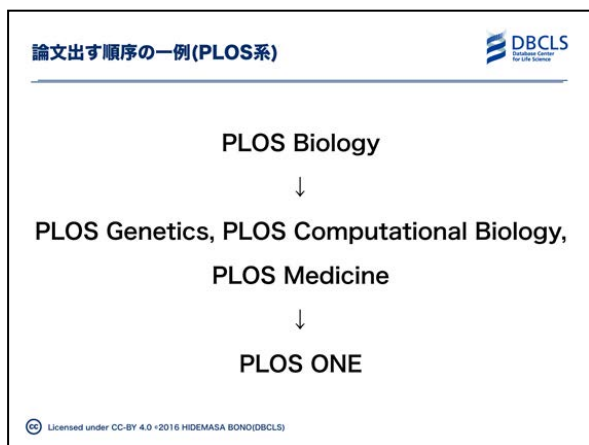
では、ジャーナルはどんなところに出したらいいのかとよく皆さんから質問を受けるので、それを解決するサービスが PubMed のテキストからできないかということをつくったのが Journal Suggestion using MedliNe (JSMN) です (図 12)。これは PubMed のアブストラクトとの類似度からどの雑誌に投稿するのがよいかサジェストするものです。今は動いていないのですが、かつて DBCLS が東大の中にあつたとき、統合牧場という学生アルバイト組織があり、その技術開発部が JSMN をつくりました。その中で生まれてきたのが統合 TV です。



(図 9)



(図 11)



(図 10)

どこのせせばいいのか?

- PubMedのAbstractとの類似度から suggest できないか?
- ⇒ JSMN: Journal Suggestion using MedliNe (ジャスミン)

JSMN  
Journal Suggestion using MedliNe

統合牧場 (2008/07-2014/02) DBCLS の学生アルバイト組織

- 技術開発部: 主に自然言語処理技術開発(上記JSMNなど)
- 統合TV制作部: <http://togotv.dbcls.jp/> 統合TV TOGO TV
- 美術制作部: Togo picture gallery <http://g86.dbcls.jp/togopic>

統合牧場  
Togo 2014

DBCLS  
Database Center  
for Life Science

(図 12)

今はこのようなサービスを雑誌社や査読を引き受ける会社がつくっていて、Edanz Journal Selector、Journal Guide、Elsevier Journal Finder などがあります。そこに自分の論文のタイトルとアブストラクトを貼り付けると、どんな雑誌がいかを出してくれますが、自分のところの雑誌ばかり出てなかなか面白いです。

まとめると、生命科学研究者はやはり Cell、Nature、Science に固執してしまっている感があります。それが無理なら姉妹誌、それが駄目なら PubMed に収録されるオープンアクセス誌にという傾向がかなりあると思います。

PubMed に収録されるオープンアクセス誌は、大学によらず全文が見てもらえるし、ネットでインデックスされるので検索したら出てくるようになっています。しかし、それを出すときは APC が掛かり、それは自分の研究費から出すことがほとんどなので、ますます減っている研究費からそれを捻出するのは非常に大変になっています。

### **グリーンオープンアクセスの余地は？**

受け皿をつくってオープンアクセスにすることは、やった方がいいと思います。私もデータを使う人間ですから、そういうものを蓄えることで、今後のテキストマイニングに使えるような材料となればと思います。

PubMed 収録、DOI 付与など、サイテーションの手段を強化して、誰がどのようにアクセスしているか、いつも統計を取って見えるようにすれば、インセンティブになって、皆さんもっとやる気が起きるのではないかと思います。しかし、研究者からしてみれば、一回論文を出したからもういいではないかという気持ちが強いです。論文を一回出してしまうと、続きの研究をされる場合もありますが、忘れてしまっている研究者も多いです。それは、今の論文が、掲載前にピアレビューで評価するという昔のシステムを引きずっているからではないかと思っています。評価は後から皆さんがするものだという意識がもっと広まってくると、変わっていくのではないのでしょうか。

そういう意味で、オルトメトリクス (altmetrics) のツールなどは非常に有効なのではないかと思っています。よく使うようにしていますが、それすら面倒ということやらない人も多いです。

そういうものは研究者サイドからやるのは無理で、ある程度トップダウンで強制になってしまうのではないかと思います。自発的な動きははっきり言って期待できません。アメリカ国立衛生研究所 (NIH) のように、「研究費をもらっているからには出せ」と強制にしなければ無理ではないかと考えています。その辺はいろいろな意見があると思いますが、この後で議論できたらと思います。

2016 年 10 月 29 日に、NII のすぐ近くで「国立遺伝学研究所 公開講演会 2016 出張!! 遺伝研」が行われます。生命科学者の生態が都内で見られますので、ぜひよろしくお願ひします。