



正しい知識が捏造を防ぐ

データを正確に解釈するための 6つのポイント

No. 3

客観的な判断のむずかしい事例をどう扱うか？ 定量化の方法と代表例の選び方を主題として

齋藤成昭

はじめに

恣意的にデータを操作してはいけない。そんなことをすれば、それは“捏造”である。サイエンティフィックに公表されるデータは公正でなければならず、可能なかぎり客観的でなければならない。それでは、どんな実験結果も客観的に解析することが可能かといえば、必ずしもそうではない。観察者の主観的な判断にたよって解析せざるをえないようなたぐいの実験は数多く存在する。そのような実験を行なうたびに、筆者は、いつも不安を感じている。“自分は公正に判断して解析したつもりだが、無意識のうちに、都合のよいように判断をゆがめてしまっていないだろうか？”という不安である。

ハードウェアやソフトウェアの進歩により、さまざまな実験データを機械的かつ客観的に解析することが可能になってきている。しかしながら、実際の研究の現場には、研究者の主観的な判断にたよらざるをえない場面が、まだまだ多く残されているように思う。研究者も人間である以上、主観にたよることが恣意的なデータ操作へとつながる可能性は、つねに存在する。このような問題をどう取り扱うべきかについて明確に示したルールは、筆者が知るかぎり存在しない。そこで、本稿では、客観化しづらいデータの解析について、筆者の見解を、実際の体験もまじえながら述

べてみたい。

さきにいいわけをするようで心苦しいが、無意識のうちにデータを操作してしまう危険性をどうすれば回避できるのか、いろいろと方策を練ってみてはいるが、完全な解決策はまだ得られていない。本稿をひとつの問題提起としてとらえて、この問題について考えるきっかけとしていただければ幸いである。

客観的な解析のむずかしいデータを どう取り扱うか？

筆者らは、酵母をモデルとして染色体分配メカニズムの研究を行なっている。研究の過程では、“ある変異株で生じる染色体分配異常のタイプを分類し、その出現頻度を定量化する”というような解析を頻繁に行なう。ここでいう染色体分配異常とは、染色体が過度に凝縮していたり、核分裂がおわるまえに細胞質分裂がはじまったり (*cut* 表現型という)、核が娘細胞へ不均等に分配されたり (*mis* 表現型という)、といったような、細胞生物学的な表現型のことである。解析方法は単純で、固定・染色した変異株の細胞を顕微鏡下でひとつひとつ観察して、その表現型を分類するのである (図1)。

この“分類”作業が、観察者の主観的な判断に大きく依存する。生物を材料とした研究を行なううえではさけられないことと思うが、それぞれの細胞には、たとえゲノムが完全に同じであろうとも、さまざまな確率的な要因の差によって“個性”が生じる。たとえば、“*cut* 表現型”としてひとまとめにされる集団のなかにも、実際には、まったく核分

Shigeaki Saitoh

久留米大学分子生命科学研究所 細胞工学研究部門

E-mail : saitoh@lsi.kurume-u.ac.jp

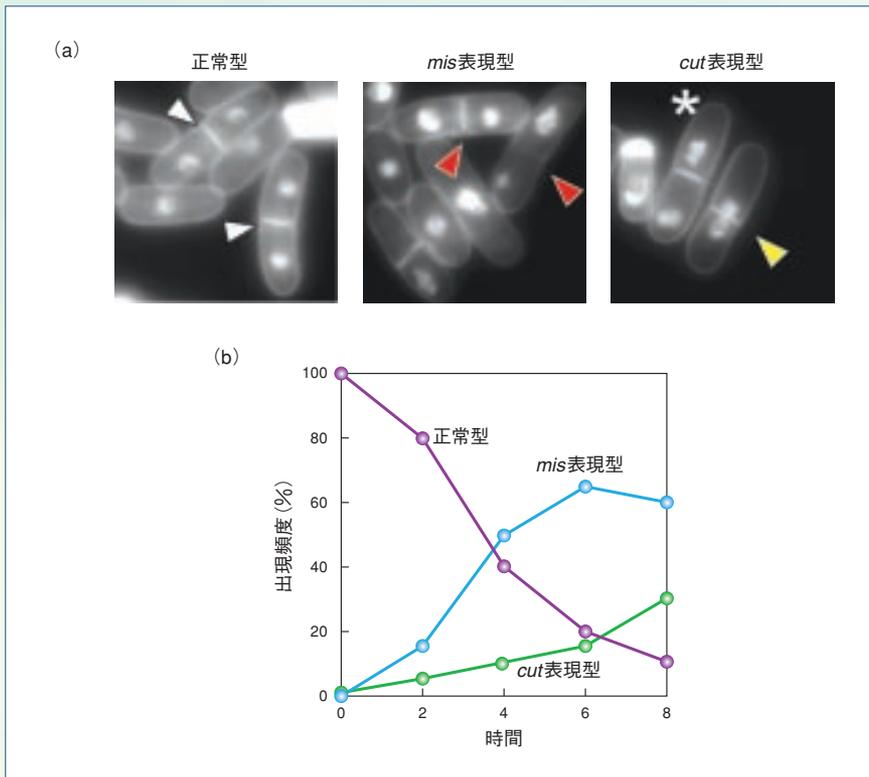


図1 架空の分裂酵母変異体Zの表現型の解析

(a) それぞれの表現型の“典型例”。DAPI染色により核と細胞壁を染色している。それぞれの表現型について矢じりで示した細胞(核分裂の直後)に注目すると、正常型と比べて、*mis*表現型の細胞では姉妹核の大きさが不均等となり、*cut*表現型の細胞では核が分裂前に細胞隔壁によって破断されている。★印は、分類がむずかしい細胞の例で、姉妹核の大きさがわずかに不均等な*mis*表現型であるようにもみえるし、正常型のようにもみえる。変形のために、核の大きさの差が明瞭でない。

(b) それぞれの表現型を示す細胞の出現頻度を、時間経過に対してプロットしたもの。染色した細胞を観察し、ひとつひとつの細胞がどの表現型であるか分類し、その出現頻度を求めた。分類がむずかしいような細胞をどう分類するかによって、結果は変化する。

裂がはじまらないままに細胞質分裂が起こってしまうものから、正常なタイミングよりわずかに早く細胞質分裂が起こってしまうものまで、さまざまな様相を示す細胞が含まれる。表現型の“スペクトラム”とでもいうべきものは連続的であり、したがって、どこからを“*cut*表現型”と分類し、どこからを別の表現型と分類するかの境界線は、観察者の判断によって決められる。さらに、既存のどのカテゴリーにもあてはまらない表現型を示す細胞が見つかった場合、その細胞のどの特徴に注目して新しくどのような分類カテゴリーをつくるかは、観察者の感性によるだろう。このような実験では、まったく同じサンプルを解析していても、観察者によってまったく結果が異なることも起こりうる。分類の基準について観察者どうしでよく話し合われていてコンセンサスが得られているとしても、やはり、結果には差がでてしまうことがある。そのような差を生み出す原因としては、観察者の経験というか、洞察力の違いが大きいように思う。“センス”といわれるものかもしれない。つまり、微妙な差異を嗅ぎ分けて、その違いを抽象化してカテゴライズする能力が備わっているかどうかで、結果はがらりと変わってしまう。このような事態は、結果に“再現性が乏しい”という点で好ましくはない。また、結果が観察者の感覚に依存することは、恣意的なデータ操作の余地を残し

ているともいえる。

このような問題を解消する手段のひとつに、紡錘体の長さや傾き、染色体の数といった、数量的なパラメーターで表わせる表現型に注目することがあげられるだろう(図2)。たとえば、“第1染色体を3本以上もつ細胞が集団中に何%含まれるか?”といった定量化ならば、観察者が代わってもほぼ同じ結果が得られるに違いない。判断基準も明確で、観察者の感覚には依存しない。あらゆる表現型に対し数量的なパラメーターがあるわけではないので万能の解決策ではないが、かなり有効な手段であると思う。ただし、場合によっては解析作業がかなり煩雑になってしまう。個々の細胞についていちいちパラメーターを正確に測って分類するのは、なかなか骨の折れる作業だ。紡錘体や染色体を可視化するための実験操作もくわわる。正直いえば、筆者もこのような解析を日常的にはあまり行ないたくない。人の手で解析を行なうかぎり、扱えるサンプルの数にも限界がある。ただし、この点に関しては、画像解析ソフトウェアをうまく応用すれば自動測定が可能になるかもしれない。作業を自動化できれば実験者の負担はかなり軽減するだろう。

数量的なパラメーターを指標にするとしても、やはり経験が“ものをいう”場合がある。筆者が大学院1年生のとき、“X変異株では染色体の不均等な分配が起こっている”とい

う可能性について検討したことがあった。FISH (蛍光 *in situ* ハイブリダイゼーション) 法で分裂酵母の第3染色体を可視化して観察したところ、対照となる正常株ではすべての細胞が第3染色体を保持していたのに対し、X変異株では30%程度の細胞が第3染色体を失っていた(ようにみえた)。この実験は、シグナルの数が0か1か、すなわち、みえるかみえないかで判断できる実験で、(少なくとも当時は)結論に疑う余地はないと考えた。そこで、自信満々に指導教授に報告したところ“よいデータだ”とほめられた。その1週間後、確認のため、今度は第2染色体のプローブを調製してFISHを行なった。“やはり30%の細胞が第2染色体を失っているに違いない”と期待して顕微鏡をのぞいたのだが、期待は裏切られた。正常株でもX変異株でも、ほとんどすべての細胞が第2染色体を保持していたのだ。悪い予感を感じつつ、もういちど第3染色体のFISHを行なってみたところ、前回とは違ってかわり、X変異株でもすべての細胞に第3染色体のシグナルがみえた。

変異株を取り違えていないかとか、培養条件を間違っていないかとか、さまざまな点を確認してみたが、なんの間違いもみあたらない。そこで、よくよく標本をみてみると、前回よりも蛍光シグナルがはるかに強く、しかも、均一であることに気づいた。何のことはない、前回の実験は失敗だったのである。なにかの手違いで、染色が全体的に弱く、しかも、不均一になってしまっており、その標本には“かろうじて染色された細胞”と“染色されていない細胞”が混ざっていたのだ。最初の実験のとき、標本のようにすがいつもと違うことに気づいて染色を失敗した可能性を考慮できていれば、もう少し慎重になっていたかもしれない。しかし、“変異株Xでは染色体が不均等に分配されているはず”という思い込みもあり、その可能性には思いがいたらなかった。

“あんなに自信たっぷりに報告したのに”と、顔から火の出る思いであった。怒られるのを覚悟しつつ、前回の結果が間違いであったことを指導教授に伝えたところ、“自分で間違いを正したのだからエライ”とほめられた(なぐさめられた)。

筆者がこの件から得た教訓は、“機械的にサンプルを観察しているだけではだめで、顕微鏡をのぞきながら、さまざまな要因を考慮できるようにならないといけない”ということであった。なんらかの理由で実験が失敗している可能性だってある。その可能性について思いがいたらなければ、アーティファクトにだまされることになる。なにがアーティファクトでなにがアーティファクトでないかを見極められるようになるには、やはり、ある程度の訓練と経験(とくに、失敗の経験)とが必要だと思う。なお、この一件以降、筆者は少し疑り深くなったかもしれない。“予想どおりの結果がでた場合には、まず疑う”ことにしている。

細胞の表現型を、感覚的に分類するにしても、なにかのパラメーターを指標にして分類するにしても、信頼できる結果が得られるようになるには、ある程度の経験と訓練とが必要である。しかし、経験豊かだからといって過信してしまうのもよくない。どんなに訓練をつんでも、思い込みのせいでデータがゆがんでしまうことは起こりうる。ここで例にあげたような表現型の頻度の測定にかぎらず、観察者の主観的な判断が結果に大きく反映されるような測定は、なるべく、複数の人間がチームを組んで何度もくり返すようにしたほうがよいと思う。経験の浅い者は経験の豊かな者とチームを組むようにしたほうがよい。そうすれば、“ひとりの思い込みや経験不足のためデータがゆがんでしまう”という事態を、ある程度までさけることができる。複数の測定値を統計的に処理することで、測定の“ぶれ”を正す

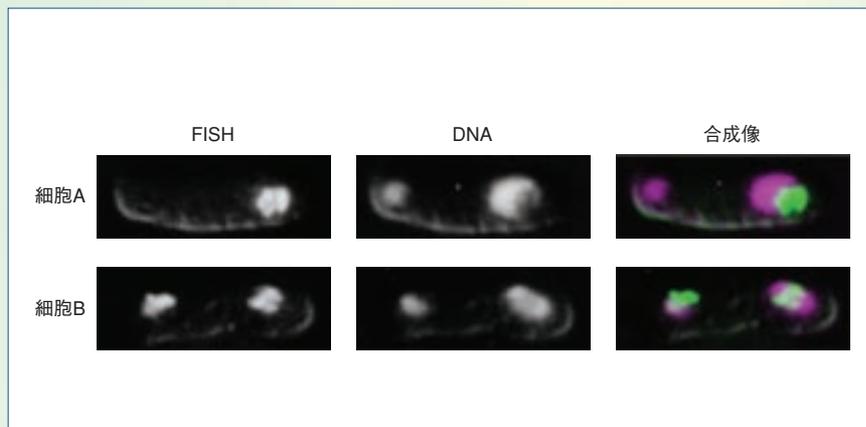


図2 数量的なパラメーターを指標にした表現型の解析

FISH法を利用した染色体の不均等分配の検出の例。特異的なプローブを用いたFISH法により、核に第3染色体が存在する(=シグナルが1つ以上)か、存在しない(=シグナルが0)かを判定することができる。核分裂の直後に、姉妹核の片方にしかシグナルがなければ第3染色体が“不均等”に分配されたことを示し(細胞A)、シグナルが両方であれば“均等”に分配されたことを示す(細胞B)。染色がきちんと行なえてさえいれば、誰が観察しても、ほぼ同じ精度で“不均等分配”表現型の発生頻度を測定することができる。ただし、この場合、第3染色体以外の染色体がいかに分配されたかまではわからない。

ことも可能であるし、異常な測定値を検出することも可能である。もちろん、複数の人間が観察を行なう場合、それぞれの判断基準が変わらないよう十分に話し合っておくことも重要である。

このような測定を行なう者は、つねに“無意識のうちに結果をゆがませてしまう”可能性を意識し、行なっている実験の“危うさ”をはっきりと理解しておくべきだと思う。きわめて慎重になるべきだ。そして、可能ならば、まったく別の実験手法を使って、何重にも結論を確認しておくべきだろう。

知らないうちに誤った定量化を行なってしまう危険性

少し話がそれるかもしれないが、定量化ということに関していうと、用いている測定法の弱点と限界を理解しておくことも大切であろう。それらに注意をはらわず実験を行なっていると、知らず知らずのうちに大きく誤った結果をだしてしまうかもしれない。

たとえば、細胞内のある蛋白質の存在量を調べるような実験を考えてみる。いろいろな測定方法が考えられるが、筆者らがよく用いるのは、ウェスタンブロットで目的の蛋白質を特異的に検出し、その電気化学発光をX線フィルムに焼きつけて、検出されたバンドの濃さを測定する方法である。この場合、バンドの濃さを測るのは機械であり、つねに客観的で公正な測定値が得られるような気がする。しかし、ここで本当に知りたいのは、バンドの濃さそのものではない。蛋白質の量である。目的とする蛋白質の量がX線フィルム上のバンドの濃さとして直線的に反映されている保証がないかぎり、バンドの濃さが測定できても蛋白質の量を測ったことにはならない。フィルムカメラで写真を撮影したことがある人ならわかると思うが、フィルムに応じた適正な露光量を守らないと、フィルム上のコントラストはすぐに落ちてしまう。つまり、被写体の明るさがフィルム上に正しく反映されなくなってしまう。同じように、X線フィルムの場合でも、露光量(=発光の強さ×露光時間)を間違えると、バンドの濃さは蛋白質の量を直線的に反映しなくなる。

どのX線フィルムにもとづいて測定を行なうかは実験者の裁量によるが、露光量を間違ったX線フィルムをもとにして蛋白質の量を測っても正しい測定値は得られない。この測定方法の弱点はここにある。この場合、弱点を克服す

るためには、測定に用いようとするX線フィルムの露光量が適正であることを確認すればよい。具体的には、段階希釈したサンプルを用いて検量線を描き、それが直線であり、そして、測定の対象となるバンドの濃さがその検量線上にあることを確認するのである。例にあげたような測定を行なう場合には必ず、検量線を描くための対照をおかないといけないのだが、意外に無頓着な研究者が多いようだ。ときおり、露出オーバー気味の真っ黒なバンドの横に濃度の測定値が記載された図をみかける。きちんと対照がおかっていたのか疑問に思うことがある。

この例にかぎらず、どのような測定方法にも弱点や限界がある。また、どんなに精度の高い測定機械を用いても、その使い方を誤れば正しい測定値が得られるわけがない。論文などに示される結果は、その実験が正しく行なわれていることを前提としている。間違った方法で測定して得られた誤った測定値を発表することは、“不正”とまでいえないまでも、好ましいことではない。このようなエラーをさけるためには、日ごろから、自分が使っている測定方法の原理をよく理解しておくことが重要である。

“典型例”“代表例”とは何か？

論文に掲載したり学会で発表したりする“代表例”“典型例”の選び方にも、“主観的な要素”が大きくかかわっている。たとえば、ある表現型を示す細胞の出現頻度を表わしたグラフの横に、その表現型を示す細胞の写真が掲載されているような図をよくみるだろう(図1a)。そして、“a representative image is shown...”というような常套句がそえられていることが多い。筆者も、よくそのような図を作成する。しかし、そこに掲載されている写真は、何を“代表”しているのだろうか？言葉のとおりうけとめれば、もっともありふれた“平均”的な写真が選ばれているはずだと思う。しかし、実際には、その論文の主張を後押しするような、“誇張”をちょっと含むような、写真が掲載されていることが多いのではないだろうか(図3)。何枚もある生データの写真のなかからどの写真を掲載するかを選ぶのは、完全に論文著者の判断による。もちろん、捏造された写真を掲載することは論外である。しかし、正しく行なった実験の結果として得られた写真を掲載するのであれば、それが極端な例であっても、一概に不正とはいえないのではないだろうか。いわゆる“チャンピオンデータ”を選りすぐって論文に掲載したり学会で発表したりすることは、とがめら

れることなのだろうか？ この点に関してどう考えるべきなのか、筆者にも、まだよくわからない。

多くの場合、“平均”的な写真を提示すると主張したい事柄が不明瞭になる。つまり、中途半端なのだ。“何をいいたいのか、どこに注目しているのか”を明確に伝えることができず、読者や聴衆にあいまいな印象をあたえてしまう。あいまいであるがゆえに、ミスリードしてしまうことだってあるだろう。実際に、筆者には、論文の初稿であいまいな写真を掲載したためレビューアを勘違いさせてしまい、その改訂にひどく苦勞した経験がある。同じような経験をした人も多いのではないだろうか。それならば、多少の誇張があるにしても、主張したい点をクリアに示した例を“代表”として示すほうがよいのではないだろうか。筆者も、いろいろ悩んだすえにそうすることが多い。

細胞の表現型の観察の話に戻すと、さきに述べたように、あるひとつのタイプの表現型に分類されている細胞でも、実際には、そのひとつひとつが異なっている。この場合、その表現型を示す細胞の典型例を示そうとしても、すべてが違うのだから本当は典型例など選びようがない。結局、その表現型の特徴をもっともクリアに示した細胞の写真を“典型例”として選ぶことになる。このような“典型例”“代表例”を選ぶことは、事実をねじ曲げようとするのでなければ許されると思うのだが、みなさんはどう考えるだろうか？ ただし、本来、きわめて低い確率でしか現われない

表現型の写真を“代表例”としてひとつ示して、あたかもすべての細胞がそのようにみえるかのような記述をするのは欺瞞であり、完全に不正である。仮に、明記していないにしても、そのような印象をあたえてしまうこと自体、好ましいとは思えない。故意ではないにしても、自説を強調するため、あまりにも極端な例を提示してしまい、結果的に誤った印象をあたえてしまうことはあるだろう。あいまいすぎてもいけないし、強調しすぎてもいけない。このあたりのさじ加減というかバランスというかが、非常にむずかしいと感じる。

再現性や出現頻度の低い表現型をどのように扱うべきかについては、なかなか判断がむずかしい。そういったデータは公表する十分な価値がないものとして扱うことも、ひとつのやり方だと思う。“間違わない”ということを重視すれば、もっとも安全なやり方であろう。しかしながら、めったに現われない表現型のなかにブレークスルーとなるようなきわめて重要な発見がかくれているのはよくあることだし、再現性の低い結果でも、それは単に、実験者に結果を再現できるだけの技量が備わっていないからであるかもしれない。じつは、その結果が真実であることも十分に考えられる。筆者は、そういったデータが仮説やモデルを立てるうえでキーポイントとなるのであれば、公表しても構わないのではないかと思っている。もちろん、それらの再現性がとりにくいことや、めったに現われない表現型であることを

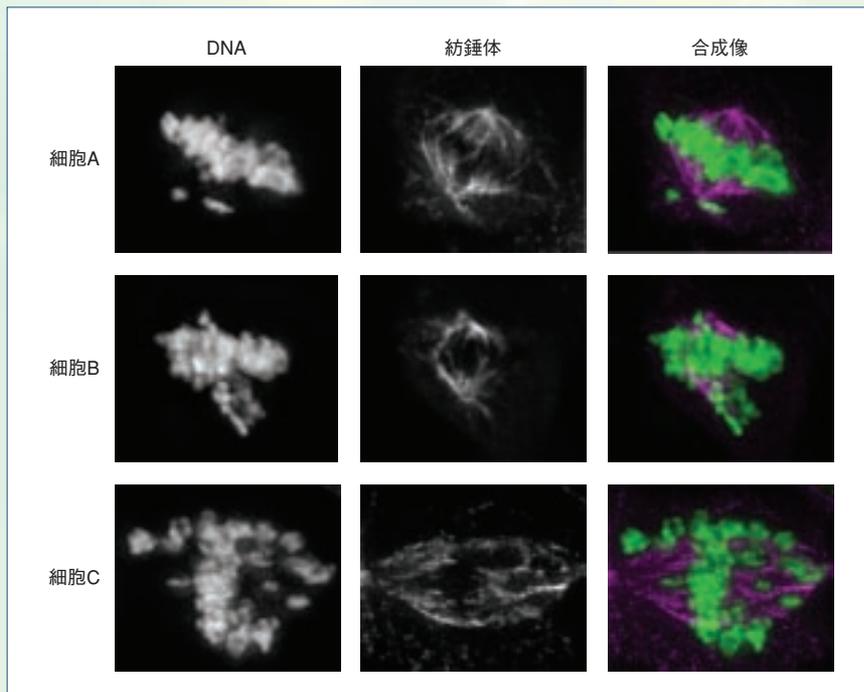


図3 どの写真を“代表例”として選ぶか？

低濃度の微小管重合阻害剤で処理した培養細胞のM期の像。DNAと紡錘体を染色している。細胞A、細胞B、細胞Cのいずれも、“紡錘体の中央に整列できない染色体が存在する”という共通の表現型を示している。細胞Bや細胞Cならその異常性は一目瞭然だが、実際には、細胞Aのように比較的軽微な異常しか示さない細胞も多い。“代表例”“典型例”としてどの写真が示されるかによって、うける印象は大きく異なる。

正直に述べることは大前提だ。自信のない結果は、自信のないことを正直に伝えればよい。そのうえで、それらのデータの信憑性や価値を判断するのは、レビュアー、あるいは、読者や聴衆の役割であろう。再現性や出現頻度に関して正しい情報が提供されているのであれば、仮に、あとからそれらが間違いであったことが判明したとしても、それらの公表自体が“不正行為”になるとは思わない。どんなに優秀な科学者であっても間違えることはある。間違っても、あとでそれをきちんと訂正すれば、それでよいと思う。

●おわりに

科学者は、ただ実験を行なって結果を得るだけでよいわけではない。いくつもの実験結果を解釈し、独自の“ストーリー”をつくりあげる。そして、そのストーリーを公表し、読者や聴衆からの批評をうける。公表するからには、誰しも自分の“ストーリー”を高く評価してほしいものだ。少しでもインパクトファクターの高い雑誌に論文を掲載したいし、学会会場の聴衆からは拍手喝采をあげたい。全員とまではいわないが、多くの科学者がそういう“願望”をもっていると思うし、その願望が科学を続けるためのモチベーションのひとつとなることは健全なことだと思う。

ひとつの“ストーリー”をつくりあげるためには、無数の実験方法のなかから必要なものを選びだし、多くの実験デ

ータから重要と思われるものを取捨選択し、そして、選択したデータをロジックにしたがって並べていく、といった一連の作業が必要となる。それぞれの段階で、科学者個人の“主観的な判断”が求められる。そういった選択作業は、単に客観的でないという理由だけでは否定されるべきではないだろう。同じ研究テーマでも、データの選び方や並べ方によってできあがるストーリーは千変万化し、おもしろくもつまらなくもなってしまう。そういった一連の“主観的な選択”こそ、科学者の“腕のみせどころ”だと思う。

しかしながら、“おもしろいストーリーにしたい”という願望が強すぎ、データを改ざんしたり隠蔽したりして嘘をついてしまうと、そのとたん、科学ではなくなってしまう。嘘は、それがいかに巧妙であろうとも、必ずばれる。科学者が科学者でありつづけるためには、つねに高い倫理感を保ちつづけなければならない。そして、“自分は何か不正を犯してしまっていないだろうか”と、いつもすこし臆病になって自戒しつづけることが大切だと思う。

齋藤成昭

略歴：1999年 京都大学大学院理学研究科博士課程 修了、同年 米国 Scripps 研究所 Research Associate、2002年 久留米大学 分子生命科学研究所 助手、2005年 同 講師を経て、2009年より 同 准教授。