

DUE

DILIGENCE

ASBMB TODAY

THE MEMBER MAGAZINE OF THE AMERICAN SOCIETY
FOR BIOCHEMISTRY AND MOLECULAR BIOLOGY



DUE DILIGENCE

Kaoru Sakabe is the data integrity manager for ASBMB publications: *Journal of Biological Chemistry*, *Molecular & Cellular Proteomics*, and the *Journal of Lipid Research*.

OFFICERS

Natalie Ahn
President

Gerald Hart
President Elect

Jennifer DuBois
Secretary

Toni M. Antalis
Treasurer

ASBMB TODAY EDITORIAL ADVISORY BOARD

Rajini Rao
Chair

Charles Brenner
Past Chair

Ana Maria Barral
Floyd "Ski" Chilton

Henrik Dohlman

Peter J. Kennelly

Beronda Montgomery

A. Maureen Rouhi

Melissa Vaught

Binks W. Wattenberg

COUNCIL MEMBERS

Squire J. Booker

Victoria J. DeRose

Wayne Fairbrother

Rachel Green

Blake Hill

Susan Marqusee

Celia A. Shiffer

Takita Felder Sumter

JoAnn Trejo

ASBMB TODAY

Angela Hopp
Executive Editor,
ahopp@asbmb.org

Comfort Dorn
Managing Editor,
cdorn@asbmb.org

John Arnst
Science Writer,
jarnst@asbmb.org

Valery Masterson
Designer,
vmasterson@asbmb.org

Ed Marklin
Web Editor,
emarklin@asbmb.org

Allison Frick
Media Specialist,
africk@asbmb.org

Barbara Gordon
Executive Director,
bgordon@asbmb.org

For information on advertising, contact Pharmaceutical
Media Inc. at 212-904-0374 or mperlowitz@pminy.com.



www.asbmb.org/asbmbtoday

PRINT ISSN 2372-0409

Articles published in ASBMB Today reflect solely the authors' views and not the official positions of the American Society for Biochemistry and Molecular Biology or the institutions with which the authors are affiliated. Mentions of products or services are not endorsements.

DUE DILIGENCE

2

A FIGURE WORTH 1,000 WORDS

ASBMB Data Integrity Manager Kaoru Sakabe gives an overview of how to best prepare figures for publication.

5

THE MYTH OF PERFECTION

Always ask yourself, "Is this figure a true representation of the original capture or image?"

6

PIXEL PERFECT

Thinking about figure resolution from the start prevents problems later on.

8

FOCUS ON EXPOSURE

Make sure you nail an image's acquisition parameters.

9

COMBATING COMPRESSION

JPEG files can compromise your hard-earned data.

11

(POWER)POINT (NOT) TAKEN

Don't blur the lines with this go-to software.

12

PLAY BY THE RULES — AND SPOT THE ERROR

What's wrong with these figures?

14

KEEP IT SAFE

Protect the data you've labored to gather.

15

Q&A WITH ORI'S KATHY PARTIN

A discussion about misconduct and questionable lab practices.

A figure worth 1,000 words

By Kaoru Sakabe

As the data integrity manager for the American Society for Biochemistry and Molecular Biology, I wear two hats. I investigate manuscripts submitted to and published in ASBMB journals for violations of ASBMB policies on publication ethics. (The ASBMB publishes the *Journal of Biological Chemistry*, the *Journal of Lipid Research* and *Molecular & Cellular Proteomics*.) I also educate authors regarding ethical issues in publishing and how best to handle them. The first role, albeit necessary, can be a roller coaster. I've



heard a lot of different excuses from authors I've investigated for violations, such as erasing blemishes and bands, reusing data from different publica-

tions and cutting and pasting bands to create data that never existed. These excuses run the gamut from somewhat credible to incredible — although I haven't yet heard that someone's dog ate it.

Educating authors about ethics is vitally important. I realize that not everybody has the exposure I had as a Ph.D. student. My mentor instilled zero tolerance for misconduct in all of his trainees. There was also a great culture in the lab of sharing best practices for data presentation.

As a publisher, the ASBMB can

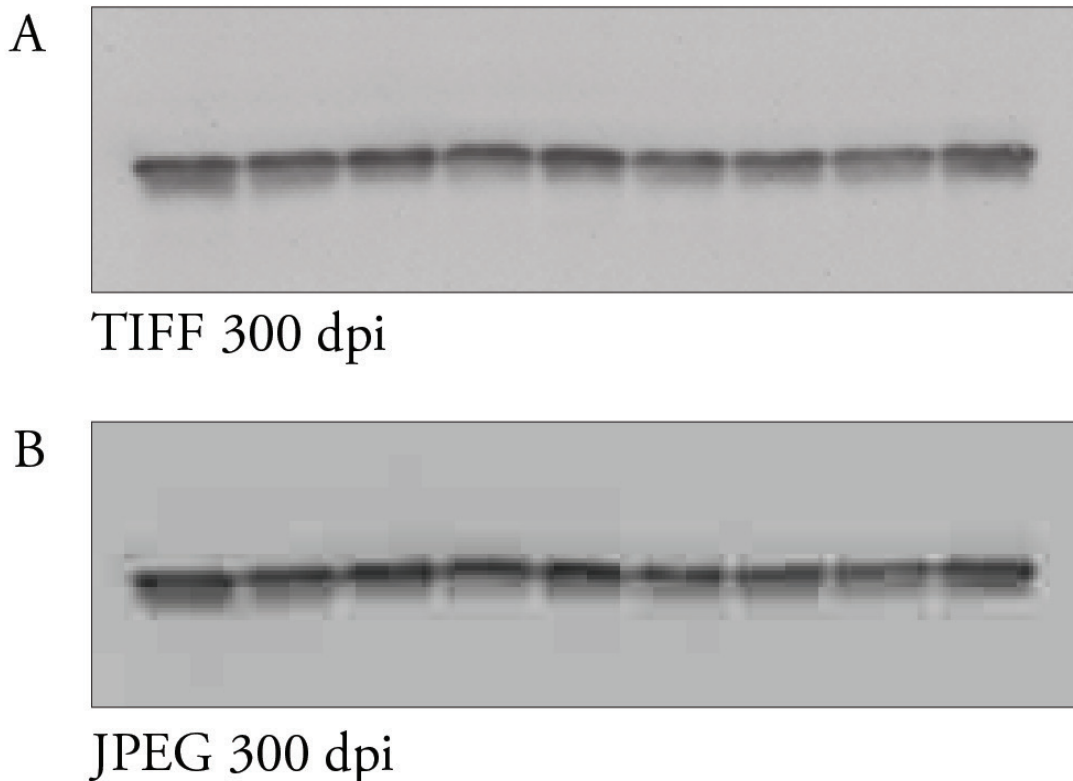


Figure 1: Images should be saved in TIFF format. The same image was saved at the same resolution of 300 dpi, but A was saved in TIFF format, while B was saved as a JPEG. Note the pixelation in B.

help fill that gap for authors unfamiliar with these practices in figure presentation, since everyone may not have had this kind of exposure as a student or postdoctoral researcher. Learning best data-presentation practices doesn't end with your formal training, though. I'm still learning, especially as publishing standards continue to evolve.

Over the next few months, I will be writing a series for ASBMB Today in which I will tackle different topics regarding images and figures and delve into ethical issues.

For now, I'll start with the basics — how to best prepare manuscript figures for submission. A manuscript is like a picture book that tells a narrative (your research) with the aid of some pictures (your figures). In telling your story, you need to present the pictures in a clear manner so that reviewers and, eventually, readers will be able to understand and interpret your data. Here are a few pointers:

Read the instructions for authors

This may seem like a no-brainer,

but you always should read the instructions to authors for the journal to which you plan to submit. The instructions contain valuable information about what the journal expects. This way, you avoid the frustration of having your manuscript sent back for formatting issues or because a reviewer can't make out a blot.

Figure preparation begins at data acquisition

Preparing publication-quality figures begins during data acquisition, long before you have a story, much less know where to submit your work. Whether it's scanning a film or taking a picture, overexposing or underexposing an image leads to loss of the fine details in the data. How can you tell your image is over- or underexposed? Take a look at the histogram. The histogram graphically displays the tonal distribution of an image by showing the number of pixels that are black, white and all the different shades of grey in between. Ideally, the pixels should be distributed throughout the range and not clustered at either end

of the spectrum. While it is tempting to acquire a clean-looking image with no background or speckles, reviewers know what real data look like. Additionally, the images should be acquired at a minimum resolution of 300 dots-per-inch.

Save images using loss-less compression

Scientific images should be saved in the TIFF format, because it uses a loss-less compression algorithm to save your data. Avoid the JPEG format because it uses an algorithm that results in loss of data (lossy compression). Lossy-compression algorithms approximate the original data, which can result in parts of your data being discarded. Although saving an image as a JPEG may save you computer disk space, the problems that this compression method may introduce, by essentially throwing out information, are not worth the benefit of more disk space or faster upload time (Figure 1).

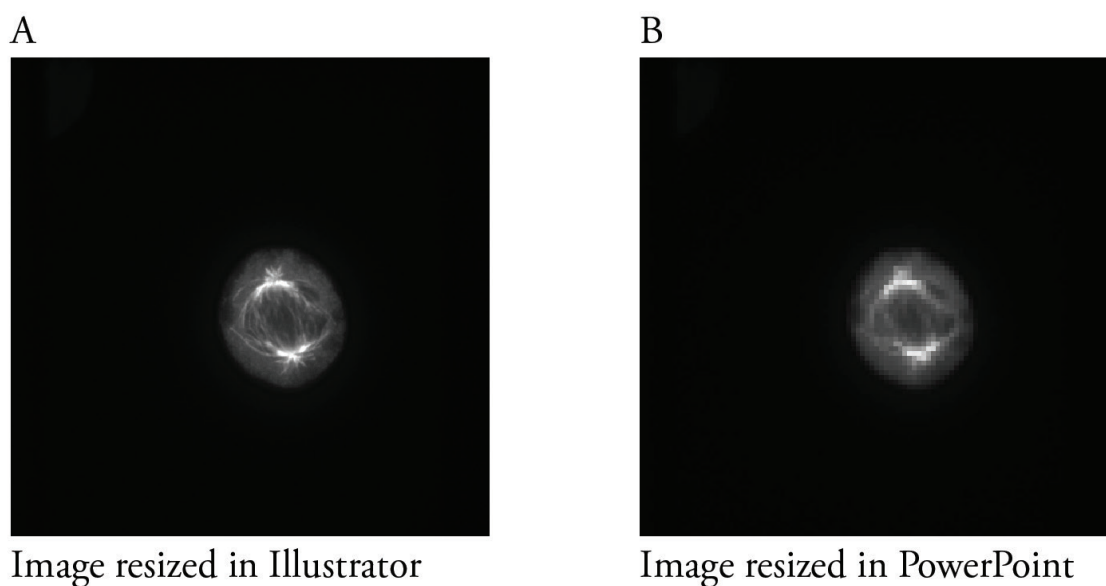
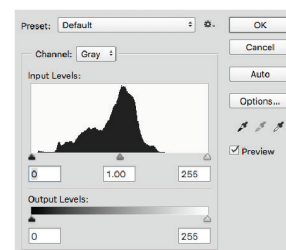


Figure 2: Figures should be created using appropriate software. The same image was resized, but A was resized in Adobe Illustrator, while B was resized in PowerPoint. Note the pixelation of the image in B. A free alternative to Adobe Illustrator is Inkscape.

A



B

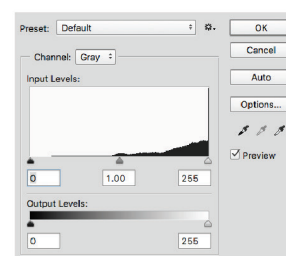
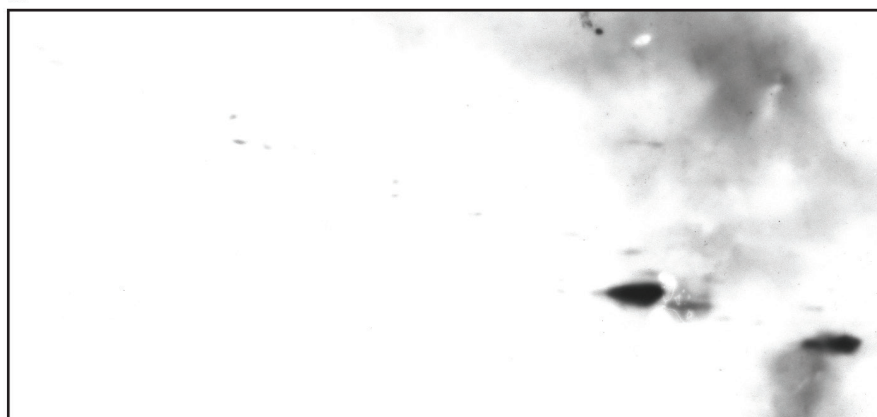


Figure 3: Avoid excessive manipulation. The original unmanipulated scan is shown in A along with the accompanying histogram. In B, the brightness and contrast were adjusted excessively. Note the absence of background, the disappearance of some spots and the shift in the accompanying histogram.

Prepare figures using appropriate software

PowerPoint is an attractive option for generating your figures, but avoid PowerPoint. The reason is that PowerPoint is designed for an onscreen resolution of 72 dpi and not print, which requires at least 300 dpi. Resizing images using PowerPoint can lead to loss of data, since it applies a lossy compression (Figure 2). Adobe Illustrator and Inkscape are good options for preparing figures.

Avoid excessive manipulation

This topic will be covered in more depth in future articles, but, in brief,

you should manipulate your image as little as possible when preparing the figures for publication. Your final image should be a true representation of the film or image when you captured the original. Aggressively contrasting your image or adjusting the levels to reduce the background may draw questions from reviewers and readers. Again, take a look at the histogram to make sure you are staying within acceptable limits. That pesky band or spot that you find troubling actually may be very informative for readers. It could indicate the performance of a certain antibody, or it could be a differentially modified form of your protein of interest (Figure 3). Importantly, those bands or spots are the actual data! Hiding or omitting them misrepresents your experimental

results to the reader.

Check your figures by printing them

It's a good idea to print out your figures before submitting them. If you have a hard time viewing your images, chances are so will the reviewers.

Submit!

And try to relax until the reviews come in.



Kaoru Sakabe (ksakabe@asbmb.org) is the data integrity manager at the ASBMB.

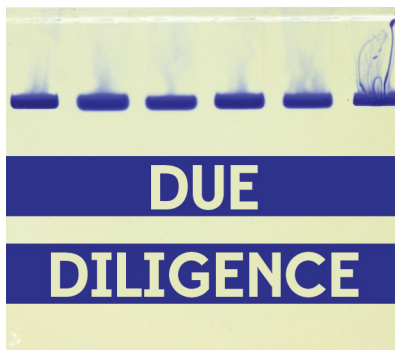
The myth of perfection

By Kaoru Sakabe

With the release of the imaging software Adobe Photoshop in the 1990s, “Photoshop-ping” entered the English lexicon. Like Google, Photoshop seamlessly has integrated itself into the scientific enterprise. Scientists use the software to tweak images and to generate publication-quality figures. It’s just so easy to create a blemish-free image. But there are guidelines to what is and isn’t acceptable to do with the software. There are a few simple rules to remember.

First, ask yourself whether any changes are needed. The best-case scenario is to be able to present your original, unaltered data in the figure. However, journal editors realize that sometimes the best case isn’t possible — an overly dark H&E stain or an overly bright Coomassie stain of a gel are two examples.

Once you’ve decided it’s appropriate and necessary to make changes, make sure your adjustments are linear. Most journals, including the journals published by the American Society for Biochemistry and Molecular Biology, require that adjustments be made uniformly to every pixel in the entire image. That means using the brightness and contrast functions in Photoshop is acceptable within reason,



since these functions apply a linear adjustment to each pixel in the image. Also, go easy on moving the slider (see the figure). Overadjusting the brightness or contrast can hide background features, which is a misrepresentation of your data. Nonlinear adjustments include adjusting the gamma settings or using the “Curves” function in Photoshop. These actions are discouraged, since they do not apply changes equally to the pixels in the image. If these adjustments are used, then you must disclose their use in the figure legend.

Speaking of data misrepresentation, specifically enhancing, removing or obscuring features would fall into this category. Worried that a faint band won’t support your conclusions? Bothered by the cell debris in the corner of your image? Concerned that the reviewers may say that the co-localiza-

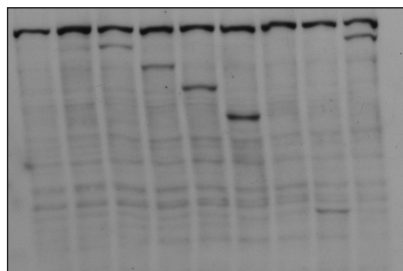
tion or the co-immunoprecipitation isn’t strong enough? The temptation to enhance or remove these features is real, but this type of manipulation falls into the misconduct category and could have serious consequences.

The final image should look like your original data, warts and all. You always should inspect your final figure and ask yourself if it is a true representation of the original capture or image. If your answer is no (or kind of), you should re-evaluate your figure.

Practically speaking, if any of these issues are discovered during the review of your paper or even after it is published, they could delay publication of your article, result in a correction, or even end in a retraction. More importantly, these issues go deeper and speak about the reproducibility of the work and your integrity as a scientist. Other researchers will not be able to replicate the results shown in your article if some of the data have been enhanced or hidden selectively. Presenting your data in a transparent manner ensures that you have done your due diligence.



Kaoru Sakabe (ksakabe@asbmb.org) is the data integrity manager at the ASBMB.



Original data



Adjusted data

Aggressively overadjusting the brightness and/or contrast misrepresents the actual data that were obtained and can mask potential biologically relevant results.

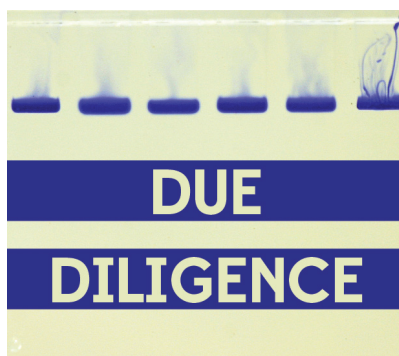
Pixel perfect

By Kaoru Sakabe

When the results started rolling in and a story began emerging, my thesis adviser usually instructed us to start assembling figures for a paper. At first, it was daunting to see blank figure panels nestled between the data we had. But the process made it easier to identify missing experiments and to see the logical progression of the story as the holes started to fill in.

With so much attention focused on building a scientific argument, the last thing on my mind while assembling figures was the final figure resolution. However, forgetting to keep resolution in mind from the start can cause problems later on. To avoid any potential issues down the road, I offer a few tips.

Let's start off with some basics. When reading submission guidelines for journals, they often throw around terms, such as "minimum resolution," "dpi," "ppi" and "vector graphics," which all seem irrelevant when you are eager to write up your manuscript. So what is a pixel, the first "p" in "ppi"? A pixel, derived from "picture element," refers to the most basic unit composing an image. Each pixel contains information telling the computer what color or shade of gray to display.



The film you have scanned, the immunofluorescent image you've snapped or the Western blot image you've exported from an imaging system — that image is composed of many pixels

arranged in an x, y grid such that the final image will show coimmunoprecipitation of your protein of interest or mislocalization of your protein upon treatment with an inhibitor.

The resolution of the image refers to the density of pixels. It is the number of pixels that make up your image. The greater the resolution, the more information an image contains and the clearer your image will be.

This quantity is expressed as pixels per inch, or ppi. For publication purposes, most journals will require that you submit your final figures with a minimum resolution of 300 ppi. You often

will see dots per inch, or dpi, used interchangeably with ppi, but dpi actually refers to printer output, or how many dots of ink are found per inch of a printed document. Since we are talking about digital images, ppi is the more relevant term to use.

The last bit of information you need to know is that your image data can be either raster or vector data (Figure 1). Raster data is simply an image made up using pixels as building blocks as discussed above. Vector data, on the other hand, is not composed of pixels but rather is a set of instructions that tells the computer to display lines and curves. This type of data is useful for graphs or models, since it remains smooth no matter how much you zoom in. Conversely, raster data

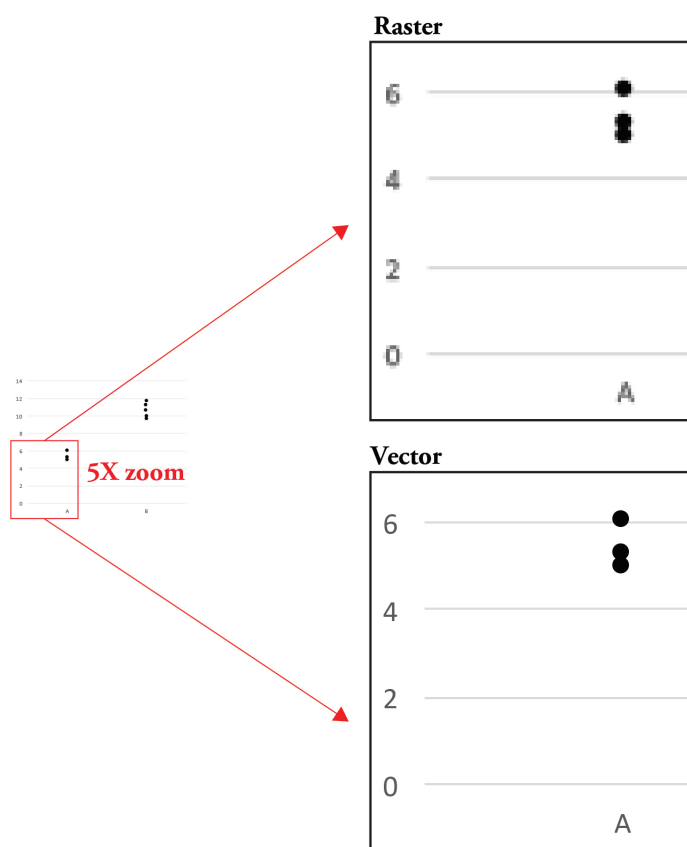


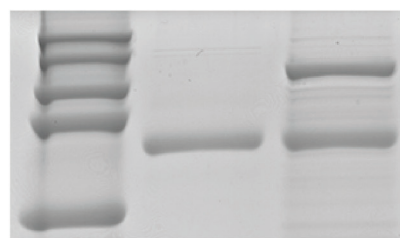
Figure 1. Raster data vs. vector data. Raster data becomes pixelated as you zoom in, whereas vector data remains clear.



10 ppi



72 ppi



300 ppi

Figure 2. Resolution matters! Even if you scan your film at 300 ppi, if you use PowerPoint, you effectively are changing it to a 72 ppi image.

becomes pixelated as you enlarge the image, making the gridlike pattern of pixels obvious.

OK, we've got the basics. Now how do you apply this information to create awesome figures?

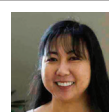
Tip 1: Remember that figure preparation begins at data acquisition. Make sure you are acquiring your image at the proper resolution. Whether you are scanning a film or exporting a file from an imaging system, keep the minimum resolution of 300 ppi in mind. It's never fun when you realize that you have to find a particular film to rescan at the proper resolution months (or even years) after you performed the experiment, or worse, conduct the experiment again if you can't find it. For graphs, make sure you are exporting the data in nonraster format, such as *.pdf, *.eps or *.svg. Exporting in these types of formats will prevent your graphs from looking pixelated and keep text legible no matter how you resize it later.

Tip 2: Use appropriate software when laying out your figures. PowerPoint may be user-friendly, but it is meant to work at screen resolution, which is only 72 ppi. When you export images from this program, you end up with a 72 ppi image that needs to be converted into a 300 ppi one (Figure 2). Conversion to a higher resolution image can result in an image that is too small for publication or one that is extremely pixelated. Additionally, depending on how you upscale, or increase the number of pixels in your image, your software program may introduce pixels into your image, thereby creating artifacts. Adobe Illustrator usually is recommended for figure assembly, but Inkscape and CorelDraw are good alternatives. These programs are meant to combine raster and vector data into a single figure and can do so without affecting the pixels found in raster data.

Tip 3: Set your canvas size to the physical dimensions provided by the

journal. Most journals provide two or three size options: single-column width, double-column width and occasionally 1.5-column width. Once you insert your graphics into the figure using the appropriate software, you usually don't have to worry about image resolution; however, be careful when increasing the size of a raster image. If you insert a 300 ppi image and decide you want to double its size, the resulting resolution of that image will be 150 ppi and likely will look less clear than the original.

Keeping track of image resolution shouldn't be a hassle. By incorporating these suggestions into your workflow, you can rest assured you have done your due diligence.



Kaoru Sakabe (ksakabe@asbmb.org) is the data integrity manager at the ASBMB.

Focus on exposure

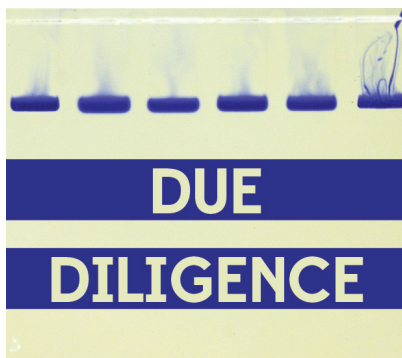
By Kaoru Sakabe

In past columns, I've made the point that figure preparation begins at data acquisition, but I haven't really explained my reasoning in depth. So here, I'll fill you in. Once you've snapped your picture or exposed your Western blot, that image becomes the version of record for your experiment. If the data you've collected is poor quality from the outset, your figure is already compromised.

One way to tell if you've nailed your image's acquisition parameters is to look at your image's histogram. Being able to interpret the histogram correctly can tell you if you can move forward with snapping the next picture of your mutant phenotype or if you need to tinker with the acquisition settings.

If you're a digital photography aficionado, you probably are very familiar with histograms and the information they contain. Here's a quick overview for those not yet accustomed to viewing them: A histogram of an image displays the distribution of pixels in the image, showing a graph of the number of pixels with a given intensity. For an eight-bit grayscale image, there are 256 possible intensities ranging from 0 (black) to 255 (white) for each pixel in the image. The histogram will not tell you how these pixels are distributed in space, just the distribution of the pixel intensity.

Ideally, you want the pixels to lie between the two extremes. This ensures that the fine details of your



images are captured. If the pixels are clustered at either end, you've likely oversaturated or underexposed your image. For example, aggressively adjusting the black levels of an immunofluorescence image to reduce the background eliminates hallmarks of a true experimental image. On the other hand, oversaturation leads to loss of fine details and makes it impossible to quantify the signal. Why? From the point of view of the detector, i.e., the camera or film, once it has recorded the maximum amount of signal, it cannot register any more. If you've hit the limit on either end of the histogram, the detector won't be able to tell you if a band or a cell feature is two times or 20 times more intense than a neighboring band or cell.

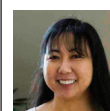
If you're acquiring images on a microscope or gel-documentation system, the hard part already is done for you, because these instruments typically show you the histogram of the image you've just acquired. If you are using film, take multiple exposures of your blot to make sure you are

within the linear range of the signal so you can properly quantify it. Once you've scanned your film, you can use either Photoshop or ImageJ to look at the histogram of your image. A telltale blip at either end of the histogram will tell you that you need to adjust your acquisition settings or use a different exposure of your film (Figure 1).

The histogram is also useful in telling you if your image has been over-adjusted during figure preparation. After you've adjusted the brightness or contrast settings of your image, make sure to check the histogram one final time. If the histogram has shifted too far to the left or to the right, you've likely truncated the pixels that were at the ends of the distribution, and your image is now overly adjusted (Figure 2). If your histogram shifts too far to either end, the resulting image may raise flags with reviewers or the journal, because it may look like you're trying to hide something. Remember, there's no need to hide your true experimental results!

Doing your due diligence at the image-acquisition phase will save you time as you prepare your figures for publication, which could be months or even years after you initially acquired your data. Going back and repeating an experiment because an immunofluorescent image was underexposed or a band was completely blown out can be frustrating, to say the least, so use these tips to make the most out of your data.

Doing your due diligence at the image-acquisition phase will save you time as you prepare your figures for publication.



Kaoru Sakabe (ksakabe@asbmb.org) is the data integrity manager at the ASBMB.

Combatting compression

By Kaoru Sakabe

I love reading blog posts about what the next generation will never experience because of changes in technology. Isn't it crazy that, at one point in time, data storage used to mean using floppy disks holding less than 1 MB? Or connecting to the internet involved using a phone line to dial into a server? Nowadays, cloud-based servers and online collaborative programs allow researchers to share large amounts of raw data quickly. External hard drives that can store terabytes of data can be purchased cheaply.

With the decreased cost of storing digital data and the ability to rapidly share files electronically, minimizing file size should no longer be a factor when deciding in what format to store your data. What you should keep in mind is that your electronic data should be stored in a universal format that does not alter its original information in any way, thus preserving your high-quality image data. In other words, you should be saving your files in a way that uses a lossless compression. For images, your go-to should be TIFF, or tag image file format, and



not JPEG, or joint photographers expert group format. Although there are other lossless file types, such as RAW, BMP or PNG, ideally you should save as a TIFF, because it is uniformly supported across different software platforms.

Because disk space and transfer speed were great limitations many years ago, scores of authors chose to save their images as JPEG files. But beware: JPEG files can compromise your hard-earned data. Technically, JPEG is not a file format but rather a method that specifies how the image will be compressed. You will see the extension JPG or JPEG when you save files this way, but there is no difference

between these two extensions. When an image is saved using JPEG compression, it is broken up into 8x8 pixel blocks, and a transformation then is applied to each block independently of the rest of the image to reduce the file size. This transformation also separates the color information from the brightness and discards more of the color information. Ultimately, JPEG is a lossy compression method (see the Due Diligence column in the January issue of ASBMB Today), which means that every time you save the file, you are discarding information. I'll demonstrate the reasons why you should avoid this format, and hopefully I can convince you to avoid using JPEGs altogether.

First, saving as a JPEG fundamentally alters the image in a way that cannot be restored. Take, for example, the original TIFF image shown in Figure 1. In last month's column, we discussed how informative histograms can be. Looking at the histogram of the TIFF image, we can see that the image contains many white pixels, some black pixels and a few pixels of various shades of gray. For JPEGs,

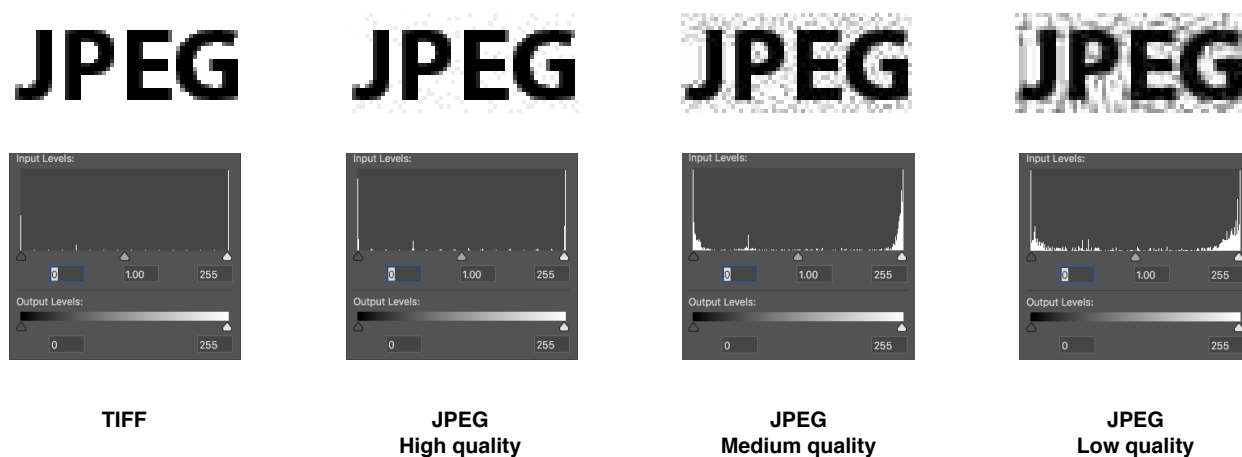


Figure 1. Saving your data as a JPEG changes the pixels in your image.

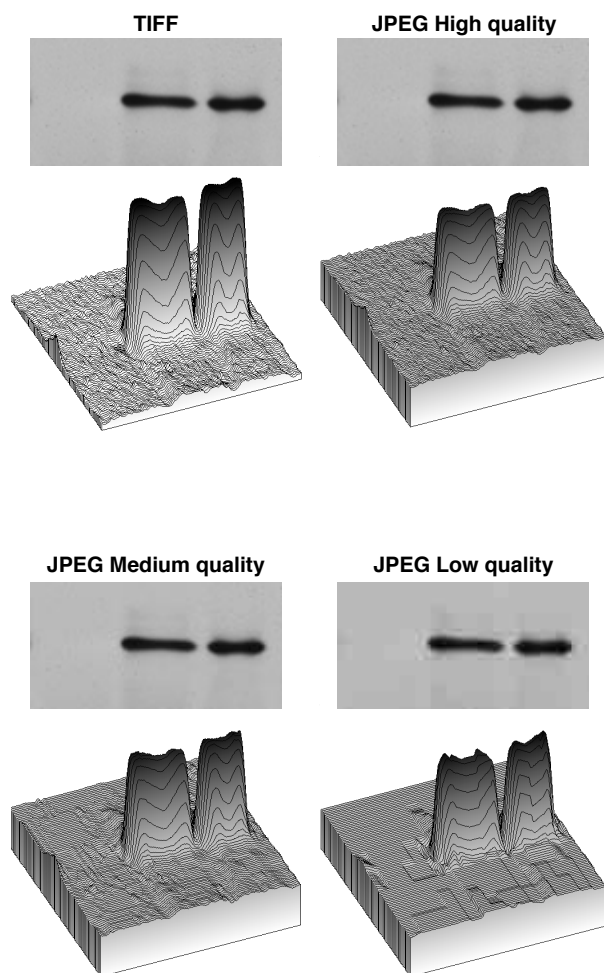


Figure 2. JPEG compression introduces artifacts.

high quality means little compression and a larger file size; low quality means high compression and smaller file size. Saving the same image as a JPEG at different quality levels introduces pixels that were not present in the original, creating a distorted image.

Now how does this translate into a scientific image? In Figure 2, I've taken a TIFF image and saved it at three different JPEG qualities. Visually, there doesn't appear to be a huge difference between the TIFF and the high-quality JPEG; however, if you analyze the image with a surface-plot analysis, you'll notice appreciable differences between the two images. As you compress the image further,

blocks start to appear, the background looks less like a real experiment and the bands seem pasted in. These artifacts occur especially in areas of high contrast, such as a dark band on a clean background.

Another issue is that each time a JPEG image is saved, the compression is applied. Repeatedly saving an image during editing can introduce artifacts. For example, in Figure 3, I've taken an image of a dividing cell and saved it 100 times at maximum quality. By the 100th save, several anomalies have appeared, and it no longer looks the same as the original. While this exercise is almost certainly an exaggeration of what's happening in the lab, it illustrates that each time you save in

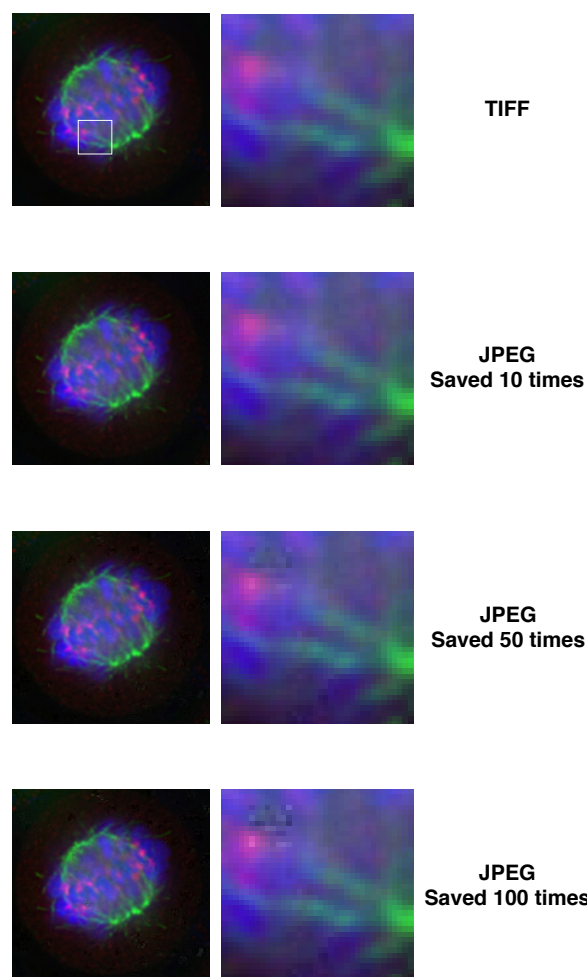


Figure 3. Repeatedly saving as a JPEG introduces artifacts.

the JPEG format, you are changing your data.

Finally, remember that by snapping the picture of the cell or scanning your film, you are recording the results of your experiment. Saving the image in a lossy file format, such as JPEG, distorts the actual results you obtained. Don't get stuck assembling a figure with muddled data. By saving your image initially in a lossless format, such as TIFF, you will be doing your due diligence in preserving your data.



Kaoru Sakabe (ksakabe@asmb.org) is the data integrity manager at the ASBMB.

(Power)Point (not) taken

By Kaoru Sakabe

For some authors, PowerPoint is the go-to software for preparing figures. Despite its popularity, PowerPoint has some serious shortcomings. Here, I will discuss why PowerPoint is not a good tool for illustrating your data. Instead, you should be using software dedicated to graphics editing, such as Illustrator or Inkscape, to ensure that figure quality is not compromised.

PowerPoint is a great program for making slide presentations. It's just not so great for making figures. PowerPoint works at screen resolution, which is 72 ppi. So when you export or save slides as TIFF files from



PowerPoint, it exports them with a resolution of 72 ppi, which is insufficient for print.

Now, some people use work-arounds. For example, they open the exported TIFF files in Photoshop and simply change the resolution from 72 ppi to 300 ppi (or more). But beware — there are pitfalls to going this route.

When you go to the “Image Size” box in Photoshop, you will see the “Resample” box. If this box is checked when you upscale or increase the resolution, Photoshop will keep the physical dimensions of the image the same and create additional pixels to generate a higher-resolution image. The additional pixels are created by a method called resampling, and the resulting image may therefore not appear the same as the original.

On the other hand, if you uncheck the box, the program will not create any new pixels but will instead decrease the size of the image to increase its resolution. This new image may be extremely small and may not be usable by the pub-

lisher (Figure 1).

There are some other ways of turning PowerPoint files into TIFFs, but they suffer similar drawbacks. For example, some users might take a screenshot of the PowerPoint slide and then convert the screenshot into a 300 dpi TIFF in Photoshop, but this method has the same issues as above.

Another method I've seen used is to save the PowerPoint as a PDF and then convert the PDF into a TIFF. This method risks compromising image quality, as certain types of PDF conversions apply a JPEG compression to the file. If you recall from my May column, JPEG compression can introduce artifacts into your images and should be avoided at all costs.

There are other work-arounds that involve changing registries or doing some conversions that I won't get into, but the easiest way to avoid these problems from the outset is to use software intended for combining high-resolution images with vector graphics.

The bottom line is that PowerPoint is software designed to give screen presentations, not to produce publication-quality figures. In fact, the output from PowerPoint needs to go through multiple conversion steps to arrive at a final product with possibly compromised integrity. Using appropriate software, such as Illustrator or Inkscape, during figure composition will ensure you are presenting your data in the best possible way — with due diligence.

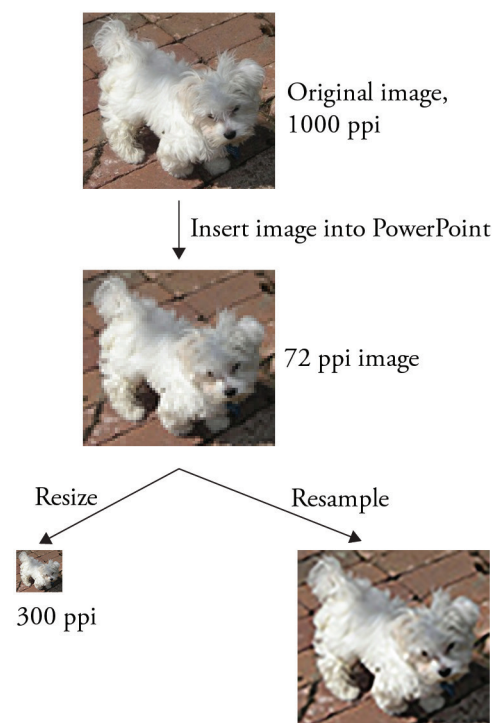


Figure 1: I've taken a 1000 ppi photo of my dog, Kim Chee, and converted it into a 72 ppi image, a scenario that would happen if you inserted your image into PowerPoint. To convert the PowerPoint into a 300 ppi TIFF, Photoshop will either resize or resample the image, and both processes may yield unusable images.



Kaoru Sakabe (ksakabe@asbmb.org) is the data integrity manager at the ASBMB.

Play by the rules — and spot the error

By Kaoru Sakabe

Over the course of this year, I've offered advice on figure presentation and assembly. For this month's installment of Due Diligence, I thought it would be helpful to discuss the American Society for Biochemistry and Molecular Biology's expectations for creating technically and ethically sound figures as well as caution against some common image adjustment errors. I'm also letting you, the reader, play reviewer by showing you several figures (below) to see if you can spot the presentation error. Answers are on page 48.

To ensure the integrity of the papers they publish, many journals provide guidance for acceptable practices when it comes to image manipulation. These standards were



introduced in response to an increase in inappropriate adjustments of images in figures. Most journals have adopted the Journal of Cell Biology's stance on image manipulation because it's thorough and rigorous:

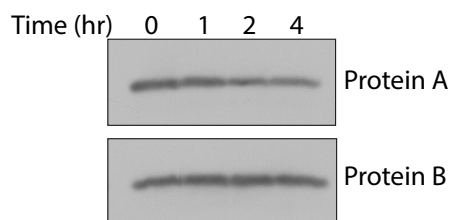
"No specific feature within an image may be enhanced, obscured,

moved, removed, or introduced. The groupings of images from different parts of the same gel fields or exposures must be made explicit by the arrangement of the figure (e.g., using dividing lines) and in the text of the figure legend. Adjustments of brightness, contrast, or color balance are acceptable if they are applied to every pixel in the image and as long as they do not obscure, eliminate, or misrepresent any information present in the original, including the background. Nonlinear adjustments (e.g., changes to gamma settings) must be disclosed in the figure legend."

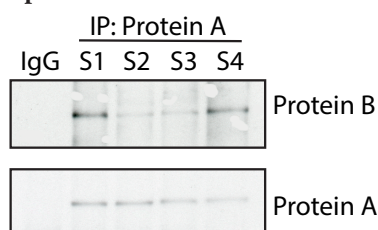
The nuts and bolts of this policy are that authors should limit the beautification or touch-up work they apply to an image. Here are the key points:

What's wrong with these figures?

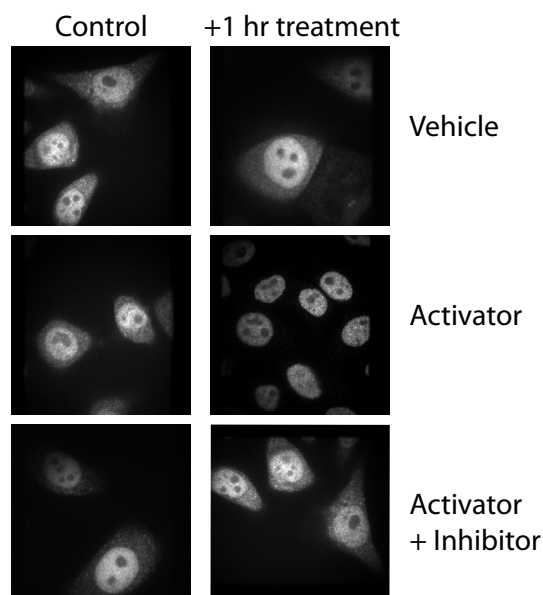
Example 1



Example 2



Example 3



- Your figure should look like the original capture of the image. Changing portions of the image to make it aesthetically pleasing or to enhance your data improperly is not allowed.
- If an adjustment is absolutely necessary, it must be applied to the entire image and not just a selection.
- Along the same lines, while authors most commonly think of the brightness and contrast tools to tweak a blot or micrograph, resizing should be applied to images with the same care as these other adjustments. If an image size needs to be adjusted, make sure you maintain the aspect ratio. If you stretch or compress an image in just one direction, you are not treating the pixels in the image equally.
- If you make any nonlinear adjustments, such as changes to the gamma

setting, you should declare these adjustments in the figure legend. These types of adjustments do not increase or decrease pixel intensity levels evenly across the image. The low and high tones are adjusted at a different rate than the midtones, meaning that, again, not all pixels in the image are treated equally.

- The image background is part of the data and never should be removed or misrepresented in the figure. Remember, background is a hallmark of authentic data.
- Finally, you should be as transparent as possible in the figure and figure legends. Disclosing gel splices, reuse of control data and image acquisition settings allows readers and reviewers properly to assess your data.

Now that you have the rules of the

game down, let's see how well you do with spotting a presentation error. See if you can figure out what each author has done wrong.

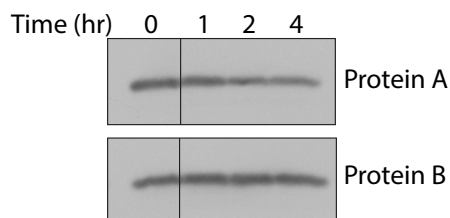
These examples, as well as many other types of manipulation, are caught easily by a trained eye and imaging software. Depending on the severity of the manipulation, these alterations could be damaging not only to your reputation but to those of all the co-authors on your paper. Make sure you are preserving the integrity of the scientific record by playing by the rules of the game.



Kaoru Sakabe (ksakabe@asmb.org) is the data integrity manager at the ASMB.

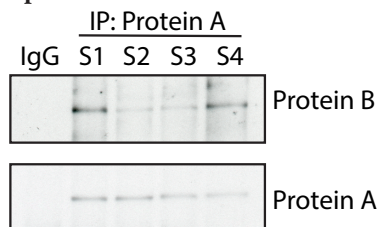
Answers

Example 1



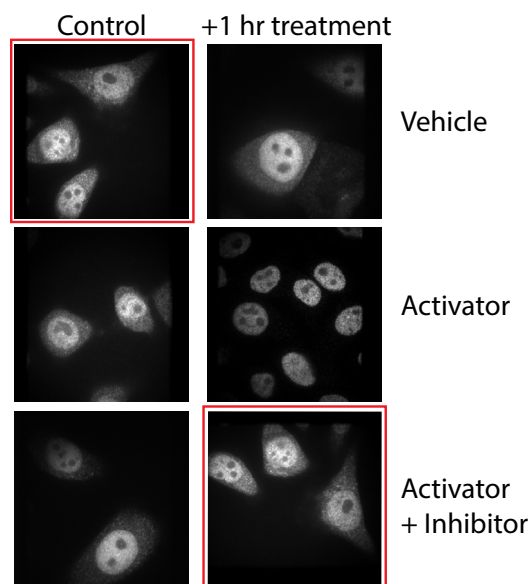
In this instance, the authors omitted the dividing lines to indicate that different sections of an immunoblot were spliced together to create the final image.

Example 2



Here, the authors removed some background spots in the immunoblot for Protein B to beautify the image. To show a true representation of the original capture, the authors should have left the spots in the final figure.

Example 3



In this example, the authors rotated the control, vehicle-treated panel and inadvertently inserted it into the activator-plus-inhibitor panel. Be sure to label your data effectively so that when you are assembling your figures later, you can determine which image goes with which experimental condition. Simple 1, 2, 3 labels can be hard to decipher a year or even a month after you've generated the data. Be sure to check the final figure against the original data to ensure that nothing was switched inadvertently during assembly.

Keep it safe

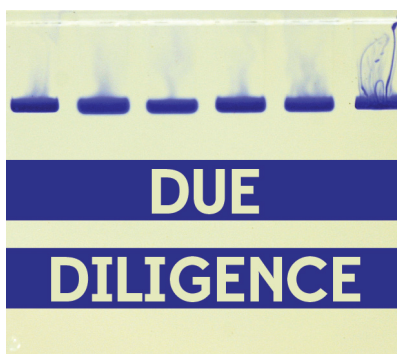
by Kaoru Sakabe

Over the course of the year, I've discussed best practices in image acquisition and manipulation, but I haven't really talked about data and data management. I recently had the opportunity to hear Brian Nosek from the Center for Open Science speak, and he summed it up perfectly: "The report is advertisement for research. The data is the actual research."

What steps are you taking to ensure that the research record is maintained and that the data you've labored to gather are protected?

I'm sure you've heard many times about the importance of keeping good research records and storing samples appropriately in obligatory classes on responsible conduct of research or from your mentor, so I won't belabor these points. Nothing is more frustrating when you're assembling figures or writing up your results months or years after data collection than not being able to piece together how an experiment was performed. Additionally, proper storage ensures that samples can be used later, which may be particularly important when sample size is limited or you need to return to these samples years down the road. But what about film, pictures or any other data files that you've collected?

With regard to blots or micrographs, I suggest multiple forms of storage. If you are acquiring images from blots using an imaging system, retain the native, system-generated file in addition to the exported image file in TIFF format. The same goes for micrographs. TIFF files are a universal file format that can be opened by anybody without specialized software. The same cannot be said for



the system-generated file. If you are scanning film, save it in TIFF format and be sure also to save the film. If the computer collecting the data is a shared computer, make sure you keep a copy for your records. With multiple users, you cannot guarantee that your file won't get accidentally deleted or corrupted.

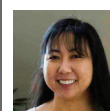
What about data from an instrument using obsolete software? Some people keep a legacy system so they can open the original system-generated file long after the instrument stops running. If at all possible, export the data into a universal file format. Often you can export column runs or FACS data as Excel or PowerPoint files.

When you are saving images, make sure you save the entire blot or gel, because it tells the whole picture of the experiment. You can always make a copy of the image to crop if the additional space isn't necessary for the paper. The original always should be retained, as it's better to keep something rather than throw away information that may be useful later. For example, a reviewer may ask you to indicate molecular weight markers or to show the full scan of the blot, and you would be at a loss if you only saved a portion. Also, while adjusting

the settings of your image may be necessary when assembling figures, make sure to save a completely unaltered image for your records.

Finally, it's important to back up your files and to store them in a safe location. Ideally, a backup should be kept separate from the original data in case something catastrophic should happen at the original location. Cloud storage as well as the availability of cheap, portable hard drives make storing safe, secure backups a relatively easy thing to do.

How long are you supposed to retain your data? There's no clear-cut answer. Institutions generally have their own policies in place. If your research is funded by the National Institutes of Health, then the general rule is that data must be retained for at least three years after filing the final financial report for the grant. Another number often thrown around is six years, as this is the statute of limitations for pursuing cases adopted by the Office of Research Integrity, the oversight office for research funded by the U.S. Public Health Service. What is likely less known is that any time an author benefits from the work, such as citing the work in a grant application or paper, the clock resets for the ORI to investigate cases of falsification, fabrication and plagiarism. Change also is on the horizon as the NIH develops policies on data retention and sharing. Remember, though, there is no limit to how long an article may be read and cited, so why should there be a time limit for retaining data?



Kaoru Sakabe
(ksakabe@asbmb.org) is the data integrity manager at the ASBMB.

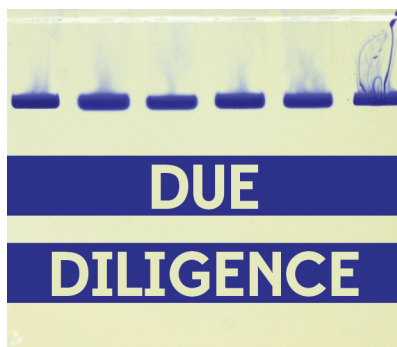
Q&A with the ORI's Kathy Partin

By Kaoru Sakabe

Throughout this series, I hope I have conveyed the importance of performing your due diligence when it comes to data presentation. For my final column, I asked Kathy Partin, director of the Office of Research Integrity, or ORI, to share her thoughts on questions you may have had while reading the series. For those not familiar with the ORI, this government agency oversees and directs U.S. Public Health Service research integrity activities. Partin, who joined the ORI in December 2015, will be detailed to the Uniformed Services University of the Health Sciences in early December. Her responses have been edited for length and style; the full responses may be found online at asbmb.org/asbmbsmbtoday.

What are questionable research practices? Is there a way to easily spot these? How should these be handled by individuals?

Questionable research practices, or QRPs, are those that fail to conform to well-accepted best practices and lead to lack of rigor and reproducibility but do not rise to the level of data falsification and/or fabrication. Examples could include the use of inappropriate controls, neglecting negative outcomes or using inappropriate statistics to support a hypothesis. These practices are not necessarily easy to spot, but they tend to be obvious with appropriate oversight, such as when one regularly compares raw data to analyzed data. Many believe that a nonrigorous laboratory culture can lead to a slippery slope



that will allow trainees to more readily make decisions about data acquisition or publications that could include misconduct. When lab members see questionable research practices, it is important to consider whether the training missed critical elements and to assess if there is adequate supervision. On the other hand, if a trainee continues to engage in questionable practices, despite good training, it might be appropriate for lab members to scrutinize the data for possible misconduct.

What are your recommendations to students/postdocs if they believe they are witnessing research misconduct?

Each of us plays a role in setting the ethical climate of the research environment, and with that comes the responsibility of taking appropriate action. If you see something, say something. However, it is important to understand that saying something sometimes puts the complainant at risk. Sometimes the complainant believes there is misconduct when there is not. Ideally, trainees have a trusted mentor to help think through

potential explanations for what is being observed. A mentor can be sure a trainee understands the implications of alleging misconduct, particularly against his or her research advisor.

If someone suspects misconduct and doesn't have anyone to talk to, there are options to consider. Read the institution's research misconduct policy. Consider contacting the institution's research integrity officer or the ORI with questions or hypothetical scenarios. Many institutions use anonymous hotlines. It is generally wise to avoid direct confrontation with the person, such as accusing him or her. It is a good idea to document details related to the possible misconduct. The time may come to make an official allegation, and then consulting with the research integrity officer is essential.

How should mentors handle suspected research misconduct in their laboratory or department?

Mentors might think that they have an obligation to investigate at the first inkling of a problem. Not true. It is better to take even early concerns to the research integrity officer so any steps taken will comply with regulations and ensure a fair process. Often, allegations can be handled more efficiently and even dispelled when the integrity officer is consulted early. Documenting concerns, including information about hard drives, file names, notebooks, dates and locations, is very helpful. Mentors who investigate outside the institutional process risk unnecessary disclosure of allegations. Confidentiality is critical to protect both the

complainant and the accused. It is everyone's responsibility to ensure the integrity of research, but we must also protect others from being unfairly identified as performing misconduct. The key is to take action but discuss your concerns only with the appropriate institutional officials, such as the dean or research integrity officer.

What can mentors do to prevent questionable research practices or research misconduct in their laboratory?

Effective mentoring is key. In any mentor-mentee relationship, mentors accept responsibility for ensuring that their mentee is adequately trained and supervised. A critical responsibility of a good supervisor is effective communication before, during and after experiments are performed. Good supervisors are respectful of their trainees, supportive, available, prepared and honest. Supervisors need to be sure that they properly instruct on research methods, convey responsible research practices, foster intellectual development and routinely check that the trainee is following through with what has been taught. A good mentor-mentee relationship creates an invisible safety net for the entire research group.

Certain terms are often heard, such as research misconduct, ethics violations, unethical behavior and fraud. Can you elaborate on the differences? Are there differences in how an institution may handle these versus the ORI?

There is definitely some confusion about terms. From the federal govern-



COURTESY OF KATHY PARTIN

Kathy Partin is director of the Office of Research Integrity, the government agency that oversees and directs U.S. Public Health Service research integrity activities.

ment's perspective, there is only one definition of research misconduct: "fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results." Institutions may have a much broader definition and might include terms like "serious deviation from accepted policies," misappropriation, misrepresentation, etc. An institution may make a finding of misconduct based on its broader definition, and the ORI will not because we are statutorily required to use a narrower definition. Globally, other countries and institutions also have widely varying definitions of research misconduct. Private foundations or sponsors of research might have their own terms that govern what an institution must do if research misconduct is alleged.

Are there agencies like the ORI in other countries? Does their definition of research misconduct differ from the ORI's?

Yes, there are regulatory agencies in Europe, Asia, Africa and North and South America that govern how research misconduct is adjudicated. Not only do definitions of misconduct vary, but the processes that govern the handling of allegations also

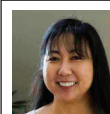
vary widely. Recently, countries have been forming networks of research integrity officers to share best practices and novel approaches and also to begin to think about how to harmonize some of the practices. The ORI collaborates and communicates with such associations as much as possible.

How does our current incentive system play a role in research misconduct or QRPs?

There are those who believe that the competitive nature of funding and publishing decisions forces perverse incentives for success. Ultimately, good research requires rigor and integrity. Trainees who are going to be successful, productive scientists and who are resilient in the face of challenges must understand that the foundation of success in research is research integrity. Trainees must believe that quality of research, not quantity, is what is needed for a successful career. Quality research comes from places that value research integrity; trainees should think about the climate of the lab they might select as well as the quality of the research being done.

Interested in getting advice from Kaoru?

Kaoru and other JBC staff will be providing advice and answering questions in webinars about scientific writing and publishing. To sign up to receive information about upcoming webinars, visit <https://asbmb.realmagnet.land/JBCwebinars>.



Kaoru Sakabe (ksakabe@asbmb.org) is the data integrity manager at the ASBMB.