© 2020. Published by The Company of Biologists Ltd | Journal of Cell Science (2020) 133, jcs250027. doi:10.1242/jcs.250027

OPINION

SUBJECT COLLECTION: IMAGING

Hypothesis-driven quantitative fluorescence microscopy – the importance of reverse-thinking in experimental design

Eric C. Wait, Michael A. Reiche and Teng-Leong Chew*

ABSTRACT

One of the challenges in modern fluorescence microscopy is to reconcile the conventional utilization of microscopes as exploratory instruments with their emerging and rapidly expanding role as a quantitative tools. The contribution of microscopy to observational biology will remain enormous owing to the improvements in acquisition speed, imaging depth, resolution and biocompatibility of modern imaging instruments. However, the use of fluorescence microscopy to facilitate the quantitative measurements necessary to challenge hypotheses is a relatively recent concept, made possible by advanced optics, functional imaging probes and rapidly increasing computational power. We argue here that to fully leverage the rapidly evolving application of microscopes in hypothesis-driven biology, we not only need to ensure that images are acquired quantitatively but must also re-evaluate how microscopy-based experiments are designed. In this Opinion, we present a reverse logic that guides the design of quantitative fluorescence microscopy experiments. This unique approach starts from identifying the results that would quantitatively inform the hypothesis and map the process backward to microscope selection. This ensures that the quantitative aspects of testing the hypothesis remain the central focus of the entire experimental design.

KEY WORDS: Experimental design, Hypothesis, Image analysis, Microscope choice, Microscopy, Quantitative analysis

Introduction

Advancements in optical engineering, labeling technologies, and computational capacity have turned fluorescence microscopy into an indispensable tool in the life sciences. Its unique capacity to probe biological questions across a large range of biological length scales has made it a popular tool in cell biology, neurobiology and developmental biology, as well as many other fields of research. Modern microscopy can reveal valuable information on molecular ultrastructure, dynamic biological processes and biological functions. Yet, the appeal of seemingly limitless promises, the myriad of technical details and the rapid development of computational capabilities has also created confusion for many seeking the right combination of imaging tools. As has been previously pointed out by Jonkman and colleagues (Jonkman et al., 2020), biologists can spend considerable time and resources acquiring huge amounts of data without proper planning, only to realize later that the data cannot appropriately address a particular biological question. This usually occurs when the design of a microscopy experiment is not guided by a suitable hypothesis, the experimenter gets side-tracked by new observations or the experiment starts without a design at all. The method proposed here aims to assist the gathering of appropriate data that directly addresses a quantitative hypothesis. The intent is to give the reader a better understanding of the process and potential issues that arise in quantitative experiments.

The importance of fluorescence microscopy lies in its ability to serve both as an exploratory and a quantitative tool. In other words, microscopy has a combined capacity that enables a biologist to both formulate hypotheses based on observation and to perform quantitative measurements to test those hypotheses. For example, one might easily observe the localization of a target protein within a mitochondrial compartment. However, it takes a shift in mindset to design an appropriate experiment capable of quantifying this localization change in response to an oxidative stress. Quantitative measurements, however, can only produce results that directly address a proposed hypothesis when the experiment is designed appropriately. In fact, even an accurate, quantitative set of data that has been generated with the best practices will not necessarily yield biologically meaningful results. An image acquired with a digital detector is inherently a data map – an array of values. While any digital image can be quantified, these measurements are only biologically meaningful when they are pertinent to the hypothesis. Take for example a study that investigates the rates of filopodia extension during cell migration. Data revealing the super-resolved, 3D actin filaments are not sufficient for determining the rate of filopodia extension. However, an experiment that captures the change in location of the filopodial tip will provide the necessary data. In other words, when testing a quantitative hypothesis, informative data are quantitative, but not all quantitative data are informative.

Reliable and informative results require high-quality image data and relevant analyses. Fortunately, there is no shortage of excellent reviews in the literature that offer step-by-step guidance to perform microscopy experiments, from image acquisition to quantitative image analysis (Berg et al., 2019; Jonkman et al., 2014; McQuin et al., 2018; North, 2006; Rueden et al., 2017; Swedlow, 2013; Van Den Berge et al., 2019; Waters, 2009; Weigert et al., 2018). The task now lies in ensuring that data acquisition and analyses can be translated into biologically meaningful information, capable of challenging a hypothesis. We argue that this must be achieved through rational experimental design.

Designing a hypothesis-driven experiment is a vital step in the overall experimental scheme, but it is often over-simplified and represented by a single step. The conventional workflow of an imaging experiment, as astutely observed by Lee and Kitaoka (2018), is adapted in Fig. 1A. In this generalized diagram, the execution of the experiment begins with sample preparation after experimental design. The images are acquired, and the data will then be processed and analyzed – usually followed by several iterations of optimization – before the final results are presented. What is

Advanced Imaging Center, Howard Hughes Medical Institute, Janelia Research Campus, Ashburn, VA 20147, USA.

*Author for correspondence (chewt@janelia.hhmi.org)

E.C.W., 0000-0002-1614-9219; M.A.R., 0000-0002-3426-1204; T.-L.C., 0000-0002-3139-7560

Journal of Cell Science (2020) 133, jcs250027. doi:10.1242/jcs.250027

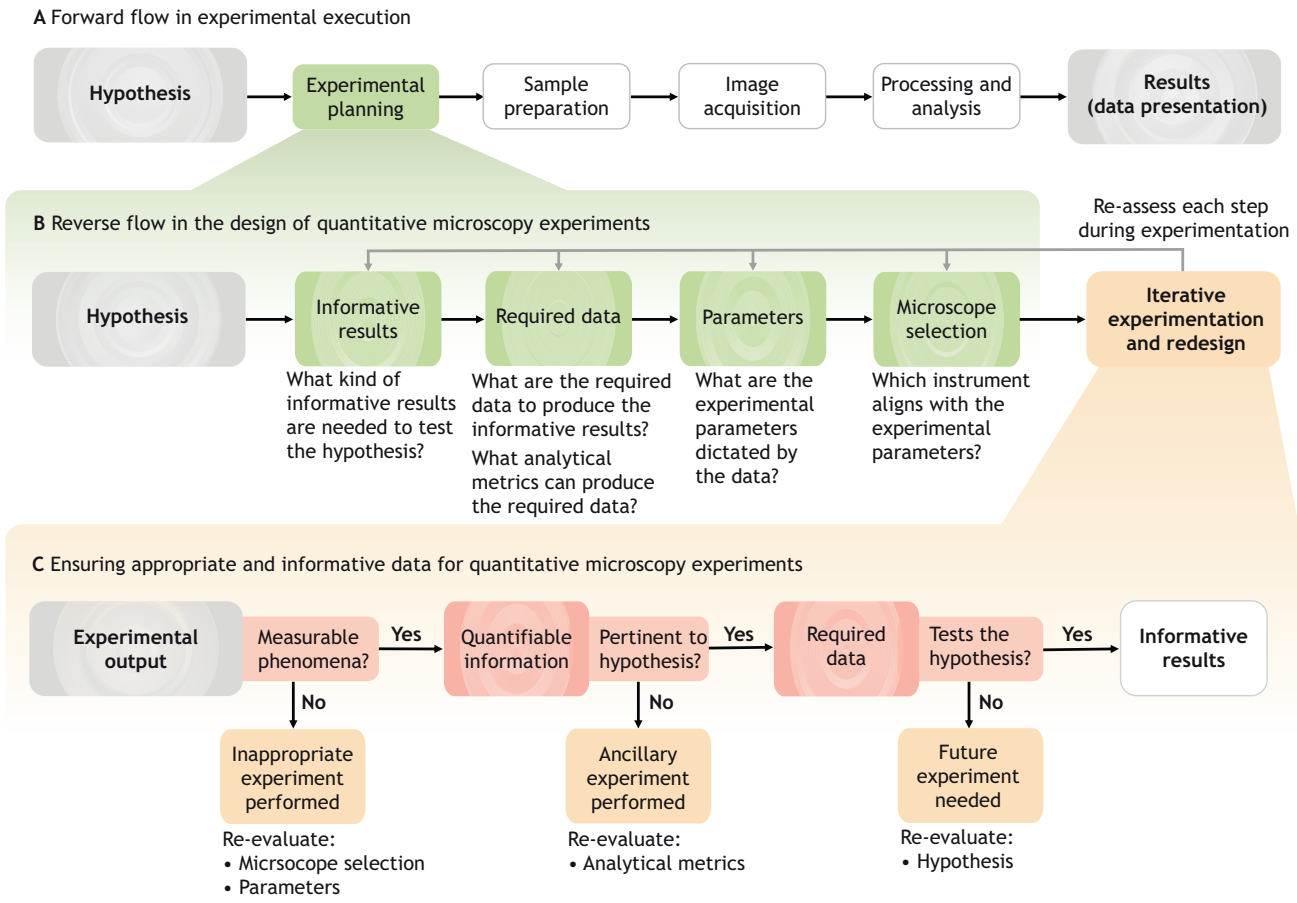


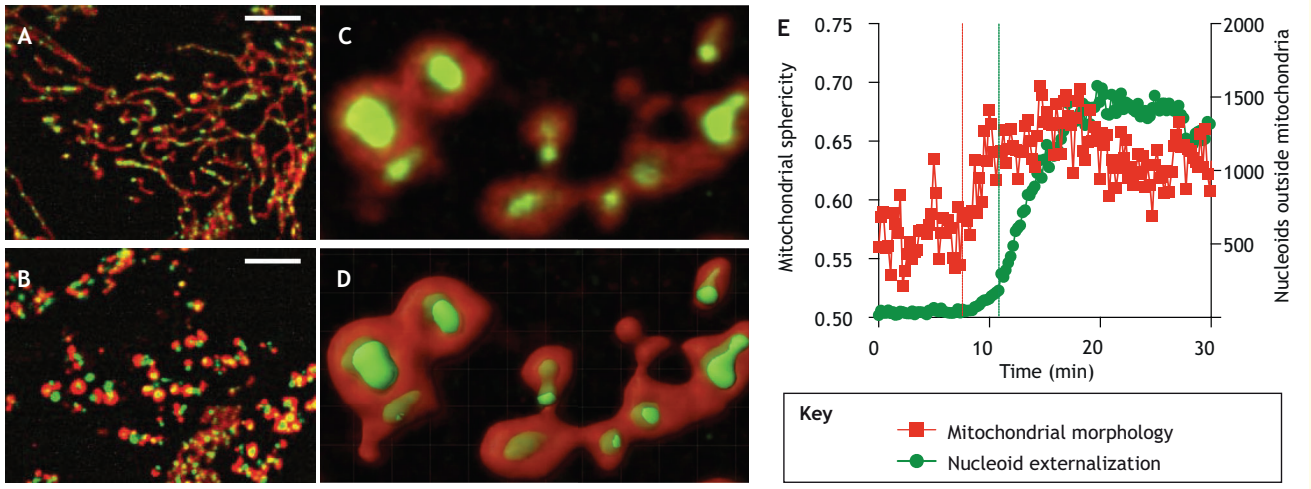
Fig. 1. Conducting and designing quantitative fluorescence microscopy experiments. (A) Typical workflow in microscopy experiment. This workflow is forward-facing, progressing from the formulation of a hypothesis to the eventual presentation of the data as results. Adapted with permission of American Society for Cell Biology from Lee and Kitaoka (2018); permission conveyed through Copyright Clearance Center, Inc. (B) A focused view of the experimental planning phase. We propose that experimental design would be more efficient and effective by adopting a reverse-facing workflow. Here, the hypothesis should determine what the necessary results should be. From there, the experimenter can plan backward from the required data to the point where the experiment can be executed. The processes outlined in A and B are iterative, and the experimenter should re-evaluate whether the best decision has been made at each step. (C) A flow diagram to determine whether the experimental output generated from the microscope will lead to informative results. Answering the questions outlined here will identify the corresponding step in the design that needs re-evaluation. Reaching the 'Informative results' box would indicate that the data acquired were most likely collected in a manner that would directly test the hypothesis. Alternatively, the bulleted lists provide insight into which step in the design process requires re-evaluation to be improved in subsequent design iterations.

important to note is that experimental design is appropriately singled out as the key first step (Fig. 1A). Yet, in stark contrast to the wealth of technical guides, there is a paucity of discussion in the literature on the logic of rational experimental design and how it can be harnessed to successfully perform a hypothesis-driven, quantitative experiment. This is an unfortunate omission, partly due to the difficulty in summarizing a logical scheme that is sufficiently general to be applicable to most biological questions. In this Opinion article, we aim to fill this important gap and focus on rational, hypothesis-driven experimental design. This guide is aimed toward biologists interested in learning how to design quantitative experiments that are geared toward testing their hypotheses. It embodies our experience in steering imaging projects from hypotheses to quantitative, informative results at the Advanced Imaging Center at HHMI Janelia Research Campus (Chew et al., 2017). We include in Box 1 a case study of how we have successfully steered the development of such a quantitative microscopy project.

The success of a microscopy-based quantitative experiment hinges on the appreciation and understanding of (i) how the underlying biological query and defined hypothesis directs the experimental design, and (ii) how experimental design and instrument choice are related to the way in which image data will eventually be analyzed. For this reason, we outline a logic that exemplifies these themes (Fig. 1B). We propose, in this Opinion article, that the very first step of experimental design, following the formulation of a hypothesis, is to determine the informative results that can quantitatively test that hypothesis. In other words, informative results are the ultimate goal of the designed experiment. Therefore, an experiment that has been developed to specifically generate data pertinent to the biological query will produce informative results. As such, the production of the required data will necessitate that a certain set of experimental parameters be met, which would in turn prescribe the features of the instrument needed to make such measurements. Overall, such a systematic workflow ensures that the hypothesis remains central to the experiment and that the experiment yields information capable of challenging the hypothesis. This will help chart the roadmap of how microscopy-based experiments should be designed for quantitative analyses. We will not replicate the many superbly written reviews

Journal of Cell Science (2020) 133, jcs250027. doi:10.1242/jcs.250027

Box 1. Case study – visualizing mitochondrial DNA release during apoptosis



This case study partially summarizes one of the quantitative experiments performed by McArthur and colleagues (McArthur et al., 2018). Preliminary observations indicated that the mitochondrial network of cells deficient in induced myeloid leukemia cell differentiation protein (MCL-1), a Bcl-2 family member, broke down during apoptosis (A in the box figure), followed by the presence of mitochondrial DNA (mtDNA) in the cytoplasm (B in the box figure). This observation led to the conceptualization of the working model – 'during apoptosis, mitochondrial morphology changes prior to the release of mtDNA into the cytoplasm'.

To properly plan a quantitative experiment to test this model, we used our reverse-logic to steer the following experimental design:

1. A more-defined hypothesis was formulated – 'during apoptosis, the mitochondrial sphericity increases prior to an increase in the number of externalized mtDNA'. Note how the initial descriptive semantics have been translated into quantitative semantics that will guide subsequent measurements.
2. Two sets of informative results were essential to test this hypothesis: (i) mitochondrial sphericity, and (ii) mtDNA externalization, both measured as a function of time.
3. To achieve these informative results, the required data must include time-lapsed, volumetric images of labeled mitochondria and mtDNA.
4. To produce these data, the following experimental imaging parameters had to be met:
   - high-speed volumetric imaging to accurately track 3D mitochondrial network reorganization
   - high signal-to-noise ratio and resolution in order to accurately measure the 3D structures of the mitochondria
   - near-isotropic resolution to precisely characterize the sphericity of mitochondria and mtDNA extrusion
   - two-color acquisition to provide information on both the mitochondria and the mtDNA.
5. While both lattice lightsheet microscopy (LLSM) and 3D structured illumination microscopy (SIM) met these benchmarks, it was also important to meet the biological requirements. Pilot studies established that two-channel volumes of 50 slices each, acquired approximately every 10 s for a total of 50 min would be necessary to capture and follow this rare process in its entirety. Photoxicity could affect the mitochondrial biology, introducing artifacts. To mitigate phototoxicity, the gentle illumination of LLSM established it as the clear choice. To further reduce light exposure, brighter fluorescent labels, such as mNeonGreen (Shaner et al., 2013) and HaloTag™ (Promega, USA) with Janelia Fluor® 646 (Grimm et al., 2017) (instead of EGFP and mCherry), were used. Note that the experimental design process was iterative and benefited from pilot studies used to identify the necessary imaging parameters, suitable fluorophores, and the optimal microscope.

C to E in the box figure illustrate the successful completion of this quantitative experiment. The LLSM micrograph (C) shows mtDNA extrusion from mitochondria. These images were used to create 3D segmentations (D) and were quantified. The mitochondrial sphericity and mtDNA externalization were measured over time, and plotted in E. This graph shows that an increase in mitochondrial sphericity (thin red line) preceded the onset of mtDNA extrusion (thin green line) – providing the informative result that ultimately supported the hypothesis.

The box figure shows morphological changes of mitochondria and mitochondrial DNA release during apoptosis; images were previously published in McArthur et al. (2018) and are reused here with permission. Scale bars: 5 µm.

and guides in the literature here, but rather aim to help readers better utilize these guides, as we embark on our journey of experimental design.

Observation-driven exploration versus data-driven analyses

The capacity of modern optical microscopy to support both visual exploration and content-rich measurement has made it a versatile biological research technique. Unfortunately, it is also one that is commonly misunderstood. Biologists are keen observers, exceptional in recognizing patterns, finding anomalies and identifying new phenotypes. In fact, when it comes to studying structures and processes, visualization by itself is often sufficient to prompt biologists to formulate working models of the observed systems, and these working models provide abstract representations of the observation. The descriptive semantics used in these working models have served as powerful tools in life sciences and enable biologists to organize and communicate information about the complexity of the living systems (Courtot et al., 2011). Indeed, specific follow-up questions can often already be framed by experienced biologists as soon as the initial images appear on their monitor; and this is the inception point of many biological queries. This is the essence of observation-driven, empirical inferences – 'I know it when I see it', and this is where the power of microscopy has historically been leveraged. Observational biology will continue to play an important role, and it is certainly true that not all biological hypotheses must be quantitatively tested. However, there is no denying that with the advent of modern experimental methods, hypotheses in general have become, and are increasingly expected to be, formulated in

more quantitative terms. Addressing these increasingly focused hypotheses is where the quantitative capacity of microscopy has the most impact and is the core of this Opinion article.

If one were to accept the idea that 'seeing is believing' with microscopy as an exploratory instrument, then surely one must also accept the notion that 'measuring is knowing' when using microscopy as an analytical technique. The challenge here is to reconcile observation and quantification using the same instrument. Quantitative measurement is intrinsically analysis-rich and semantics-agnostic (Shasha, 2003). However, this is where the disparity between observation and quantification often arises. It is common to see proposed microscopy studies with phrases such as 'to analyze the spatial-temporal dynamics of an organelle'. There is unfortunately no specific analytical metric for the 'dynamics' of an organelle or any other biological structure. Dynamics is an ambiguous term that is often used to encapsulate several different metrics that together describe a particular observation. To transform vague biological queries such as this into quantifiable goals for microscopic analysis, we need to consider how intuitive biological semantics can be reformulated. With this in mind, we will begin by exploring how hypotheses shape the rationale of microscopy-based experiments.

Testable hypothesis

The cornerstone of the classical scientific method is to determine whether evidence supports or negates a postulated idea. Hypotheses, at the experimental level, must therefore be negatable by observation or measurements (Popper, 2005). A clearly stated, verifiable hypothesis will guide every step of an experiment and will provide invaluable checkpoints. More importantly, the negatable hypothesis will impart the necessary restraint to mitigate being side-tracked from the initial question. This disciplined approach does not preclude future exploration of other observations, but it serves to balance both the exploratory and the analytical priorities of an experiment (Fig. 1C). This is why a hypothesis such as 'condition X will increase the rate of mitochondrial fission' has stronger semantic specificity than 'condition X will affect the spatial-temporal dynamics of mitochondria'. The latter hypothesis cannot be tested because the experimental variables (i.e. fission events) that either support or negate it are not defined.

Interestingly, such cautionary advice is rarely needed for biochemical and molecular biology assays. These are assays that are uniquely quantitative and do not usually serve as observational tools, and biologists learn these techniques extensively during their training. As a result, biologists formulate testable hypotheses and perform quantitative analyses with ease using assays such as immunoblots, PCRs, ELISAs or enzyme kinetic assays. What differentiates these assays from microscopy is that they are explicitly linked to well-defined sets of output. For example, an immunoblot yields specific information on molecular mass and abundance. In contrast, a vast plethora of information can be derived from microscopy data, including molecular abundance, spatial location, movement behavior, morphological changes, structural features, molecular association, enzymatic activity, and the list goes on. Microscopy is therefore not a single assay; instead, it is a collection of assays that vary depending on how the experiment is designed. Without a defined boundary, the scope of an experiment can quickly become too ambitious and unnecessarily complex. This underscores the importance of identifying the appropriate experimental output that addresses the hypothesis early in the design process.

Compared to biochemical and molecular biology assays, the complexity of microscopy is further compounded by the variability in the nature of the sample. In comparison to molecular biology assays that use defined samples for input, such as nucleic acids or proteins, microscopy can accommodate a wide variety of complex samples (from purified molecules to a multitude of model organisms at various stages of development, for example) that in turn change the requirements and implementation of the experiment. Thus, it does not come as a surprise that the experimental scheme and sample choice often have to be considered in parallel due to their interdependencies (Galas et al., 2018). Sample compatibility is a complex issue that comprises both the specimen and fluorescent labels. Likewise, the labeling strategy and sample viability are critically important factors to the success of an experiment, and these topics have been extensively discussed in the literature (Albrecht and Oliver, 2018; Dean and Palmer, 2014; Frigault et al., 2009; Heppert et al., 2016; Icha et al., 2017; Kiepas et al., 2020; Lambert, 2019; Schneider and Hackenberger, 2017; Specht et al., 2017; Thorn, 2017). Overall, the compatibility of a sample will be determined by all aspects of the experiment and demands careful consideration. As a result, the hypothesis and the associated experiment will be heavily influenced by what can be realistically achieved given the nature of the sample. Once the hypothesis has been appropriately defined, rather than proceeding directly to performing microscopy experiments, the most critical step at this point is to evaluate what it means to challenge the hypothesis.

Embarking on the journey of experimental design

Informative results

Not all results can adequately test a hypothesis. It is important to differentiate between a 'desired outcome' and an 'informative result'. The desired outcome would naturally be for the evidence to support the hypothesis. Continuing with the example of mitochondrial fission stated above, the informative result in this case would be the number of mitochondrial fission events as a function of time, both in the presence and absence of condition X. This is in contrast to the 'desired outcome' of finding an increased rate of mitochondrial fission given condition X. In addition, to be informative, the required data should encompass appropriate controls and sufficient replicates to support statistical analyses. The informative result is not designed to affirm one's intuition; it is required to support or negate the hypothesis.

Required data

As depicted in Fig. 1B, experimental design involves a reverse-thinking workflow that begins with informative results and concludes with microscope choice. This reverse-flow provides the necessary logic for designing a quantitative experiment. The essence of efficient experimental design is to home in on the appropriate assay from the multitude of possibilities offered by fluorescence microscopy. It is therefore imperative that the experimenter identifies what the necessary data are, as this will ultimately define the appropriate assay. This underscores the importance of thinking in reverse, as the necessary data can only be defined by informative results. While results and data are sometimes used interchangeably elsewhere, they are distinctly different in this context. Results refer to the final analytical metrics compiled from a set of related experiments. In contrast, a set of data generated from the microscope, by itself, is insufficient to speak to the validity of a hypothesis.

The transition from data to results requires certain translational steps. A good example of such translation is the process of connecting coordinates of a moving object, be it a cell or particle, between time points into a defined track. Without further analyses,

the tracked data of a moving object is only minimally informative; it merely indicates that the object has moved. If one were to hypothesize that the object would change its migratory behavior under certain conditions, then one would need to consider which measurements could describe that behavior. These informative measurements, when performed on the data, are referred to as the analytical metrics. In this example of characterizing migration patterns, the analytical metrics may include directionality, velocity and motion persistence (Aaron et al., 2019). Informative results are produced when these analytical metrics are applied to the appropriate data.

Adhering to our reverse-design approach, the types of analytical metrics that will lead to the informative results are the next factor an experimenter must consider. Table 1 shows how common biological objectives dictate the relevant analytical metrics, which in turn prescribe the necessary experimental tools. Analytical metrics is a form of semantics. What sets it apart from the semantics used in working models is that, in analytical metrics, the semantics are quantitative and specific rather than descriptive. What should be clear from Table 1 is that careful consideration is required to choose the appropriate analytical metrics. In fact, as reflected in the mitochondrial fission example, analytical metrics (mitochondrial fission rate) should be central to the hypothesis, so that it can be tested. An additional example where the choice of the correct analytical metric would affect the results is in colocalization studies. One must first determine whether measuring the degree of overlap (co-occurrence) of the two signals is more appropriate than measuring the extent of their correlation. This decision will dictate the analytical metric that should be used (Aaron et al., 2018). Likewise, if a certain treatment is postulated to increase the dissemination of cancer cells from a cell cluster, it is important, from a mechanistic standpoint, to properly frame the testable hypothesis. This can be accomplished by avoiding vague descriptions such as ‘dissemination’ and instead frame the descriptor in quantitative terms, such as velocity, directionality and persistence of the cellular movement (Aaron et al., 2019). This is how descriptive semantics should be translated into quantitative semantics, thereby enabling the underlying biology to be measured.

Interestingly, and perhaps ironically, many of the analytical metrics listed in Table 1, such as velocity, directionality, or curvature, collectively describe ‘spatial-temporal dynamics’. Yet, owing to various limitations of individual microscope design, it is impossible to capture them all in one experiment (see the section on microscope selection below). Similarly, it is often counter-productive to acquire more data than one needs, as this complicates data analysis and also compounds the problem of data storage (Andreev and Koo, 2020). Added complexity can lead to the experimenter being side-tracked from the original goal and makes data interpretation more difficult. Fig. 1C shows how the iterative evaluation of the experimental output will ensure that these readouts stay pertinent to the hypothesis and allow room for observational biology to take place. Parsimonious selection of analytical metrics will focus the scope of the experiment, generating data that can test the hypothesis. However, the well-considered selection of analytical metrics only fulfills half of the data requirement. One also needs to consider the validity of the data. In other words, how to ensure that the data set is accurate and reproducible.

Accuracy and reproducibility together describe the rigor of the experiment. While highly related, it is possible that accurate data are not reproducible, and reproducibility does not ensure accuracy (Payne-Tobin Jost and Waters, 2019). Too often, the accuracy and reproducibility of microscopy data is only an afterthought, which can potentially jeopardize an entire experiment. There are two places in which rigor can be compromised: during data generation and in the experimental design. Great care should be taken to ensure unbiased sampling, appropriate use of standards and controls, uniform instrument performance and consistent data processing pipelines. In this light, preserving accuracy and reproducibility during image acquisition has been extensively covered (Jonkman, 2020; McQuin et al., 2018; Payne-Tobin Jost and Waters, 2019), and is beyond the scope of our discussion. Nevertheless, this is extremely important advice and should be followed closely.

However, identifying the appropriate constraints for a rigorous experimental design can be equally challenging. How experimental controls and baselines are chosen can alter the data and the results, and therefore cannot be taken lightly as it can skew data interpretation. In stark contrast to physics, in which absolute numbers of various universal constants can be mathematically derived, biology is a comparative science. In biology, it is the change of experimental readouts in response to a modification of the experimental variables that is the important factor. As previously mentioned, modern microscopes will always generate quantifiable data because a digital image is intrinsically a data map. However, not all quantifiable digital images are meaningful. An absolute number derived from a colocalization experiment (for example, a calculated Pearson’s correlation coefficient of 0.75) between two proteins is quantitative, but utterly meaningless as a stand-alone piece

Table 1. Selecting analytical metrics based on biological questions

Objectives	General examples (descriptive semantics)	Analytical metrics (quantitative semantics)	Common tools
Characterizing biological motion	Organelle movement, cell locomotion	Velocity, directionality, persistence	Object tracking
	Saltatory movement or processivity of motor-driven movement	Velocity, directionality, persistence, diffusion constant	Particle tracking
	Molecular turnover in a compartment	Dissociation constant, diffusion constant	FRAP, photoconversion
Interpreting biological interaction/ signaling	Target protein in an organelle	Co-occurrence coefficient, correlation coefficient	Colocalization analyses
	In situ levels of protein modifications (e.g. phosphorylation)	Image ratio	Ratiometric imaging
	Microenvironment of target molecules	Anisotropy	Anisotropy measurement
Measuring shape/size of biological structures	Aspect ratio during cell spreading or cell differentiation	Shape descriptors	Morphometry
	Filopodial dynamics, membrane ruffles in 3D	Turnover, angular deflection, surface curvature	Deformation measurements

FRAP, fluorescence recovery after photobleaching.

Journal of Cell Science (2020) 133, jcs250027. doi:10.1242/jcs.250027

of data. It has to be compared to controls to become biologically informative – has the Pearson's coefficient changed in response to a variation in the experimental condition? The importance of establishing an experimental baseline for comparison cannot be overstated. Owing to our inherent tendency to look for the desired outcome, experimental bias occurs in the absence of a rigorous baseline. Validation of an experimental pipeline will ensure the measurements accurately represent the biological truth. This can be achieved by the effective use of controls and standards (Payne-Tobin Jost and Waters, 2019). While this sounds cliché, we found that comparative baselines are often forgotten. By articulating the necessary controls for a given hypothesis, the underlying nature of the experiment can become more apparent. This in turn can be used to refine the hypothesis and home in on what the biologist seeks to test. Stringent controls will indeed make for better experiments.

Parameters

When an experiment is driven by a hypothesis, the hypothesis itself will define the requirements of the experiment. These, in turn, will define the parameters that subsequently circumscribe the rest of the microscopy assay. The key parameters in any microscopy experiment will include one or more of the following: (i) lateral and axial spatial resolution, (ii) temporal resolution, (iii) tolerance to phototoxicity and photobleaching, (iv) field of view, (v) imaging depth, (vi) multiplexing capacity to acquire a combination of colors, and (vii) spectroscopic imaging capabilities. In a perfect world, a microscope will encompass all these parameters. Unfortunately, in reality, such a microscope does not exist as every microscope requires trade-offs (Combs, 2010; Lemon and McDole, 2020; Scherf and Huisken, 2015; Schermelleh et al., 2010). Occasionally, the trade-off can come at an exorbitant price, and this is especially the case with super-resolution microscopy. To gain the extra resolution, these modalities either completely sacrifice the capacity to image live phenomena or incur unacceptable doses of illumination light that rapidly induces phototoxicity (Schermelleh et al., 2019). Thus, the trade-off of an otherwise suitable microscope may render it incapable of producing the required data.

In order to avoid such situations, it is best to understand what needs to be captured by the microscope before selecting an instrument. This can be achieved by changing the ambiguous, descriptive semantics (e.g. 'membrane 3D dynamics') to those that are framed in the semantics of analytical metrics (e.g. 'filopodial angular deflection', 'membrane surface curvature') (see Table 1). By identifying the necessary metrics, the required imaging parameters can be prioritized. For example, the analytical metrics required to sufficiently measure the 3D membrane ruffles of a cell (Fritz-Laylin et al., 2017) include angular deflection, surface curvature, volumetric changes and the turnover rate of these membranous structures. These metrics will mandate the following imaging parameters: (i) high volumetric imaging speed (multiple volumes per min); (ii) improved axial resolution producing near or true isotropic resolution in all three axes, so that the ruffling structures can be resolved and segmented accurately; (iii) gentle illumination to minimize phototoxicity; (iv) live-cell-compatible imaging; and (iv) labeling of the cell membrane that is capable of facilitating the high number of image acquisitions. Box 1 also provides a case study of how analytical metrics influence microscope choice. Specific analytical metrics do not preclude the experimenter from observing (and even exploring) the biology; instead, they help winnow the imaging parameters down to the bare essentials. Together, quantitative metrics and experimental parameters will guide the user to the optimal microscope(s).

Microscope selection

The task of microscope selection can be bewildering to novices, and at times is confusing to even experienced microscopists. Biologists often face multiple hurdles in identifying suitable microscopes for an experiment through no fault of their own. These include (i) the lack of access to the desired instrument, (ii) ill-informed demand from reviewers to use the latest technology in the name of innovation, (iii) over-promise of instrument capabilities from the manufacturers, (iv) under-reporting of the instrument limitations, and (v) insufficient or erroneous reporting of published results that render experimental conditions irreproducible. Table 2 summarizes the features of various commonly used microscope modalities, as well as their relative advantages and shortcomings in our experience. Biologists have access to a wide range of modalities beyond standard widefield epifluorescence microscopes: total internal reflection fluorescence microscopy (Mattheyses et al., 2010),

Table 2. Performance comparison of various microscope modalities

	Imaging parameters or requirements					
	Lateral resolution	Imaging speed	Sample exposure	Live volumetric	Axial resolution	Imaging depth
Widefield epifluorescence	Good	Very good (2D)	Very good	Moderate	Moderate	Moderate
TIRF	Good	Very good (2D)	Excellent	N/A	N/A	N/A
Gaussian beam lightsheet	Good	Excellent (3D)	Excellent	Excellent	Moderate	Very good
Bessel beam lattice lightsheet	Good	Excellent (3D)	Excellent	Excellent	Good	Good
Spinning disk confocal	Good	Excellent (3D)	Good	Excellent	Moderate	Very good
Laser scanning confocal	Good	Good	Good	Good	Good	Very good
Two-photon fluorescence	Good	Good	Good	Good	Moderate	Excellent
Image scanning	Very good	Good	Very good	Very good	Good	Very good
SIM	Very good	Moderate	Moderate	Good	Very good	Moderate
STED	Excellent	Good	Poor	Good	Excellent	Very good
RESOLFT	Excellent	Moderate	Poor	Good	Excellent	Very good
SMLM	Excellent	Poor	Poor	Poor	Moderate	Moderate

Each block represents the relative qualitative capacity, in our experience, of each modality. The various imaging parameters or requirements cover a wide variety of uses and applications. Although other modalities exist, here we only highlight common, commercially available techniques. TIRF, total internal reflection fluorescence; SIM, structured illumination microscopy; STED, stimulated emission depletion; RESOLFT, reverse-saturable optical fluorescence transitions; SMLM, single-molecule localization microscopy; N/A, not applicable.

lightsheet microscopy (Chatterjee et al., 2018; Chen et al., 2014; Power and Huisken, 2017), confocal microscopy (Claxton et al., 2011; Conchello and Lichtman, 2005; Jonkman et al., 2020; Oreopoulos et al., 2014), two-photon excitation fluorescence microscopy (Benninger and Piston, 2013; So et al., 2000) and image scanning microscopy (Gregor and Enderlein, 2019), as well as super-resolution techniques (Demmerle et al., 2017; Sahl et al., 2017; Schermelleh et al., 2019; Sydor et al., 2015; Vicidomini et al., 2018). What should be immediately obvious from their comparison is that there is no 'winner' or 'loser' (Table 2). No microscope scores equally well or poorly across the various parameters, reinforcing the notion that every microscope compromises a combination of parameters in order to excel at others. As a result, the process of microscope selection is rarely linear. Many instruments have overlapping capabilities that obscure the selection process and will require that more than one instrument be considered at a time. By defining the required parameters beforehand, they can be used to filter the selection down to the most appropriate instrument(s), as exemplified in the case study presented in Box 1. The process of microscope selection is aided by a good understanding of the necessary imaging parameters. Ultimately, the justification for an instrument lies solely on the ability of that microscope to provide the necessary analytical metrics and the data informative of the biology.

It is impractical to expect biologists to understand the myriad of technical nuances of these rapidly evolving technologies. Likewise, most advanced imaging systems are usually concentrated in shared microscopy facilities, managed by experienced microscopists. This makes it all the more important for biologists to communicate, precisely and concisely, the desired analytical metrics and the corresponding parameters required for a successful experiment. It is sometimes difficult to appreciate that the latest imaging technology is not always the most appropriate. A super-resolution microscope or advanced lightsheet microscope may not necessarily be more suitable than a widefield epifluorescence microscope for a particular experiment. A microscope can only enhance certain parameters, and it is only beneficial if the enhanced parameters are utilized wisely. Even though structured illumination microscopy (SIM) offers improved resolution (see Table 2), it does not enhance the data of cell tracking studies over what can be achieved with a standard epifluorescence widefield microscope. It is also important to note that sometimes no single existing imaging technology may be able to produce the required data, necessitating the use of multiple instruments, or even the modification of the testable hypothesis. However, the availability of a new technology can open up the possibility of previously unfeasible analytical metrics that make it possible to address different biological queries.

Perspectives

The microcopy literature has no shortage of excellent reviews on the technical aspects of various imaging modalities, as well as tutorials on how to generate quantitative and reproducible data. However, topical discussion of best practices and optics does not necessarily engender a coherent framework of how these sets of information can be integrated to facilitate a hypothesis-driven, quantitative experimental design. Here, we present not only a roadmap of how to use these guides in the literature, but we break with the convention and argue that microscopy-based quantitative experiments should be designed in reverse, starting with determining the informative results needed to challenge a hypothesis.

Despite the promises of the latest technologies, no microscope is perfect. Usually, a feature gained in a technique comes at the cost of other key parameters. The essence of experimental design is never about the inclusion of every parameter the experimenter wants; rather it is about the careful exclusion of unnecessary parameters. This will allow accurate measurements to be performed and will ensure that the parameters relevant to the information the experimenter needs are maintained. This is the core concept of our approach. The essential parameters must be determined by what is required to test a hypothesis. These parameters will, in turn, naturally shape the rest of the elements of an experimental pipeline (Fig. 1A). A hypothesis-driven experimental design must be just that – driven by the hypothesis. It should be based on the biological question at hand, and not by the lure of the latest technologies. Fortunately, this process is an iterative feedback loop. The key questions left unanswered due to lack of technology inspire the development of novel microscopes. New technologies then reciprocally inform biology so that new hypotheses can be formulated. This cyclical process, however, does not negate the fact that experiments should be framed within the confines of existing technologies.

This Opinion article does not, by any means, diminish the exploratory power of microscopes and the well-honed acumen of biologists to observe and deduce. On the contrary, most hypotheses are synthesized following keen observation. The scope of this discussion is to focus on the process of quantitatively verifying a hypothesis. We have not addressed here how the power of modern microscopy has been harnessed in big-data scientific exploration. Such experiments are usually hypothesis-free; instead machine-learning algorithms are employed to search for patterns beyond what human perception can efficiently discern (Chessel and Carazo Salas, 2019; Piccinini et al., 2017).

Quantitative microscopy experiments are not easy to design, as they require knowledge at the confluence of optics, imaging probes, data analysis and how the biological samples interact with the microscope. It is therefore of paramount importance for biologists to seek and heed the advice of expert microscopists and data scientists, especially those in core facilities, who are experienced in the application of microscopy. The conventional practice of generating a lot of data first, followed by data analysis as a secondary consideration should be avoided. Microscopy-related experiments demand careful planning and continued, iterative evaluation before the optimal approaches can be implemented. The fact that this message is echoed in every review and guide cited here is because it is important, and unfortunately, because it is also commonly overlooked. The perils of ignoring it cannot be overstated.

Acknowledgements

We thank Dr Christopher Obara as well as the members of the Advanced Imaging Center for their thoughtful discussion and insightful contributions.

Competing interests

The authors declare no competing or financial interests.

Funding

The Advanced Imaging Center at Janelia Research Campus is generously supported by the Howard Hughes Medical Institute and the Gordon and Betty Moore Foundation.

References

Aaron, J. S., Taylor, A. B. and Chew, T.-L. (2018). Image co-localization – co-occurrence versus correlation. *J. Cell Sci.* **131**, jcs211847. doi:10.1242/jcs.211847

Aaron, J., Wait, E., DeSantis, M. and Chew, T. L. (2019). Practical considerations in particle and object tracking and analysis. *Curr. Protoc. Cell Biol.* **83**, e88. doi:10.1002/cpcb.88

Albrecht, R. M. and Oliver, J. A. (2018). Labeling considerations for confocal microscopy. In *Basic Confocal Microscopy* (eds W. Jerome and R. Price), pp. 99-134. Cham: Springer International Publishing.

Andreev, A. and Koo, D. E. S. (2020). Practical guide to storage of large amounts of microscopy data. *Micros. Today* **28**, 42-45. doi:10.1017/S1551929520001091

Benninger, R. K. P. and Piston, D. W. (2013). Two-photon excitation microscopy for the study of living cells and tissues. *Curr. Protoc. Cell Biol.* **59**, 4.11.1-4.11.24. doi:10.1002/0471143030.cb0411s59

Berg, S., Kutra, D., Kroeger, T., Straehle, C. N., Kausler, B. X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M. et al. (2019). ilastik: interactive machine learning for (bio)image analysis. *Nat. Methods* **16**, 1226-1232. doi:10.1038/s41592-019-0582-9

Chatterjee, K., Pratiwi, F. W., Wu, F. C. M., Chen, P. and Chen, B.-C. (2018). Recent progress in light sheet microscopy for biological applications. *Appl. Spectrosc.* **72**, 1137-1169. doi:10.1177/0003702818778851

Chen, B.-C., Legant, W. R., Wang, K., Shao, L., Milkie, D. E., Davidson, M. W., Janetopoulos, C., Wu, X. S., Hammer, J. A., Liu, Z. et al. (2014). Lattice light-sheet microscopy: imaging molecules to embryos at high spatiotemporal resolution. *Science* **346**, 1257998. doi:10.1126/science.1257998

Chessel, A. and Carazo Salas, R. E. (2019). From observing to predicting single-cell structure and function with high-throughput/high-content microscopy. *Essays Biochem.* **63**, 197-208. doi:10.1042/EBC20180044

Chew, T.-L., George, R., Soell, A. and Betzig, E. (2017). Opening a path to commercialization. *Opt. Photonics News* **28**, 42. doi:10.1364/OPN.28.7.000042

Claxton, N. S., Fellers, T. J. and Davidson, M. W. (2011). *Laser Scanning Confocal Microscopy (LSCM)*: Springer Reference.

Combs, C. A. (2010). Fluorescence microscopy: a concise guide to current imaging methods. *Curr. Protoc. Neurosci.* **50**, 2.1.1-2.1.14. doi:10.1002/0471142301.ns0201s50

Conchello, J.-A. and Lichtman, J. W. (2005). Optical sectioning microscopy. *Nat. Methods* **2**, 920-931. doi:10.1038/nmeth815

Courtot, M., Juty, N., Knüpfer, C., Waltemath, D., Zhukova, A., Dräger, A., Dumontier, M., Finney, A., Golebiewski, M., Hastings, J. et al. (2011). Controlled vocabularies and semantics in systems biology. *Mol. Syst. Biol.* **7**, 543. doi:10.1038/msb.2011.77

Dean, K. M. and Palmer, A. E. (2014). Advances in fluorescence labeling strategies for dynamic cellular imaging. *Nat. Chem. Biol.* **10**, 512-523. doi:10.1038/nchembio.1556

Demmerle, J., Innocent, C., North, A. J., Ball, G., Müller, M., Miron, E., Matsuda, A., Dobbie, I. M., Markaki, Y. and Schermelleh, L. (2017). Strategic and practical guidelines for successful structured illumination microscopy. *Nat. Protoc.* **12**, 988-1010. doi:10.1038/nprot.2017.019

Frigault, M. M., Lacoste, J., Swift, J. L. and Brown, C. M. (2009). Live-cell microscopy - tips and tools. *J. Cell Sci.* **122**, 753-767. doi:10.1242/jcs.033837

Fritz-Laylin, L. K., Riel-Mehan, M., Chen, B.-C., Lord, S. J., Goddard, T. D., Ferrin, T. E., Nicholson-Dykstra, S. M., Higgs, H., Johnson, G. T., Betzig, E. et al. (2017). Three-dimensional actin-based protrusions of migrating neutrophils are intrinsically lamellar and facilitate direction changes. *eLife* **6**, e26990. doi:10.7554/eLife.26990

Galas, L., Gallavardin, T., Bénard, M., Lehner, A., Schapman, D., Lebon, A., Komuro, H., Lerouge, P., Leleu, S. and Franck, X. (2018). "Probe, Sample, and Instrument (PSI)": the hat-trick for fluorescence live cell imaging. *Chemosensors* **6**, 40. doi:10.3390/chemosensors6030040

Gregor, I. and Enderlein, J. (2019). Image scanning microscopy. *Curr. Opin. Chem. Biol.* **51**, 74-83. doi:10.1016/j.cbpa.2019.05.011

Grimm, J. B., Muthusamy, A. K., Liang, Y., Brown, T. A., Lemon, W. C., Patel, R., Lu, R., Macklin, J. J., Keller, P. J., Ji, N. et al. (2017). A general method to fine-tune fluorophores for live-cell and in vivo imaging. *Nat. Methods* **14**, 987-994. doi:10.1038/nmeth.4403

Heppert, J. K., Dickinson, D. J., Pani, A. M., Higgins, C. D., Steward, A., Ahringer, J., Kuhn, J. R. and Goldstein, B. (2016). Comparative assessment of fluorescent proteins for in vivo imaging in an animal model system. *Mol. Biol. Cell* **27**, 3385-3394. doi:10.1091/mbc.e16-01-0063

Icha, J., Weber, M., Waters, J. C. and Norden, C. (2017). Phototoxicity in live fluorescence microscopy, and how to avoid it. *BioEssays* **39**, 1700003. doi:10.1002/bies.201700003

Jonkman, J. (2020). Rigor and reproducibility in confocal fluorescence microscopy. *Cytom. Part A* **97**, 113-115. doi:10.1002/cyto.a.23924

Jonkman, J., Brown, C. M. and Cole, R. W. (2014). Quantitative confocal microscopy. In *Methods in Cell Biology* (eds Jennifer C. Waters and Torsten Wittman), pp. 113-134. Academic Press Inc.

Jonkman, J., Brown, C. M., Wright, G. D., Anderson, K. I. and North, A. J. (2020). Tutorial: guidance for quantitative confocal microscopy. *Nat. Protoc.* **15**, 1585-1611. doi:10.1038/s41596-020-0313-9

Kiepas, A., Voorand, E., Mubaid, F., Siegel, P. M. and Brown, C. M. (2020). Optimizing live-cell fluorescence imaging conditions to minimize phototoxicity. *J. Cell Sci.* **133**, jcs242834. doi:10.1242/jcs.242834

Lambert, T. J. (2019). FPbase: a community-editable fluorescent protein database. *Nat. Methods* **16**, 277-278. doi:10.1038/s41592-019-0352-8

Lee, J.-Y. and Kitaoka, M. (2018). A beginner's guide to rigor and reproducibility in fluorescence imaging experiments. *Mol. Biol. Cell* **29**, 1519-1525. doi:10.1091/mbc.E17-05-0276

Lemon, W. C. and McDole, K. (2020). Live-cell imaging in the era of too many microscopes. *Curr. Opin. Cell Biol.* **66**, 34-42. doi:10.1016/j.ceb.2020.04.008

Mattheyses, A. L., Simon, S. M. and Rappoport, J. Z. (2010). Imaging with total internal reflection fluorescence microscopy for the cell biologist. *J. Cell Sci.* **123**, 3621-3628. doi:10.1242/jcs.056218

McArthur, K., Whitehead, L. W., Heddleston, J. M., Li, L., Padman, B. S., Oorschot, V., Geoghegan, N. D., Chappaz, S., Davidson, S., Chin, H. S. et al. (2018). BAK/BAX macropores facilitate mitochondrial herniation and mtDNA efflux during apoptosis. *Science* **359**, eaao6047. doi:10.1126/science.aao6047

McQuin, C., Goodman, A., Chernyshev, V., Kamentsky, L., Cimini, B. A., Karhohs, K. W., Doan, M., Ding, L., Rafelski, S. M., Thirstrup, D. et al. (2018). CellProfiler 3.0: next-generation image processing for biology. *PLoS Biol.* **16**, e2005970. doi:10.1371/journal.pbio.2005970

North, A. J. (2006). Seeing is believing? A beginners' guide to practical pitfalls in image acquisition. *J. Cell Biol.* **172**, 9-18. doi:10.1083/jcb.200507103

Oreopoulos, J., Berman, R. and Browne, M. (2014). Spinning-disk confocal microscopy. present technology and future trends. In *Methods in Cell Biology* (eds Jennifer C. Waters and Torsten Wittman), pp. 153-175. Academic Press Inc.

Payne-Tobin Jost, A. and Waters, J. C. (2019). Designing a rigorous microscopy experiment: validating methods and avoiding bias. *J. Cell Biol.* **218**, 1452-1466. doi:10.1083/jcb.201812109

Piccinini, F., Balassa, T., Szkalisity, A., Molnar, C., Paavolainen, L., Kujala, K., Buzas, K., Sarazova, M., Pietiainen, V., Kutay, U. et al. (2017). Advanced cell classifier: user-friendly machine-learning-based software for discovering phenotypes in high-content imaging data. *Cell Syst.* **4**, 651-655.e5. doi:10.1016/j.cels.2017.05.012

Popper, K. (2005). *The Logic of Scientific Discovery*. Routledge.

Power, R. M. and Huisken, J. (2017). *A guide to Light-Sheet Fluorescence Microscopy for Multiscale Imaging*. Nature Publishing Group.

Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T. and Eliceiri, K. W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics* **18**, 529. doi:10.1186/s12859-017-1934-z

Sahl, S. J., Hell, S. W. and Jakobs, S. (2017). Fluorescence nanoscopy in cell biology. *Nat. Rev. Mol. Cell Biol.* **18**, 685-701. doi:10.1038/nrm.2017.71

Scherf, N. and Huisken, J. (2015). The smart and gentle microscope. *Nat. Biotechnol.* **33**, 815-818. doi:10.1038/nbt.3310

Schermelleh, L., Heintzmann, R. and Leonhardt, H. (2010). A guide to super-resolution fluorescence microscopy. *J. Cell Biol.* **190**, 165-175. doi:10.1083/jcb.201002018

Schermelleh, L., Ferrand, A., Huser, T., Eggeling, C., Sauer, M., Biehlmaier, O. and Drummen, G. P. C. (2019). Super-resolution microscopy demystified. *Nat. Cell Biol.* **21**, 72-84. doi:10.1038/s41556-018-0251-8

Schneider, A. F. L. and Hackenberger, C. P. R. (2017). Fluorescent labelling in living cells. *Curr. Opin. Biotechnol.* **48**, 61-68. doi:10.1016/j.copbio.2017.03.012

Shaner, N. C., Lambert, G. G., Chammas, A., Ni, Y., Cranfill, P. J., Baird, M. A., Sell, B. R., Allen, J. R., Day, R. N., Israelsson, M. et al. (2013). A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum. *Nat. Methods* **10**, 407-409. doi:10.1038/nmeth.2413

Shasha, D. E. (2003). Plant systems biology: lessons from a fruitful collaboration. *Plant Physiol.* **132**, 415-416. doi:10.1104/pp.102.019588

So, P. T. C., Dong, C. Y., Masters, B. R. and Berland, K. M. (2000). Two-photon excitation fluorescence microscopy. *Annu. Rev. Biomed. Eng.* **2**, 399-429. doi:10.1146/annurev.bioeng.2.1.399

Specht, E. A., Braselmann, E. and Palmer, A. E. (2017). A critical and comparative review of fluorescent tools for live-cell imaging. *Annu. Rev. Physiol.* **79**, 93-117. doi:10.1146/annurev-physiol-022516-034055

Swedlow, J. R. (2013). Quantitative fluorescence microscopy and image deconvolution. In *Methods in Cell Biology* (eds Greenfield Sluder and David E. Wolf), pp. 407-426. Elsevier Inc.

Sydor, A. M., Czymmek, K. J., Puchner, E. M. and Mennella, V. (2015). Super-resolution microscopy: from single molecules to supramolecular assemblies. *Trends Cell Biol.* **25**, 730-748. doi:10.1016/j.tcb.2015.10.004

Thorn, K. (2017). Genetically encoded fluorescent tags. *Mol. Biol. Cell* **28**, 848-857. doi:10.1091/mbc.e16-07-0504

Van Den Berge, K., Hembach, K. M., Soneson, C., Tiberi, S., Clement, L., Love, M. I., Patro, R. and Robinson, M. D. (2019). RNA sequencing data: Hitchhiker's guide to expression analysis. *Annu. Rev. Biomed. Data Sci.* **2**, 139-173. doi:10.1146/annurev-biodatasci-072018-021255

Vicidomini, G., Bianchini, P. and Diaspro, A. (2018). STED super-resolved microscopy. *Nat. Methods* 15, 173-182. doi:10.1038/nmeth.4593

Waters, J. C. (2009). Accuracy and precision in quantitative fluorescence microscopy. *J. Cell Biol.* **185**, 1135-1148. doi:10.1083/jcb.200903097

Weigert, M., Schmidt, U., Boothe, T., Müller, A., Dibrov, A., Jain, A., Wilhelm, B., Schmidt, D., Broaddus, C., Culley, S. et al. (2018). Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat. Methods* **15**, 1090-1097. doi:10.1038/s41592-018-0216-7