SACRED BOVINES

THE VERY REPRODUCIBLE (BUT ILLUSORY) MESOSOME

DOUGLAS ALLCHIN, DEPARTMENT EDITOR

Science, we hear, is facing a Reproducibility Crisis. Famous psychological experiments have been redone but did not yield the same results. Replications of clinical studies on drugs and cancer therapies have "failed," too. Ultimately, the fear is that we cannot trust scientific claims.

Underlying this skepticism is the concept of reproducibility. Researchers expect, on principle, that if one follows the same procedures and gets the same results, somehow the conclusions are also secure. For example, an editor of *Science* magazine asserted, "Science advances on a foundation of trusted discoveries. Reproducing an experiment is one important approach that scientists use to gain confidence in their conclusions" (McNutt, 2014). A later editor echoed her sentiment, "The ability to test validity by replicating experiments and comparing results is a cornerstone of science" (Berg, 2019). Elsewhere: "replication—the confirmation of results and conclusions from one study obtained independently in another—is considered the scientific gold standard" (Jasny et al., 2011).

Here, I reconsider this widely held belief (this month's Sacred Bovine) that reproducibility is the primary factor in ensuring that scientific conclusions are reliable. Mere repetition alone does not establish valid conclusions. Illusions, too (ironically), are very reproducible. Rather, interpretation is essential. Additional observations—for example, controlled experiments—help clarify and contextualize what one is observing. They help us separate fact from coincidental but meaningless artifact (see also Sacred Bovines, March 2020). Context and error analysis, not replication, yields trustworthy conclusions.

○ "Discovering" Mesosomes

Consider the case of the mesosome (Culp, 1994; Hudson, 1999; Rasmussen, 1993). The development of the electron microscope (EM) in the first half of the 20th century enabled higher resolution, yielding unprecedented views of the cell. "New" structures—previously unobserved—became visible. One such structure was the bacterial mesosome, first imaged in 1953: a spiral shaped membrane near the edge of the cell (Figure 1).

But were mesosomes "*real*"? Cytologists certainly knew that microscopic images could be misleading. Early microscopists in the 19th century, for example, discovered how lenses refract light and introduce chromatic or spherical aberrations. Chemical stains highlighted some structures but hid others. So, one might imagine, to safeguard against error and regard mesosomes as authentic, investigators needed to reproduce them.

Cytologists thus compared the new EM images against observations of similar samples using the more familiar light

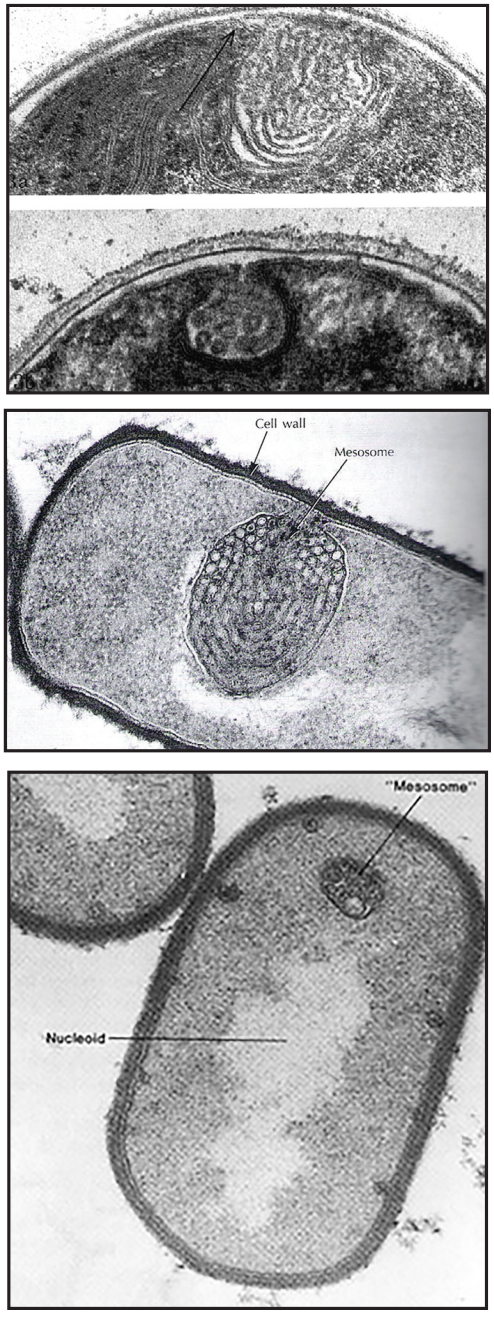

Figure 1. Electron micrographs of mesosomes. Is seeing believing? Mesosomes are artifacts of preparation techniques, not real structures in native bacteria. (From Allen, 1972; muhadharaty.com; vrchemistry.chem.ox.ac.uk.)

The American Biology Teacher, Vol. 84, No. 5, pp. 321–323, ISSN 0002-7685, electronic ISSN 1938-4211. DOI: https://doi.org/10.1525/abt.2021.84.5.321

microscope: a form of calibration. They varied the preparation procedures to see if the mesosome phenomenon was stable. It appeared to be robust. Not that mesosomes did not depend on how samples were prepared. Microscopists varied the presence of sucrose, glycerol or calcium ions, prefix time, temperature, form of protection against freezing, fixative, and method of viewing (e.g., thin section vs. freeze-fracture). They gradually learned the optimal conditions for producing reliably "good" mesosomes in contrast to "poor" mesosomes. Mesosomes were (and are) quite reproducible. Did that truly "test validity" in the apparent discovery of a new organelle?

Beginning in the early 1960s, textbooks featured pictures and diagrams of mesosomes. Still, their role was a mystery. Were they related to cell division? To oxidative phosphorylation? To membrane biogenesis? To secretory processes? Or to something else? Biochemists joined in to help analyze their function. By 1975, the status of mesosomes was summarized in an extensive 58-page review. But their function was still unknown (Greenwalt & Whiteside, 1975).

Alas, mesosomes are now regarded as ***artifacts***. That is, they are ***created*** by the preparation techniques. They are not merely made visible by them (Ebersold et al., 1981). Despite the micrographic images, mesosomes are not native structures in bacteria. Even if they are reproducible. A puzzle for science's supposed "gold standard," perhaps?

○ From Fact to Artifact

How was the error detected, then? Reservations about the authenticity of mesosomes emerged in several labs beginning in the late 1960s and early 1970s. Nanne Nanninga had been checking the relatively new freeze-etch (now freeze-fracture) technique. He encountered problems with another structure, the nucleoplasm. He resolved the ambiguities using comparisons from phase-contrast light microscopy on living cells (Nanninga, 1971). But for mesosomes, discrepancies between old and new techniques persisted. He isolated one difference to the use of the prefixative osmium tetroxide (OsO_4). But here (he acknowledged) that you could only see mesosomes using OsO_4, so how could you independently test how it possibly distorted the image? M. Silva (1971) echoed these concerns. Citing inconsistencies from different methods he, too, adopted a new posture of uncertainty. He used visible differences as evidence. Both investigators began to dislodge anchors that established the ***meaning*** of the cryptic structure in the micrographs.

But these were just suspicions. One needed to explain instead what one ***was*** observing. Namely, just how were mesosomes created? Possible explanations followed quickly. Nanninga (1973) hypothesized an enlargement of small membrane features due to chemical or physical impairment, although without offering any detailed mechanism. He cited the shape of mesosomes, their placement, and the failure of anyone over more than a decade to identify a clear function. Michael Higgins' lab, by contrast, documented experimentally that glutaraldehyde could cross-link proteins. They proposed that the fixative caused small, peripheral membrane units to coalesce into one, oversized mesosome (Higgins et al., 1976). Margrit Fooke-Achterrath and her group (1974) again linked large mesosomes to OsO_4. They also showed that chilling could prevent this, while admitting that "the precise mechanism by which the artifacts arise is unknown" (p. 282). All these researchers explained large mesosomes only.

Silva's lab (1976), however, went deeper. They characterized the exact conditions that produced mesosomes. Those conditions significantly altered the cells from their original state. First, they showed that the longer the exposure to OsO_4, the larger and more numerous the mesosomes. One could virtually track their development. Further, they proposed that OsO_4 damaged membranes. This they demonstrated using OsO_4 on simple vesicles without cell walls (protoplasts). They also measured the diffusion of potassium ions to show that, independently of any microscopy technology, there was membrane damage. They also considered other chemicals besides OsO_4. Only those that damaged membranes generated mesosomes. Debate on mesosomes continued for at least another decade. Yet the ultimate resolution echoed Silva's 1976 sketch. Mesosomes are "real," perhaps, but they appear only when the bacterial membrane is damaged in preparing cells for EM. They have not been found in untreated cells.

In summary, then, fixatives (osmium tetroxide, in particular) were correlated with the observations of mesosomes. Increasing the time of fixation yielded larger mesosomes. The fixative damaged the permeability of the membrane. Mesosomes were (are) a product of artificial cross-linkage between membrane surfaces. Understanding mesosomes depended on articulating the context of the images—a constellation of data and experimental controls—not merely on being able to replicate them.

Mesosomes are a fact, perhaps—although an uninteresting or irrelevant fact. They exist. Under certain conditions, at least. But they are not genuine, or "real," structures in bacteria.

Yet at the same time they ***are*** perfectly reproducible. Unfortunately, as molecular biologist Walter Gilbert once cautioned, "you can reproduce artifacts very, very well" (quoted in Judson, 1981, p. 170). Ironically, reliance on reproducibility hid, rather than exposed, mesosomes as an error.

○ The Status of Reproducibility

Recall, now, all the declarations about the centrality of reproducibility in science. How essential is this principle? Efforts at replication are certainly common in science. But why? Most importantly, perhaps, for one to build on innovative findings, one needs to learn and master the relevant methods. Repeating earlier work is thus integral pragmatically to stepwise progress. Such replication is not really about checking for errors or validating conclusions, however.

Ironically, perhaps, failure to replicate is also common. In a recent survey, more than 70% of researchers reported having tried and failed to replicate someone else's experiment (Baker, 2016). More poignantly, perhaps, over half had tried—and failed—to replicate ***their own*** experiments! (Of course, this tells us nothing of the successes.) The implicit expectation that replication is (or should be) a definitive check is misinformed.

Consider the historical case of Joseph Priestley's discovery of the "restoration of air" by plants in 1771 (see Sacred Bovines, September 2012). In today's terms, Priestley demonstrated that photosynthesis by sprigs of mint or other herbs generates oxygen. Many contemporaries were eager to witness the remarkable result for themselves. Yet their efforts often went unrewarded. Later, even Priestley himself failed. Ironically here, despite the ***lack*** of reproducibility, the initial findings were indeed correct. Priestley had missed noticing an important variable: light from a nearby window. That was all sorted out in subsequent investigations. But the confusion illustrates a problematic dilemma when replications "fail": Is the original at fault,

or the attempted copy? If key variables are overlooked and hence go unreported, other labs will be stymied by an "incomplete recipe." Efforts to mimic the original may go astray in other ways. Reagents may change subtly, but significantly. The subjects of study—genetic strains, perhaps, or sample populations—may not be exactly parallel. Conditions that at first seem similar, may ultimately not be. Namely, faithful replication is far from simple (Bryan et al., 2019; Gilbert et al., 2016; Nosek & Errington, 2017).

How, then, should we interpret the conspicuous "failures" in recent years to replicate many landmark studies (e.g., Open Science Collaboration, 2015; Kaiser, 2021)? One finds that on many occasions, the original protocols were altered or adapted (to reduce cost, to accommodate local conditions, convenience of data analysis, and so on). The original investigators did not always view the revised study design as sufficiently similar. When that endorsement was absent, the proportion of successful replications plummeted by roughly 75% (Gilbert et al., 2016). Replication is not easy.

Statistical analysis is another stumbling block. Ironically, if one relies on a *p*-value of 0.05 (the conventional cutoff), one may *expect* failures 1 out of every 20 times. In other cases, the sample size of the original study, or of the replication, is too small. It lacks the statistical power to be a true test. All these factors help mitigate concerns that a full-fledged crisis is signaling the demise of science. The remarkably rapid research on the coronavirus, its treatment and vaccines in recent years demonstrates that science still seems to be progressing unimpaired.

The challenges of replication, however, underscore that the near-mythic "scientific method" (so often promulgated to students in classrooms) is simplistic and woefully misleading. The notion that experiments are uncomplicated either-or tests, unambiguous and easily confirmed through repetition, is utterly unfounded. As vividly exemplified in the mesosome case, science depends on multiple studies and an extended process of reciprocal criticism. Science education needs to engage students with these complexities to nurture reasonable expectations about how science works and when.

Trust in science is easily damaged by sensational claims or colorful anecdotes. The flurry of worry known as the Reproducibility Crisis seems based on an idealized notion of science, not on intimate understanding of actual scientific practices. The story of the ill-fated mesosome and how it illustrates the limits of replication may potently remind teachers, yet again, of the importance of teaching the nature of science and of the virtues of historical cases in interpreting today's scientific practices.

References

Allen, M.M. (1972). Mesosomes in blue-green algae. *Archiv für Mikrobiologie*, *84*, 199–206.

Baker, M. (2016). Is there a reproducibility crisis? *Nature*, *533*, 452–54.

Berg, J. (2019). Replication challenges. *Science*, *365*, 957.

Bryan, C.J., Yeager, D.S. & O'Brien, J.M. (2019). Replicator degrees of freedom allow publication of misleading failures to replicate. *Proceedings of the National Academics of Science, 116*, 25535–45.

Culp, S. (1994). Defending robustness: The bacterial mesosome as a test case. In D. Hull, M. Forbes & R.M. Burian (Eds.), *PSA 1994* (Vol. 1, pp. 46–57). Philosophy of Science Association.

Ebersold, H.R., Cordier, J.-L. & Lüthy P. (1981). Bacterial mesosomes: Method dependent artifacts. *Archives of Microbiology*, *130*, 19–22.

Fooke-Achterrath, M., Lickfield, K.G., Reusch, V.M., Jr., Aebi, U., Tschöpe, U. & Menge, B. (1974). Close-to-life preservation of *Staphylococcus aureus* mesosomes for transmission electron microscopy. *Journal of Ultrastructure Research*, *49*, 270–85.

Gilbert, D.T., King, G., Pettigrew, S. & Wilson, T.D. (2016). Comment on "Estimating the reproducibility of psychological science." *Science*, *351*, 1037–38.

Greenwalt, J.W. & Whiteside, T.L. (1975). Mesosomes: Membranous bacterial organelles. *Bacteriological Reviews*, *39*, 405–63.

Higgins, M.L., Tsien, H.C. & Daneo-Moore, L. (1976). Organization of mesosomes in fixed and unfixed cells. *Journal of Bacteriology*, *127*, 1519–23.

Hudson, R.G. (1999). Mesosomes: A study in the nature of experimental reasoning. *Philosophy of Science*, *66*, 289–309.

Jasny, B.R., Chin, G., Chong, L. & Vignieri, S. (2011). Again, and again, and again … *Science*, *334*, 1225.

Judson, H.F. (1981). *The Search for Solutions*. Holt, Rinehart, Winston.

Kaiser, J. (2021). Key cancer results failed to be reproduced. *Science*, *374*, 1311.

McNutt, M. (2014). Reproducibility. *Science*, *343*, 229.

Nanninga, N. (1971). The mesosome of *Bacillus subtilis* as affected by chemical and physical fixation. *Journal of Cell Biology*, *48*, 219–24.

Nanninga, N. (1973). Freeze-fracturing of microorganisms. Physical and chemical fixation of *Bacillus subtilis*. In E.L. Benedetti and P. Favard (Eds.), *Freeze-Etching Techniques and Applications*. Paris: Société Française de Microscopie Electronique, 151–79.

Nosek, B.A. & Errington, T.M. (2017). Reproducibility in cancer biology: Making sense of replications. *eLife*, *6*, 23383.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716.

Rasmussen, N. (1993). Facts, artifacts and mesosomes: Practicing epistemology with the electron microscope. *Studies in the History and Philosophy of Science*, *24*, 227–65.

Silva, M.T. (1971). Changes induced in the ultrastructure of cytoplasmic and intracytoplasmic membranes of several gram-positive bacteria by variations in OsO_4 fixation. *Journal of Microscopy*, *93*, 227–32.

Silva, M.T., Sousa, J.C.F., Polónia, J.J., Macedo, M.A.E. & Parente, A.M. (1976). Bacterial mesosomes: Real structures or artifacts? *Biochimica et Biophysica Acta*, *443*, 92–105.

DOUGLAS ALLCHIN (allchindouglas@gmail.com) taught both high school and college biology before receiving his PhD in history and philosophy of science. He is a Fellow at the Minnesota Center for the Philosophy of Science, University of Minnesota, Minneapolis, MN 55455. His collected essays from this column are available in *Sacred Bovines* from Oxford University Press.