Trends in Cell Biology

CellPress

Scientific Life

Interdisciplinary Team Science in Cell Biology

Rick Horwitz[1,*]

The cell is complex. With its multitude of components, spatial–temporal character, and gene expression diversity, it is challenging to comprehend the cell as an integrated system and to develop models that predict its behaviors. I suggest an approach to address this issue, involving system level data analysis, large scale team science, and philanthropy.

> 'The area of scientific investigation has been enormously extended.. But the assimilative power of the human intellect is... strictly limited. Hence, it was inevitable that the activity of the individual investigator should be confined to smaller and smaller section... as a result of this specialization, it is becoming increasingly difficult for even a rough general grasp of science as a whole, without which the true spirit of research is inevitably handicapped...'
> - Albert Einstein, *The World As I See It* (1949)

This quote, written late in Einstein's career, arguably also describes the present state of cell biology, sitting at a time of greatly increased understanding and enormous momentum. Research in cell biology is poised to deepen our understanding of the mechanisms that underlie the plethora of activities comprising the various cellular behaviors that dictate cell types and states. While opportunity abounds for continued great discovery and understanding, the increased specialization that will accompany the deepening investigation presents a consequent challenge in comprehending and integrating this information. In this perspective, I discuss a possible approach to address this challenge.

Cell biology, beginning largely as microscopic observations, followed the molecular biology revolution, which viewed genes, cells, and the machinery that underlies their activities as molecular systems that could be fully characterized and understood using methods of genetics and biochemistry. Viewing the cell as a complex, dynamic molecular composite brought insights from chemistry and physics to bear on biological problems. Just as the molecular genetic era was codified by the publication of Watson's book, *Molecular Biology of the Gene* [1], two decades later the *Molecular Biology of the Cell* by Alberts, *et al.* [2] served a similar purpose for cell biology.

Currently, cell biology stands in an enviable position. The Human Genome Project[i] and related activities have produced a list of molecular players, fueling the widespread use of genetic tools and producing an increasingly meaningful and useful understanding of cellular processes and their regulation. Dazzling new light and electron microscopic methods are now allowing the elucidation of complex structures in situ [3,4]. A plethora of 'omics' data on gene expression, post-translational modifications, proteins, metabolomes, and others are revealing important complex relationships and connections. Finally, the effects of the microenvironment and mutations on these processes, particularly in the context of disease, are revealing insights into basic cell biology and its misregulation [5,6].

These new approaches and the consequent rapid increase in knowledge are generating immense opportunities for individual investigators; but they are also revealing the need for theories and quantitative computational models that integrate and conceptualize this knowledge and predict cellular behaviors in response to mutation and environmental alterations. This requires quantitative data and a spectrum of modeling approaches across a range of scales - from machine learning, correlation and systems approaches to detailed physical–chemical mechanisms [7]. The data required for these models are now in sight. New gene editing methods are providing endogenous expression of tagged and mutant cells [8], and new live-cell imaging methods are promising biochemistry in living cells, measuring concentrations, dynamics, equilibria, and organization [9]. Similarly, super-resolution microscopy and cryoEM tomography, which allow structure determination and organization in situ [3,4], imaging mass spectrometry [10], and single-cell and spatially-resolved genomic approaches [11–13], among other image-based technologies, all point to a new golden era of cell biology where measurements can be performed in cells while still preserving the spatial (and often temporal) character of the cell.

Computational modeling benefits from an iterative cycle of model and experiment, ideally executed by integrated, interactive collaborations or teams. Furthermore, some of the more integrative research may require large data sets and a focus, at least initially, on 'discovery-driven science'. These challenges and complex technologies point to a need for interdisciplinary teams. This way of performing research may require new funding mechanisms, since their scope, absence of clear hypotheses, and possibly the need for large, highly integrated teams, does not readily fit into most NIH funding mechanisms for basic cell biology. How can cell biology foster this new research environment?

Philanthropy is one answer and is playing an increasing role in supporting science. In most cases, it reflects the interests and style of the donor. Many philanthropies facilitate research activities focused in a particular area, and they collaborate with the academic world to help move important cutting edge fields forward. Several successful examples of philanthropic and academic relationships exist, including the Broad Institute for studying genomics and the Wyss Institute for investigating

translational bioengineering. In addition, the Howard Hughes Medical Institute (HHMI) was founded to support promising investigators in long-term, high-risk/high-reward research. The HHMI has a large footprint, and it selects and provides major laboratory funding for over 300 academic investigators in a variety of scientific areas across the United States. This funding has greatly facilitated research in costly areas such as investigations focused on structure or genetic screens, and in specific areas, such as neuroscience through its efforts at Janelia Farm. Other funding mechanisms focus on similar efforts as the HHMI, including those supported by Packard, Pew, MacArthur, the Burroughs–Wellcome Fund, and The Paul G. Allen Frontiers Group. In addition to these trans-institution mechanisms, there are a large number of institutes associated with specific universities, often aimed at advancing a specific mission by comingling scientists in a single location.

The newly-formed Allen Institute for Cell Science complements these other institutes in style, mission, and approach. The program has aspirational goals and clearly defined milestones. The overarching mission is to understand and predict cellular behaviors in normal, pathological, and regenerative contexts. Its style is to engage in 'industrial-scale' basic research that is open and community-empowering, but cannot be performed easily in a typical academic environment.

However, the Institute is working closely with the cell biology community. Many of the Institute's activities have been informed and shaped by formal and informal conversations with members of the cell biology and allied communities. Open sharing of tools, reagents (cell lines), and data are intrinsic to the Institute's activities. Since the goal of the Institute is to characterize and integrate a range of cellular interactions, rather than to deeply investigate a single process, the Institute will add a new 'scale' or layer to our understanding of the cell and facilitate new multiscale modeling approaches. In this way the spirit of the Institute is to empower, rather than compete with individual investigators.

The Institute's first project takes an integrative approach by developing dynamic, visual data on the organization and localized activities in human induced pluripotent stem cells (hiPSCs) and cardiomyocytes derived from them. The choice to study hiPSCs was made because they are diploid, 'disease in a dish' cell and tissue/organ models that possess relatively stable genomes, propagate well, can be induced to differentiate into a number of different cell types, and can be used to exploit the vast and rapidly increasing knowledge of the human genome [14,15].

The approach is highly integrated. The immediate goal is to use gene editing to develop a set of well characterized hiPSCs expressing fluorescence tags on proteins that identify the locations of major molecular structures (molecular machines, organelles, and regulatory complexes) found in most cells. These cells are also being studied from the undifferentiated state to their differentiation into cardiomyocytes, which is a robust, relatively fast, and reproducible transition. The cells and organoids derived from hiPSCs are being imaged using a semiautomated pipeline consisting of spinning disk and super-resolution modalities. The data are produced in the context of computational models that will predict positions of the major structures and activities, derive correlations from changes in positions, and develop mechanisms for their repositioning. The goal is to produce high replicate measurements(hundreds to thousands), producing the mean and variance of the natural variation in the relative positions for the major cellular components. The measurements will also be made on cells residing in different environments, in response to perturbations, for example, drugs and mutation, and during changes in cell cycle and differentiation into cardiomyocytes. Statistical analyses of these data will reveal coupling among different organelles as well as principles and algorithms for cell organization. These correlations will generate hypotheses that can then be tested, modeled, and further investigated to determine the mechanisms by which organelles change positions during cellular reorganization. In addition, the knowledge of organelle positions and the timing of their local activities will enable modeling at many scales, from initial correlations to physical–chemical mechanisms. An 'animated cell' (a cellular 'Google Earth') is a key output for the Institute. The large numbers of replicate measurements allow us to develop positional statistics, which can be visualized in data-driven animations. Similarly, since only a few colors can be expressed and imaged in a single cell, data from different experiments will need to be integrated, using computation-derived positional referencing. These outputs as well as those from other modeling activities can be presented as data-driven animations.

Is large scale, team, or philanthropically-sponsored research the future of cell biology? Will interdisciplinary or large teams with focused goals or resource-rich centers replace the traditional, individual-initiated research that has served us so well? Likely not, because small focused research enterprises addressing specific process are the lifeblood of research, effectively crowdsourcing discovery, leading to deeper understanding, and fueling translation. But questions and issues remain that are not easily addressed by individuals or small collaborations. Therefore, I see a mix of approaches and ways of studying the cell in the future, with investigator-initiated research being in the majority. The need for other approaches, however, is becoming increasingly important. While investigator-driven research has evolved to work well to address specific hypotheses and problems, the evolution of large-scale team science is still a relatively new frontier and successful models for its effective implementation are still being developed.

Resources

[i] www.genome.gov/10001772/all-about-the–human-genome-project-hgp/

[1]Allen Institute for Cell Science, Seattle, WA, USA

*Correspondence: rickh@alleninstitute.org (R. Horwitz).

http://dx.doi.org/10.1016/j.tcb.2016.07.007

References

1. Watson, J.D. (1965) *Molecular Biology of the Gene,* Benjamin
2. Alberts, B. *et al.* (1983) *Molecular Biology of the Cell,* Garland Science
3. Liu, Z. *et al.* (2015) Imaging live-cell dynamics and structure at the single-molecule level. *Mol. Cell.* 58, 644–659
4. Asano, S. *et al.* (2016) *In situ* cryo-electron tomography: a post-reductionist approach to structural biology. *J. Mol. Biol.* 428, 332–343
5. Nelson, C.M. and Bissell, M.J. (2006) Of extracellular matrix, scaffolds, and signaling: tissue architecture regulates development, homeostasis, and cancer. *Annu. Rev. Cell. Dev. Biol.* 22, 287–309
6. Engler, A.J. *et al.* (2009) Multiscale modeling of form and function. *Science* 324, 208–212
7. Ideker, T. and Lauffenburger, D. (2003) Building with a scaffold: emerging strategies for high- to low-level cellular modeling. *Trends Biotechnol.* 21, 255–262
8. Doudna, J.A. and Charpentier, E. (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346, 1258096
9. Digman, M.A. and Gratton, E. (2012) Scanning image correlation spectroscopy. *Bioessays* 34, 377–385
10. Caprioli, R.M. (2016) Imaging mass spectrometry: molecular microscopy for the new age of biology and medicine. *Proteomics* 16, 1607–1612
11. Lubeck, E. and Cai, L. (2012) Single-cell systems biology by super-resolution imaging and combinatorial labeling. *Nat. Methods* 9, 743–748
12. Trapnell, C. (2015) Defining cell types and states with single-cell genomics. *Genome Res.* 25, 1491–1498
13. Coulon, A. *et al.* (2013) Eukaryotic transcriptional dynamics: from single molecules to cell populations. *Nat. Rev. Genet.* 14, 572–584
14. Tiscornia, G. *et al.* (2011) Diseases in a dish: modeling human genetic disorders using induced pluripotent cells. *Nat. Med.* 17, 1570–1576
15. Clevers, H. (2016) Modeling development and disease with organoids. *Cell* 165, 1586–1597