CellPress

Staying Clear of the Dragons

Johan Elf[1,*]
[1]Department of Cell and Molecular Biology, Science for Life Laboratory, Uppsala University, 75124 Uppsala, Sweden
*Correspondence: johan.elf@icm.uu.se
http://dx.doi.org/10.1016/j.cels.2016.04.008

A new, game-changing approach makes it possible to rigorously disprove models without making assumptions about the unknown parts of the biological system.

An early hope in modern systems biology was that theory, modeling and computation would solve the problems of biological complexity. However, the research program where experimentalists struggle to test, and at best falsify, the predictions of quantitative models in order to drive new theory development has not yet taken off. Considering the complexity of living matter, one may ask why theory and quantitative modeling do not have the same central position in biology as they do in the study of inanimate matter. Clearly formulated falsifiable models are in no way less needed in biology than in physics. Could it be that the theoretical tools that have been developed so far simply are too blunt to cut through the complexity of dynamical biological systems and thus fail to give real guidance to experimental research? In this issue of *Cell Systems*, new work from Andreas Hilfinger, Thomas Norman, and Johan Paulsson (Hilfinger et al., 2016a) provides a sharper tool.

One prolific modeling approach uses constraints from physical chemistry and independently measured parameters to build specific dynamic models of intracellular reactions systems that can be tested against data. For example, it is possible to calculate the expected cell-to-cell variability due to intrinsic stochastic fluctuations; the temporal responses in gene expression in response to a drug; or even include information about localization, transport, and structural rearrangements. When the model is specific and the predicted quantities are measurable, it is possible to make increasingly detailed and accurate measurements until the boundaries of the model's validity are carefully explored.

This approach has worked well in many cases, for example Hammar et al. (2014), but it has at least two problems: first, model predictions are rarely specific enough to be rigorously falsified. The range of predictions that are compatible with the model is simply too large and the biological measurements are too inaccurate. Only if the experiments are accurate and the predictions are unexpected can one learn something from not being able to falsify a model, but this is not generally the case.

Second, if there is only a slight disagreement between a model's predictions and experimental data, it is tempting to blame uncontrolled measurement errors or "biological factors and unknowns" that we cannot account for, rather than the model. This is where the cycle of "model-prediction-experiment- falsification- build new model" often stops. There are some interactions that we do not know how to account for so we cannot be certain if we have falsified the model, so let's keep believing in it. Ideally one should instead make model predictions that are independent of the unknown factors of biology. By analogy to old maps, this approach would let sea-faring navigators chart effective courses between ports while also staying clear of the dragons in the uncharted parts of the sea (Figure 1)

Hilfinger et al. (2016a and 2016b) provides biology a rigorous means to avoid its dragons; their approach avoids the need to make assumptions about the myriad uncharted, influential reactions that may go on in the cell. They present a framework for testing models that, though they may contain an arbitrarily small number interactions between components, nonetheless capture real biological relationships. The authors derive relations between measurable quantities that must be satisfied in order for the selected set of reactions to be correct (Hilfinger et al., 2016b). As a consequence, the relations can be used to exclude whole classes of models that have no chance to satisfy the data. It is obviously still impossible to prove specific parts of the model right, since there will always by many models that would give rise to indistinguishable results, but previously it has also been nearly impossible to prove models wrong. The strength of the work is that Hilfinger-Paulsson relations are exact and valid for highly unspecified systems. This is exceptionally rare in biology, although biology is probably where it is most needed.

The framework is based on the theory for stochastic chemical reactions (van Kampen, 2007), which infers relations between fluctuations in the numbers of different molecule species. The dynamics of the chemical processes drive the number of molecules from their average values, even when the system is at steady state. For this reason, steady-state distributions of copy-number fluctuations include information about the rates of chemical processes, and the correlations in the fluctuations between different species include information about the reactions that they share. The exceptional consequence of this is that the Hilfinger-Paulsson relations can be used on data measured in fixed cells. It is generally much easier to measure at the level of specific molecules in individual fixed cells than their dynamic behavior over time. In total, studying Hilfinger-Paulsson relations is both more experimentally accessible and logically rigorous than conventional approaches.

The possibility of rigorously discarding whole classes of models has two important consequences: first, it allows researchers to zoom in on under-appreciated classes of interactions that may still be valid and to make more specific experiments to test these interactions. Second, it will give theory a much more central position in systems biology, where it has often been hard to

Figure 1. Here Be Dragons
The "Carta Marina,"a map hand-drawn by Olaus Magnus in 1539, marks uncharted parts of the sea with dragons, allowing mariners to navigate around them. Hilfinger at al. address the problem of how to test quantitative models of parts of systems that are embedded in a sea of uncharted reactions. Image: Carta marina, a wallmap of Scandinavia, by Olaus Magnus from https://commons.wikimedia.org/wiki/File:Carta_Marina.jpeg.

rigorously invalidate interpretations of data without making assumptions that always could be questioned. Instead, the new theory guides experimental design by declaring what should be measured in order to test specific models of intracellular reactions.

Another practical advantage of Hilfinger et al.'s approach is that the quantities that end up in their relations are much more concrete and measurable than what we may have expected, especially considering all the rate constants that go into the underlying Master equation present in conventional approaches. The measurements needed by Hilfinger et al.'s approach are average abundances, reaction stoichiometries, average life times and correlation coefficients. In some rare cases (for example, Taniguchi et al., 2010) this information has already been acquired.

In summary, Hilfinger et al. is a highly recommended read, not only for the specific theorems, but for the way we approach biological science. It opens up a new way for extracting dynamic information from existing omics-scale, single-cell datasets and suggests which new experimental data we need to collect. Also, for being a paper from the Paulsson lab, it is unusually easy to understand. The paper is a game changer for quantitative modeling in systems biology. It makes it possible to falsify models rigorously, without hiding behind questionable assumptions or worrying about the hidden dragons of unspecified dynamics, and thus directing the focus on what may still be true. Beware, the theoreticians have finally climbed into the driver seat of systems biology.

REFERENCES

Hammar, P., Walldén, M., Fange, D., Persson, F., Baltekin, O., Ullman, G., Leroy, P., and Elf, J. (2014). Nat. Genet. *46*, 405–408.

Hilfinger, A., Norman, T.A., and Paulsson, J. (2016a). Cell Syst. *2*, this issue, 251–259.

Hilfinger, A., Norman, T.M., Vinnicombe, G., and Paulsson, J. (2016b). Phys. Rev. Lett. *116*, 058101.

Taniguchi, Y., Choi, P.J., Li, G.W., Chen, H., Babu, M., Hearn, J., Emili, A., and Xie, X.S. (2010). Science *329*, 533–538.

van Kampen, N.G. (2007). Stochastic processes in physics and chemistry, Third Edition (Amsterdam: Elsevier).