

NEWS AND VIEWS

A new model for circadian clock research?

Daniel Forger, Mark Drapeau, Ben Collins and Justin Blau

Department of Biology, New York University, New York, NY, USA

Molecular Systems Biology 28 June 2005; doi:10.1038/msb4100019

From a closer and purer league between...the experimental and the rational...much may be hoped. (Bacon, 1620)

Although Francis Bacon proposed the benefits of interdisciplinary science in 1620, only recently have molecular biologists and mathematicians talked with any frequency. Although no modern biologist would deny the validity of computational approaches in biology—just look at the burgeoning field of Genomics—how useful mathematical modeling will be to biologists remains debated (Lawrence, 2004; Tyson, 2004). An answer may be here with a study in this issue of MSB by Locke *et al* (2005), which highlights the advantages of being able to work effectively with models and molecules.

The study by Locke *et al* (2005) focuses on circadian rhythms, daily rhythms of behavior and physiology found in most organisms. These rhythms range from human sleep/wake cycles to leaf movements in plants. The cyclical nature and the precision of these internally driven rhythms has intrigued mathematicians and biologists alike. Yet despite working on the same questions for decades, most circadian molecular biologists have not embraced mathematical modeling. The following dialogue highlights similarities and differences in the views of a biologist (B) and a mathematician (M):

B: I don't understand how 'Math-Biology' will help my research—mathematical models seem more descriptive than predictive.

M: Well, I have an excellent paper for you to read in which the authors move freely between computer simulations and experiments. Locke *et al* (2005) used experimental data to build a model, and then tested whether this model could predict other experimental data not initially included. Their initial model had only three genes in the network and did not match the *in vivo* data—so Locke *et al* added hypothetical components to make the model more accurate. Their simulations worked so well that they were able to return to experiments and identify a strong candidate for one of the hypothetical components.

B: That does sound useful. What were they modeling?

M: The *Arabidopsis* circadian clock.

B: Oh yes, I read about that. I can guess which genes they started with: *TIMING OF CAB EXPRESSION 1 (TOC1)*, *LATE ELONGATED HYPOCOTYL (LHY)* and *CIRCADIAN CLOCK ASSOCIATED 1 (CCA1)*. Experiments have shown that TOC1 activates *LHY* and *CCA1* expression, and that *LHY* and *CCA1* proteins then feed back to inhibit *TOC1* expression and, consequently, to inhibit further *LHY* and *CCA1* expression

(Alabadi *et al*, 2001). A classical clock negative feedback loop. In fact, the clocks in *Drosophila* and mammals have two of these transcription/translation feedback loops interlocked with one another (Hardin, 2004).

M: Yes, that's right. In their paper, Locke *et al* found that the single TOC1/LHY/CCA1 loop could not explain data that they measured *in vivo*, such as weak residual 18 h rhythms in plants with both *lhy* and *cca1* mutated. So they added a second loop to the *Arabidopsis* clock, and their simulations were much more accurate.

B: But there are two loops in the *Drosophila* and mammalian circadian clocks (Hardin, 2004), so is it really a surprise to find the same in plants? That does not seem very predictive.

M: Wait: The authors' simulations predicted that RNA levels of 'Factor Y', the key player in the second loop of their model, would show two peaks of expression every day—a burst of expression at dawn, and a broader peak at dusk. Then they went back to the bench and looked at the expression profiles of a number of genes known to affect circadian gene expression but which had not yet been fitted into the molecular clock network. Since they were looking for a very brief peak of RNA at dawn, they designed their experiments to sample every hour around dawn and then less frequently over the rest of the day.

B: And?

M: They found one gene, *GIGANTEA (GI)*, whose expression paralleled the rhythms of Factor Y.

B: So is GI Factor Y?

M: Probably, because *gi* mutants have low amplitude molecular clock oscillations (Mizoguchi *et al*, 2002). It is a very strong candidate, but we will need experiments to test this.

B: Great! But how did Locke *et al* design an accurate model without knowing the abundance or half-lives of any of these proteins?

M: This is called an inverse problem in Mathematics and, rather than starting with known parameters, they have to be chosen to match experimental data. Then one runs simulations to see if the model fits experimental data. This type of parameter sampling is widespread in other areas of mathematical modeling and was used to model the mammalian circadian clock (Forger and Peskin, 2003). When I said that Locke *et al*'s one-loop model did not match the experimental data, I meant that they could not find a set of parameters that would simulate the experimental data.

B: I see. So does all of this mean that I should run a simulation before my next experiment? Not a chance!

M: Funny that you say that. This is one of the implications of Locke *et al*'s study—that simulations can help design great experiments. Remember, circadian expression profiles for much of the *Arabidopsis* genome have been available for nearly 5 years (Harmer *et al*, 2000), but the dawn peak of *GI* expression was missed because the sampling times were every four hours. I don't know if anyone would have caught this early *GI* peak unless they sampled at one hour intervals. So simulations were invaluable in this case. You know, there are models and interactive, user-friendly tools for biologists to run simulations on the Web—for example, www.amillar.org/Downloads.html, www.sbml.org or www.BioSpice.org.

B: But how would I know which model to use? It is a long time since I studied Mathematics.

M: You need to look for rigor in the model: Biological rigor—the modeler should precisely state all biological assumptions; Mathematical rigor—the modeler should describe exactly how these assumptions were converted into equations; and Numerical rigor—the modeler should justify how these equations were solved. And the model that most accurately reflects the biology may be complex. If the underlying biology is complex (many proteins, many cells, etc.), then do not expect a simple model.

B: But do you really think that this can help Biology in general? We already know about so many genes and so many pathways.

M: That is my main point. As biologists find increasing numbers of components in pathways, computer simulations will be needed to identify their relationships. With mathematical modeling, diagrams of interactions between genes and proteins take on analytical power, and can reveal insights missed by verbal reasoning. Use the power of computers for all kinds of biological research, not just for circadian biology—or at least ask people like me for help! And although we know a lot of genes in some networks, we do not understand how they work as a system. For example, how do circadian clocks keep 24 h rhythms across a range of temperatures when individual biochemical reactions are temperature-dependent?

B: Okay, last question. The *Arabidopsis* clock loop had one loop yesterday, and two today. Modelers constructed simulations of the *Drosophila* and mammalian clocks that were rhythmic with just one loop, and then added a second loop when new components were identified (Leloup *et al*, 1999; Leloup and Goldbeter, 2003). Do you think you could predict how many feedback loops there are in a circadian clock? Two, three, four?

M: Good question. Let me get back to you on that one...

References

- Alabadi D, Oyama T, Yanovsky MJ, Harmon FG, Mas P, Kay SA (2001) Reciprocal regulation between TOC1 and LHY/CCA1 within the *Arabidopsis* circadian clock. *Science* **293**: 880–883
- Bacon F (1620) THE NEW ORGANON
- Forger DB, Peskin CS (2003) A detailed predictive model of the mammalian circadian clock. *Proc Natl Acad Sci USA* **100**: 14806–14811
- Hardin PE (2004) Transcription regulation within the circadian clock: the E-box and beyond. *J Biol Rhythms* **19**: 348–360
- Harmer SL, Hogenesch JB, Straume M, Chang HS, Han B, Zhu T, Wang X, Kreps JA, Kay SA (2000) Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. *Science* **290**: 2110–2113
- Lawrence P (2004) Theoretical embryology: a route to extinction? *Curr Biol* **14**: R7–R8
- Leloup JC, Goldbeter A (2003) Toward a detailed computational model for the mammalian circadian clock. *Proc Natl Acad Sci USA* **100**: 7051–7056
- Leloup JC, Gonze D, Goldbeter A (1999) Limit cycle models for circadian rhythms based on transcriptional regulation in *Drosophila* and *Neurospora*. *J Biol Rhythms* **14**: 433–448
- Locke JCW, Southern MM, Kozma-Bognar L, Hibberd V, Brown PE, Turner MS, Millar AJ (2005) Extension of a genetic network model by iterative experimentation and mathematical analysis. *Mol Systems Biol* 28 June 2005; doi:10.1038/msb4100018
- Mizoguchi T, Wheatley K, Hanzawa Y, Wright L, Mizoguchi M, Song HR, Carre IA, Coupland G (2002) LHY and CCA1 are partially redundant genes required to maintain circadian rhythms in *Arabidopsis*. *Dev Cell* **2**: 629–641
- Tyson JJ (2004) A precarious balance. *Curr Biol* **14**: R262–R263