How can we fully realize the potential of mathematical and biological model systems to re-integrate biology?

Authors: Anna Dornhaus[1], Brian Smith[2], Kalina Hristova[3], Lauren B. Buckley[4]

[1]Department of Ecology & Evolutionary Biology, University of Arizona, Tucson, AZ 85721
[2] School of Life Sciences, Arizona State University, Tempe, AZ 85287
[3] Departments of Materials Science and Biomedical Engineering, and Program in Molecular Biology, John Hopkins University, Baltimore, MD 21218
[4] Department of Biology, University of Washington, Seattle, WA 98115

Summary:
Both mathematical models and biological model systems stand as tractable representatives of complex biological systems or behaviors. They facilitate research and provide insights, and they can describe general rules; in fact, models that represent biological processes or formalize general hypotheses are essential to any broad understanding. But at the same time, models - either mathematical or biological - do not represent all the detail of the natural system (although the degree of abstraction varies) and thus may ultimately be incorrect representations. A key challenge is that this is not just 'noise': models can be incorrect in their qualitative, broad implications if those details matter (e.g. Fussman & Blasius 2005). Our paper discusses this tension, and how we can improve our inferences from both mathematical models and biological model systems (e.g. organisms like *C. elegans*, *Drosophila*, or *Arabadopsis* in controlled environments; or even purified proteins in buffers, lipid bilayer membrane models, etc.; to some degree, most lab experiments can be seen as abstracted versions of the natural world). We advocate for further efforts dedicated to model development, improvement, and acceptance by the scientific community, all of which may necessitate a more explicit discussion of the purpose and power of models.

We argue that models should play a central role in re-integrating biology as a way to test our integrated understanding of how molecules, cells, organs, organisms, populations, and ecosystems function. We explore the following issues required to realize the integrative potential of models:

- We argue that biologists have yet to appreciate the full range of ways models allow us to make inferences and stimulate and direct other research. Fully leveraging models requires recognizing that models are not and cannot be proven to be perfectly faithful representations of the systems that they are modeling; their power lies in demonstrating or testing a principle, not the specifics.
- Researchers often struggle with a tension between tractability and 'realism' of both mathematical/computational models and biological model systems, in that they may be more or less removed from the system they represent. We argue that biologists using biological model systems face some of the same challenges as mathematical modellers in this, and should also consider what, fundamentally, the purpose of their 'model system' is, and what insights can and cannot be gained from it.

- We argue that we need improved training for both junior and senior scientists and improved funding mechanisms for theory and system development to enable researchers to capitalize on the powers, but also deeply appreciate the limitations of models and model systems, including going beyond just what is historically an 'accepted' model system towards what is a useful and appropriate model.

Main text

Our vision

We believe that to reintegrate biology, as well as to accelerate progress and deep understanding in science, it is critical that both empiricists and theoreticians fully appreciate the inferences that can and cannot be drawn from models and model systems. Any 'rules of life', or principles that transcend subdisciplinary boundaries as well as levels of organization in biology, are bound to be phrased as abstractions; and detailed mechanistic understanding of almost any system requires the study of well-controlled model systems that can represent, without fully capturing, natural variability and complexity.

If empiricists were trained and enabled to examine what purpose models serve, it would enable them to be more rigorous and effective in their own work, and give them access to a more universal language, enabling them to better transcend disciplinary boundaries and identify general principles. If theoreticians were better trained to understand and appreciate the goals of empirical science and the purpose that theory can have in improving our understanding of the world in a scientific context, communication between theory and empirical research would be strengthened, positively transforming the efficiency and generality of biological research.

These considerations also apply to field biologists vs. lab scientists studying biological model systems. That is, a field biologist would benefit from appreciating the power, tractability, and inferences that can be made from studying, for example, cells in vitro; a lab biologist studying cells in vitro would benefit from a solid understanding of how much the petri dish differs from the function and activity of cells in a living, behaving animal in the field. As with the distinction between empiricists and theoreticians, while mutual respect is often present at the surface, a deep communication is often hindered by a lack of explicit discussion of goals of research and what inferences can be made from the 'model system' towards and understanding of processes in the natural world.

What is a model

The systems that biologists ultimately want to understand are organisms, populations or ecosystems in their natural environment. Generally, what we consider as a model is a simplified representation of such a system, typically with the purpose of understanding the complex system better by studying the simplified version. Models are ubiquitous in biology, both in the sense of theory, i.e. mathematical or computational or otherwise abstractions from the real world, and in the sense of model systems, by which we mean specific organisms or preparations (e.g. cultured cells), which are considered to be representative of organisms more generally.

The purpose and benefits of models

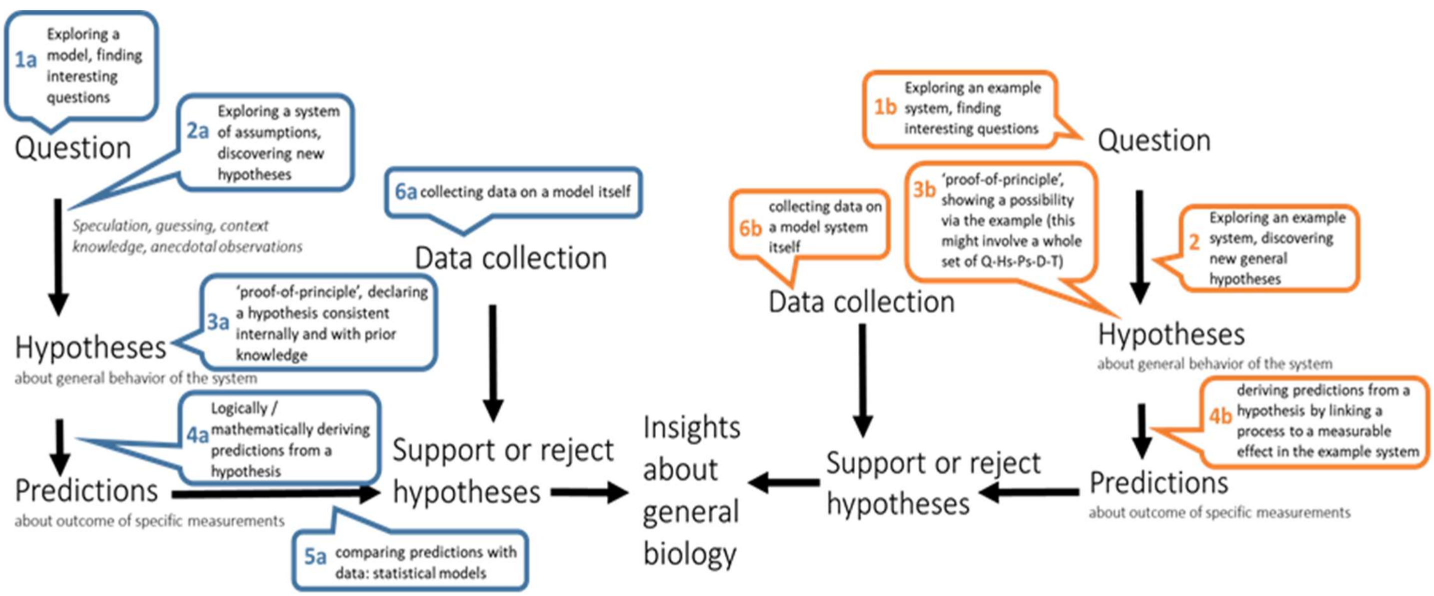
Models are a powerful tool for integrating biology by encapsulating general principles and linking levels of biological organization (molecules through organisms through behavior through populations). Models can also integrate biology by uniting empirical understanding from different experiments into a unified framework. In principle, models offer the potential of a universally understandable language that can transcend subdisciplines; however, this potential is often not realized because of a lack of training in philosophy of science (see training section below).

Key to realizing the potential of models as an integrative tool is engagement with empiricism. This means, modelers of all kinds (i.e. including people working on model systems) need to also have knowledge of the natural system that the model represents. The goal of models should be clearly defined, not only to inform their structure but also to clearly communicate across biological research communities what insights can be gained from them. Classic models (e.g. in physics) represent key hypotheses (often in mathematical form) and then generate testable predictions, i.e. usually specific, quantitative, empirically measurable outcomes (Fig. 1, #4); empirical data thus can be used to reject or support the key hypothesis that the model represents (e.g. in biology, mutations occur prior to need not in response to it, in the classic study of Luria & Delbruck 1943). Models of this type are most informative when they are inconsistent with empirical data, because it allows us to reject the underlying hypotheses. Conversely, consistent empirical data can not rule out alternative causal mechanisms (e.g. McGill et al. 2007 show how many different mechanistic hypotheses make the same prediction about species abundance curves as neutral theory; for a fun illustration of this principle see also https://twitter.com/dornhaus/status/1204902045464981505).

However, many models in biology do not serve this purpose, but instead may serve to explore a system to generate questions or hypotheses (e.g. see Servedio et al. 2014). [Footnote: The specific classification presented here is based on what was developed and taught for several years in a graduate course at the University of Arizona by Joanna Masel and Anna Dornhaus] Commonly in biology, models are used to show that some assumptions necessarily lead to certain outcomes - we term this ‘proofs of principle’ (in Fig. 1, #3). These are often used to defend the plausibility of an unpopular hypothesis (e.g. ritualized contests can evolve by individual selection despite this seeming counterintuitive, Maynard Smith & Price 1973; or simple rules can in fact generate complex food webs, Williams & Martinez 2000), or to reject the possibility of a popular (verbal) hypothesis (e.g. dominant genes do not increase in frequency, Hardy 1908). Important here is to realize that the models do not provide evidence that these processes actually occur (e.g. that ritualized conflict actually evolved by this mechanism), only that they could (or could not) occur in principle. Biological model systems also very often serve this purpose: such as when a drug at high dose has an effect on cells in culture, this chiefly demonstrates its *possible* effectiveness to cure a disease, i.e. the principle, rather than being a proof of actual effectiveness in the system to be represented (e.g. a human body). Or when a study demonstrates that a particular African Grey Parrot can understand the concept of zero this serves to demonstrate that birds can understand this concept in principle (Pepperberg & Gordon

2005); whether they commonly (or ever) do so in the wild can not be directly inferred from this ‘model’, and requires the study of the original system (here, parrots in the wild).

Fig. 1: The purpose of models. It is important to define what function a particular theory model has in the context of scientific progress. Using the ‘scientific method’ (strong inference) as the basic process, models can potentially add insights at these 6 positions. (Figure credit is Anna Dornhaus & Joanna Masel, UA, based on a lecture slide). “a”/blue boxes reflect mathematical/computational models, “b”/orange boxes reflect biological model systems. Models can inspire/direct research on the ‘real’ system of interest by helping to generate new questions or hypotheses (1a/b, 21/b); or by showing for either an example (3b) or for an abstract set of assumptions (3a) the possibility of a particular phenomenon (or impossibility in the case of 3a). They can also help to specify what to expect from a particular hypothesis in terms of measurable quantities (predictions) (4a/b). Statistical models are used to help assess how much predictions agree with data (5a). Finally, in some cases, data are collected on the model itself, in which case they are informative primarily about the model, not about the ‘real’ system the model represents.



It is critical to point out that many aspects of biological systems cannot be measured at all or precisely enough unless a model system is used; and the same justification is also used for theory models. Indeed a huge amount of progress has been made in biology based on the insights gained from effort in researching very controlled and reduced systems (e.g. purified proteins or nucleic acids in buffers, lipid bilayer membrane models, or cell cultures, but also specific model species or preparations like the proboscis extension reflex in restrained honey bees as a model for principles of learning). Such studies allow us to detect patterns and effects that would be impossible to discover in the noisy ‘real’ world. This variable reduction is thus the main strength and the main weakness of (any kind of) models.

Science as discovering new mechanisms vs. application for specific benefit
The argument above principally relates to explanatory models, that is, models that serve to better understand a not-yet understood system. It is worth pointing out that in some areas of biology, models also have an applied function which we term 'forecasting'. A forecast is a prediction in the sense of a best-guess about the future given all knowledge we currently have. For example, in some climate and population models, we use everything we already know about factors driving the system to estimate how a certain future will play out, e.g. how much the planet will warm given a particular increase in CO2 (e.g. Stainforth et al. 2005). Similarly, a model derived from detailed knowledge of processes at a cell or organ level might be used in the context of medicine to inform a patient of their prognosis. We call this 'forecasting' to distinguish from the 'predictions' relevant for testing hypotheses (sensu Fig. 1); forecasting is concerned with best-estimates based on the knowledge we have, not on gaining new knowledge about what drives the system. That is, we are not trying to discover new mechanisms, for example driving climate, we instead have an applied, practical interest in faithfully estimating how the system will behave in the future, or under certain scenarios. This implies that for this kind of model, all current estimates for factors thought to play a role are best included, often leading to fairly complex models (e.g. see latest climate models); simultaneously, this makes such models typically very hard to intuitively understand, and often precludes identifying general principles. In some cases, forecasting is done with detailed mechanistic models (based on existing knowledge of mechanisms); in some cases, statistical models of past behavior/data are as accurate or more so in order to correctly forecast system behavior (Buckley et al. 2010, Kearney & Porter 2009). In either case, the model is not usually suitable for any of the functions listed in Fig. 1; its purpose lies in an applied benefit, not in generating new knowledge.

Tractability and usefulness of models
A central challenge for models is the tradeoff of generality versus realism: models need sufficient realism to produce the important mechanisms and thus to be considered relevant, but increased realism tends to reduce generality and tractability; this tradeoff also applies when aiming to scale models across levels of biological organization, space, and time. Reconciling these demands may be easier said than done, both for theory models and for biological model systems. For example, including realistic environmental fluctuations is likely central to exploring the plasticity of phenotypes in nature, but environmental fluctuations are highly context specific and difficult to describe and empirically replicate (Dillon and Woods 2016).
A perceived lack of realism (of mathematical models or of biological model systems) is the main criticism often expressed by researchers closer to the natural system. Expectations for 'realism' in models can be too high. It is important to realize that the degree to which 'realism' is a goal is intimately linked to the purpose a model is to serve. For example, for models intended for 'forecasting' (see above), the usefulness of the model depends on its faithfulness in representing the outcomes of all relevant biological processes (although even here the relevant mechanisms need not be explicitly represented, as in a machine learning model that makes inferences from example data without knowledge of mechanism). But models used in science (for discovery of new mechanisms or to test hypotheses) often do not need to make quantitatively accurate predictions. For models demonstrating a principle, for example, often a

highly abstracted, minimal model gives the most general and powerful insight. In addition, in some cases, the purpose of a model is exploration (#1 or #2 in Fig. 1), in which case the model may represent possible worlds or imperfectly represent a natural system and still be successful in prompting new questions or hypotheses. It is thus an important lesson for empiricists/field biologists that 'realism' is frequently neither the goal nor desirable.

In many cases, suites of models of differing complexity and generality are likely to further the integrative capacity of models. Specific approaches can be deployed in a given scenario as a function of knowledge of the system, the number of interacting parts, the contribution of knowns and unknowns, and precision and accuracy in empirical measurements. Mechanistic approaches may facilitate scaling, particularly across levels of biological organization, but they can be resource- and knowledge-prohibitive to broadly implement. Inherent stochasticity in biological processes can be difficult to distinguish from experimental artifacts or unincorporated mechanism.

Cases where modelling has achieved highly detailed yet faithful models

We highlight examples across fields of how models differing in detail and generality have been employed as an integrative tool:

- In biophysics/biochemistry, most of the knowledge has been gained through investigations in highly simplified model systems, such as purified molecules in solution, where parameters can be directly assessed and quantified. This work has also greatly benefited from computational models that capture the behavior of the molecules (conformational flexibility, interactions) from first principles (Alford et al. 2017, Childers and Dagget. 2017)
- In physiology, there have been significant successes in the development of multiscale physiological models, describing interconnected systems of events and incorporating known behaviors of building blocks such as molecules, cells, etc. These models allow us to generate highly faithful models of processes such as angiogenesis, for example, based on what is known about signaling molecules and interaction networks (Qutub et al, 2009)
- In ecology and evolution, mechanistic distribution models are being developed to test our understanding of ecological and evolutionary responses to environmental change as well as forecast biodiversity shifts (Urban et al. 2016). Statistical models provide an expedient and general tool to forecast shifts in distribution and demography, but they exhibit mixed performance in extrapolation to the novel combinations of abiotic and biotic conditions that result from environmental change (Maguire et al. 2015). Mechanistic models that integrate our empirical understanding of organism-environment interactions to predict performance, fitness, and demographic shifts exhibit forecasting promise, but struggle with generality and tractability (Kearney and Porter 2009, Buckley et al. 2010). Emerging and called for collaborative efforts to develop, parameterize, and test models should further integrative capacity (Urban et al. 2016, Dietze et al. 2018).

Cultural and concrete changes

Coordinated efforts between ‘empiricists’ (researchers close to the ‘natural’ system) and ‘modelers’ (researchers working on an abstracted version, including biological model systems) for model development and interpretation are essential to overcoming these challenges and realizing the integrative promise of models. How can we realize such coordination?

The argument that empiricists and theoreticians/modelers should engage more is old. What really will transform science in a way to encourage more and better such interaction? We argue that both empiricists and modellers should be more explicit about the purpose of their model in gaining insights, and specific about how the different elements of the model contribute to that insight-gain. Matching this, readers of modeling studies should accept that models have value only in relation to the stated purpose, and should not be judged by their ability to fulfill purposes for which they were not intended (which they typically cannot).
This need for explicitly stating the purpose and possible conclusions for the natural world beyond the model itself should include studies using model systems, which need to articulate the applicability of their model systems to natural settings, and it includes studies with mathematical models, which need to clearly state the purpose of their model, how it does or does not correspond to natural settings, and the model’s potential to interact both with empirical data and other models. Integration will likely best be achieved by models with a spectrum of detail and generality, but it will also require a willingness by both empiricists and modellers to sometimes compromise on their desire for detail and generality to further integration. A focus on the type of insight gained from any particular model will achieve some of these goals automatically, by focusing all parties on whether the assumptions made in the modeling are appropriate for the claim.

In biological model systems in particular, many researchers are skeptical of new model systems, which are scrutinized for their ‘realism’ or lack of it, while at the same time unquestioningly accepting well established model systems despite their many known limitations, and despite these established systems being just as much abstracted from the natural system. Some of this is motivated by a lack of close examination of the purpose of each model and what we aim to discover. As with mathematical models, it is key to examine whether a general principle can be demonstrated, an interesting hypothesis tested, or a system explored to motivate new questions. In many such cases, ‘realism’ is not necessary as long as key assumptions are made explicit. As a community, we benefit from both improvements and diversity in model systems, and we should celebrate each improvement while being mindful of the limitations, and the specific purpose, of the model, as well as of the natural system that the model is thought to represent. Such an approach will drive progress and lead to deeper understanding of biological phenomena.

We believe there are also specific activities and training approaches that will yield concrete benefits in applying and interpreting model studies. Particularly for highly parameterized models used in forecasting, the push toward open and reproducible science, including documenting and sharing methods, code, tools, and data is central to coordinated efforts (Mislan et al. 2016). For example, such coordinated initiatives are emerging to forecast ecological responses to environmental perturbations (Ecological Forecasting Initiative: https://ecoforecast.org/, Dietze et al. 2018; Zoon R package for species distribution modelling, Golding et al. 2017). Concerted data collection and databasing efforts are often essential to parameterizing models and consistently testing their performance (e.g. as called for in biodiversity forecasting Urban et al. 2016). Such efforts will require structure and incentives for collaborations among researchers across subdisciplines including those who work in laboratories, in the field, and computationally (Buckley et al. 2018). In addition, initiatives are needed to ensure that quantitative and computational capacity and training are not limiting (see concrete suggestions below).

Operationalizing our vision

Implementation of new approaches to integration across levels of analysis in biology, including between theoretical and experimental approaches, will require reassessment of many current institutional practices. New practices must consider funding priorities as well as training opportunities at every career stage. More funding opportunities must be created that explicitly require modelers, whether theoreticians or those working on model systems highly abstracted from the natural world, to communicate with those working directly on the complexity of the real world. Often this can be achieved if projects combine theory with experiments, including integrative outcomes against which successful combination can be evaluated. This might include workshops, both small and large, where researchers who specialize in theoretical and experimental approaches are required to meet and exchange ideas. It also includes 'internationalization' of funding programs, such that researchers can more effectively collaborate across the world. This would bring funding mechanisms up to the speed and efficiency that is now possible in communication and collaboration by researchers across long distances in the 21st century. And it would allow for the formation of the best possible teams for accomplishing any desired outcome. Training opportunities include sabbatical-like possibilities for theoreticians and experimentalists to work in each others' laboratories. Moreover, changes in the near future can be made more likely by opportunities that are created for early career researchers, who need to be productive in areas that are normally viewed as a risk for them because of the need for team-based science. These changes therefore must include rethinking hiring and promotion practices at research and teaching institutions. And it should include development of undergraduate and graduate fellowship opportunities in integrating biology across levels as well as across theory and experiment, which would automatically instill in the next generation of researchers the importance of interdisciplinary, team-based science.

Concrete suggestions to funding agencies

- It is hard to get theory funded, particularly relatively abstract mathematical or computational modeling. This is a known problem. We offer the suggestion that theoreticians should be enabled to have much larger parts of their grants devoted to 'listening sessions' to learn about what questions empirical biologists have, and to 'outreach sessions' where theoreticians offer basic training in how to read their papers. We imagine this could take the form of larger national meetings or, perhaps preferably, small local (institutional) workshops.
- We also suggest that collaborations between modelers and empiricists or researchers closer to the real system should be encouraged. In the case of biological model system research, this has not traditionally been done, but we believe it is just as vital as it is in mathematical modeling. In both cases, such collaborations should be allowed high flexibility, as in many

cases broad interactions may be more fruitful than specific collaboration on a particular project; however, the latter should also be enabled. Modelers sometimes need to finish their work before the corresponding empirical study can be done, or vice versa, a model may become necessary only after the empirical work has uncovered an unexpected phenomenon. In these cases, add-ons funding or collaborative grants which do not require to be simultaneous in time would be beneficial.

- Training at the senior (faculty) level: not all training can be incorporated in graduate programs, and often scientists will only realize as their research program advances how much they could benefit from, or how much they rely on, models of all types. Broadly accessible funding for sabbaticals, to enable researchers to train themselves in a subdiscipline of their choice, might be an effective option for enabling such training. Alternatively such funding could be specific to spending time in a theory/empirical lab respectively for empiricists/theoreticians. Ultimately we believe that training must go beyond a single workshop to be effective.
- Graduate fellowship opportunities: currently, GRFPs are restrictive (only US citizens, no one with an MSc) and scarce, and thus not a major funding source for graduate students. But if students are paid either through RAships or TAships, they are more risk averse, and will likely tend to focus on generating publishable data in a narrow field. More generalized funding for graduate students, including costs of travel to other labs, would increase independence from advisor's (on principle narrowly defined) grants, and enable students to genuinely train interdisciplinarily.

Concrete suggestions to researchers in biology

- Pay more attention to curriculum at undergraduate and graduate level with regard to philosophy of science, i.e. how we know what we know, and why we need models (and what for). Most programs do not include this. Skills in actual math or data science may also be relevant, but are not sufficient: here we are talking about the skill to actually understand what is contributed by a theory paper, and what insights from an empirical study on a model system can be generalized. The often-misunderstood 'Scientific Method' framework can provide a guide here (see Fig. 1).
- Have high expectations on seminar speakers, theses, and papers reviewed to explain clearly, specifically, and in detail what inferences can be drawn from their model/model system. (But be patient when everyone speaks different languages.)
- Accept that you may have an incorrect understanding of the purpose of models generally or a particular model. Examine your own assumptions.
- Put 'teams' of PhD students and postdocs together to be innovative and interdisciplinary, and give some kind of funding mechanisms to given them the wherewithal to implement to research

Conclusion

We believe that a more explicit recognition of the purpose and power, but also the limitations, of models can serve to reintegrate biology and generate fundamental progress. There are some apparently low-hanging fruit to achieving this, but scientists at all levels of seniority might benefit from additional chances for training, and from a recognition that the purpose of models is often misunderstood, not least because it is not clearly communicated.

References

Buckley, L. B., Urban, M. C., Angilletta, M. J., Crozier, L. G., Rissler, L. J., & Sears, M. W. (2010). Can mechanism inform species' distribution models?. *Ecology letters*, *13*(8), 1041-1054.

Buckley, L. B., Cannistra, A. F., & John, A. (2018). Leveraging organismal biology to forecast the effects of climate change. *Integrative and comparative biology*, *58*(1), 38-51.

Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S., ... & Loescher, H. W. (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proceedings of the National Academy of Sciences*, *115*(7), 1424-1432.

Dillon, M. E., & Woods, H. A. (2016). Introduction to the symposium: beyond the mean: biological impacts of changing patterns of temperature variation. *Integrative and comparative biology*, *56*(1), 11-13.

Golding, N., August, T. A., Lucas, T. C., Gavaghan, D. J., van Loon, E. E., & McInerny, G. (2018). The zoon R package for reproducible and shareable species distribution modelling. *Methods in Ecology and Evolution*, *9*(2), 260-268.

Hardy, G. H. (1908). Mendelian proportions in a mixed population. Science, section:“Discussion and Correspondence”. Vol. 28.

Kearney, M., & Porter, W. (2009). Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology letters*, *12*(4), 334-350.

Luria, S. E., & Delbrück, M. (1943). Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*, *28*(6), 491.

Maguire, K. C., Nieto-Lugilde, D., Fitzpatrick, M. C., Williams, J. W., & Blois, J. L. (2015). Modeling species and community responses to past, present, and future episodes of climatic and ecological change. *Annual Review of Ecology, Evolution, and Systematics*, *46*, 343-368.

McGill, B. J., Etienne, R. S., Gray, J. S., Alonso, D., Anderson, M. J., Benecha, H. K., ... & Hurlbert, A. H. (2007). Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology letters*, *10*(10), 995-1015.

Mislan, K. A. S., Heer, J. M., & White, E. P. (2016). Elevating the status of code in ecology. *Trends in ecology & evolution*, *31*(1), 4-7.

Pepperberg, I. M., & Gordon, J. D. (2005). Number comprehension by a grey parrot (Psittacus erithacus), including a zero-like concept. *Journal of Comparative Psychology*, *119*(2), 197.

Servedio, M. R., Brandvain, Y., Dhole, S., Fitzpatrick, C. L., Goldberg, E. E., Stern, C. A., ... & Yeh, D. J. (2014). Not just a theory—the utility of mathematical models in evolutionary biology. *PLoS biology*, *12*(12), e1002017.

Smith, J. M., & Price, G. R. (1973). The logic of animal conflict. *Nature*, *246*(5427), 15.

Urban, M. C., Bocedi, G., Hendry, A. P., Mihoub, J. B., Pe'er, G., Singer, A., ... & Gonzalez, A. (2016). Improving the forecast for biodiversity under climate change. *Science*, *353*(6304), aad8466.

Williams, R. J., & Martinez, N. D. (2000). Simple rules yield complex food webs. *Nature*, *404*(6774), 180.

Alford RF, Leaver-Fay A, Jeliazkov JR, O'Meara MJ, DiMaio FP, Park H, Shapovalov MV, Renfrew PD, Mulligan VK, Kappel K, Labonte JW, Pacella MS, Bonneau R, Bradley P, Dunbrack RL Jr, Das R, Baker D, Kuhlman B, Kortemme T, Gray JJ. The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. (2017) *J Chem Theory Comput*. 13(6):3031-3048.

Childers MC, Daggett V. Insights from molecular dynamics simulations for computational protein design. (2017) *Mol Syst Des Eng*. 2(1):9-33.

Qutub AA, Mac Gabhann F, Karagiannis ED, Vempati P, Popel AS. Multiscale models of angiogenesis. (2009) *IEEE Eng Med Biol Mag*. 28(2):14-31.