The problem of scale: Finding common ground from molecules to ecosystems

Alison Bennett
Bob Mason (masonr@science.oregonstate.edu)
Juan B. Gutierrez
Stephen Deslauriers (sdeslaur@morris.umn.edu)
Cynthia Downs (cjdowns@esf.edu)
Gail Rosen (gail.l.rosen@gmail.com)
Veronica Hinman (veronica@cmu.edu)
Anna Tansik (atansik@odu.edu)

Summary: Our aim in this vision statement is to develop central concepts, currencies, and tools that will help integrate concepts across hierarchical levels of biological organization. Practically our goal is to provide a language and tools for scientists who work at different levels of organisation so that they can communicate with each other and better identify. This should allow us to understand how, when, or if mechanisms at one level predict or are related to organization at another (higher or lower) level of organization. We discuss energy, matter, and information as currencies across levels of biological and discuss interactions and emergent properties as shared processes. We end by outlining barriers, scientific and otherwise, that have hindered the proposed cross-level work in the past and suggest some concrete solutions to those barriers.

Main aim: Can we take what we know and find commonalities and can we find commonalities by talking with each other across levels of biological organization/hierarchies? To do so, we need to think generally about systems instead of thinking about all of the details that make our individual systems special

Note: Denis Noble (wrote Music of Life) talks about cell/organism out thinking as a biological processes, as opposed to reductionist or emergentist thinking. Concepts are related to what we have been discussing.

Goals of vision statement:

1) Scientists working at different scales have a way of understanding the processes at that level and crosstalk among research that work at different scales will enhance ability to understand processes within a level.
2) Being able to understand the processes that are common across scales that we can use to understand the mechanisms that drive those processes and the emergent processes that arise at levels above the focal process and the consequences of the process.
3) Synthesize across the scales by identifying common biological processes, frameworks, themes, techniques, and ways of thinking so that we are not siloed.

4) Understand how, which, when, or if mechanisms at one level predict or are related to organization at another (higher or lower) level of organization.

Our working definition of scale

Biological scale can mean many things: Spatial scale (space), temporal scale (time), how traits change with (body) size, and levels of biological hierarchy (Fig 1). We are focused on the latter.

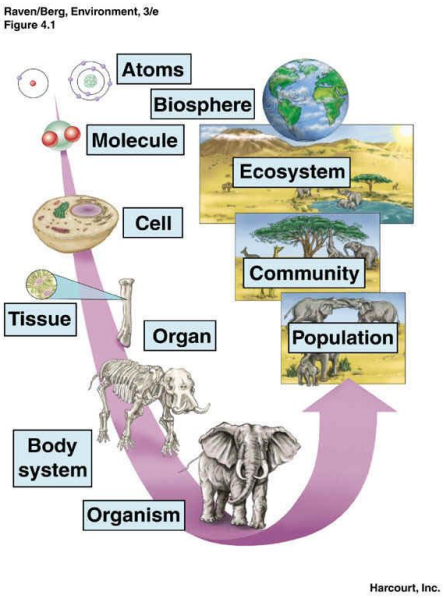

Figure 1. Level of biological hierarchy

Questions? **What might be the broader impacts/what is the potential impact?**

- When are biological hierarchies reflective of nature and when are they just constructs to help us think about nature? When does the metaphor break down?
- Help us to make predictions about system-level phenomenon without knowing all of the players in the game.
- Help us to make predictions about the response of the system to a perturbation and understand how a perturbation at one level will alter the systems at other levels of biology.

Examples of multiscale questions that need common ground:

There are a lot of “emergent properties” that no one understands because of complexity and dynamic nature of systems. Emergent properties are properties that can only be predicted from a systems point of view. The value of the property is different from the sum of the parts.

One example that illustrates the challenge of multi-scale questions is the interaction between plants and pathogens. Crop species are susceptible to a wide number of bacterial, viral, and fungal pathogens which can have significant economic impact (Fao 2017). For most of

these pathogens, effects can be observed across many scales. For example, Phymatotrichum root rot is caused by spread of the fungus *Phymatotrichum omnivora* through soil and it affects a broad range of crop species including cotton and alfalfa (Uppalapati et al. 2010). At one scale, the interaction between the fungi and root cortical tissue being colonized is of value for disease prevention. Which cellular or genetic changes take place during the course of infection? Other researchers might investigate how innate defenses of the individual plant affect susceptibility or how the yield of infected plants is affected. Rather than looking at the interaction between *P. omnivora* and an individual plant, other researchers may choose to focus on the population as a whole. How is the overall crop yield affected by spread of the fungi? Are there ways to limit spread through the soil by treatment with chemical fungicides? Are there long term effects of chemical treatment of the soil? What about from an ecological scale? Root rot is certainly prevalent in areas of natural plant growth. How does spread of the disease affect biodiversity of plants in that area? What about pollinators or herbivores? All of these questions are legitimate areas of research interest, but beyond the scope of a single research group. A variety of work has been done investigating the utility of CRISPR/Cas9 as a tool for introducing resistance genes (reviewed by Borrelli et al. 2018), but it would very helpful to know the ecological impacts of introducing these resistant plants into the field. Common ground needs to be found between these different levels of focus in order to find solutions.

Another example of a multi-scale problem would be directed evolution in response to climate change. Through genomics-assisted breeding, it is possible to generate populations that are resistant to changes in temperature, water availability, or CO_2 levels (Kole et al. 2015). This work has necessarily involved research collaboration at the level of plant genetics and genomics, physiology, and population studies. As in the previous example however, consideration must still be given to impacts on biodiversity and ecology. What might be the impacts on broader communities that these organisms are in? What are the potential pitfalls of having a “super” population or a collection of “super” populations? While a significant amount of work has been done on selectively-breeding a variety of different plant and animal species, limited focus has been given to the impacts of these individuals on the ecosystem as a whole.

We would like to give a more detailed description of a third example, that of the microbiome.

The problem of scale for microbial communities

Until recently, researchers didn’t even think about microbial communities (a.k.a. microbiomes), since studying one microbe took a lifetime of work. Now, with new sequencing technologies and molecular methods in the past 20 years, the field of exploring microbial ecosystems has exploded. Usually investigators want to know how microbial communities are interacting with the host or environment and how emergent phenomena may result in disease or phenotype of interest. However, this is usually difficult as these systems are extremely complex, and the spatial and time components compound system analysis, not to mention space/time at micro- and macro-scopic scales (https://iopscience.iop.org/article/10.1088/1478-3975/aac473/meta). For example, a widely publicized case is one where a fecal microbiome transplant (FMT) patient died after the transplant. In that case, 22 patients received a fecal transplant from a healthy donor infected

with drug-resistant E. coli (https://www.nejm.org/doi/full/10.1056/NEJMoa1910437). Why was the donor healthy and why did only 2 out of the 22 recipients developed bacteremia? This is just one (yet severe) example out of many that shows that we cannot understand emergent properties without understanding the full system, in this case — the microbiome+host. Usually analysis of complex systems involves two types of analysis, that of bottom-up or top-down.

Bottom-up

To simulate a microbial community, one must understand the individual behavior of each organism, and very few organisms are currently well-understood. Each microbial community has individual organisms, with each having its own set of genes. Metagenome studies usually take a snapshot of this. While each strain in the community has its own genome, it is possible that it has received or shared genes horizontally across the community in the past or might in the future. Beyond the genetic component, these organisms are consuming, producing, and reacting with metabolites, which can be from environmental sources or each other. These reactions are catalyzed by enzymes, which are also made by genes, and also exchange energy. A snapshot of metabolites in a community can also be taken with recent technology.

Besides the limitation of having many unknown (meaning unknown to human discovery) organisms, genes, and metabolites, it is difficult to then simulate how thousands of these microbes may interact where emergent phenomena may result. These microbes are spatially (in 3-D) related to each other and temporally related. Examples are (1) how spatial configuration can affect *A. actinomycetemcomitans* and *S. gordonii* (https://academic.oup.com/femsle/article/366/11/fnz125/5513995) metabolite exchanges spatially and (2) how bacteria in various degrees of varying degrees of antibiotic production and degradation capabilities could stably coexist in various temporal modes (stable equilibrium, limit cycles, or chaotic oscillations) without spatial separation (Kelsic et al. 2015). In addition to the microbiome changing over time, genomes are evolving at the molecular level (https://www.sciencedirect.com/science/article/pii/S0092867419312309?via%3Dihub), so we must consider different scales over time.

If individual parameters are known and understood, it would be possible to simulate individuals, their metabolite fluxes, and their dynamic behavior. Most of the time, genome-scale metabolic models (GEMs) and constraint-based reconstruction and analysis (COBRA) are used to simulate microbial communities (https://www.biorxiv.org/content/biorxiv/early/2019/07/26/716126.full.pdf). And now using observed species and metabolites through high-throughput methods, there are ways to add this information to the models (https://www.nature.com/articles/s41564-019-0491-9). This opens up avenues for engineering control theory to think about modeling the system and _intervening_ in order to control the system (https://www.nature.com/articles/s41467-019-08890-y , https://royalsocietypublishing.org/doi/10.1098/rsif.2016.0380). However, we need to know much more about microbial communities and their spatiotemporal interactions with host and environmental processes.

Top-Down

The bottom-up deciphering of the microbiome will yield the greatest insight and pave the way for drug design and therapies once the microbiome functioning is fully understood and can be controlled. However, it is a mountainous task, and there are aspects that can be deciphered despite not having every detail simulated. Meta-analyses (https://www.nature.com/articles/s41467-017-01973-8 ,https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004977) can offer interesting discoveries of microbial communities on phenotype and other emergence properties. One of the first forays into discovering important metabolic pathways and protein families across an age-balanced meta-analysis showed that the microbiome plays a role in B12 vitamin synthesis (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3869192/). Also, it may be of interest to predict individual-level phenotypes within a larger system, and recent research is very interested in predicting organism phenotypes from strain-level genomic content and organization (e.g. antibiotic minimum inhibitory concentration of species strains that have varying levels of AMR (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5765115/)).

Advances in machine learning, such as deep neural networks and other statistical methods, offer promise to help to understand systems from a high level. For example, going beyond finding just important taxa, statistical topic modeling on 16S rRNA data can be used to identify groups of co-occurring taxa (sub communities), that may be cooperating, to perform functions (https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0219235) . Also, deep learning shows promise for predicting phenotypes of microbiomes (https://www.biorxiv.org/content/biorxiv/early/2018/01/28/255018.full.pdf). However, it is difficult to interpret deep neural networks and obtain biologically meaningful information that may give us insight to a microbiome. It would be nice to have a machine learning framework that could simultaneously be able to learn to predict phenotype while learning features at multiple scales — being able to identify taxa, genes, and even genetic mutations/metabolites that contribute to different overall phenotype, sub community functions, and even taxa. This would be an ideal top-down approach where the emergent function can be predicted but also to figure out why in a multi-scale way.

Potential

But once we simulate a community, what do we expect to see? We expect to see emergent phenotypes in host (https://science.sciencemag.org/content/365/6460/1405?rss=1) and environment (https://www.frontiersin.org/articles/10.3389/fmicb.2018.01929/full).
If the microbiome and its emergent behaviors can be understood, they can be manipulated. There can be a plethora of products to “direct” the growth of our microbial communities (https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-018-0592-8). Engineering microbial communities through synthetic biology will herald a new generation of therapeutics (https://www.sciencedirect.com/science/article/pii/S1931644816000046?via%3Dihub , https://msystems.asm.org/content/4/3/e00106-19). And finally, technology at the micro and even nano-scale (https://www.nature.com/articles/s41565-019-0589-5) may be needed, if medicines/therapies are not reliable and stable. Researchers are speeding towards therapies and interventions with the few understandings that have been deciphered so far.

Ideas that will help us think about our aims in addressing the problem of scale

Are there common processes across levels of biological organization?
Interactions

In endeavoring to find ways to breach the walls between hierarchical scales, one idea which will help is that of common processes. The processes common to all life are an important part of this idea, but to date we are still limited in our understanding despite them. We are proposing a new way of looking at common processes, taking advantage of the hierarchical nature of biology. At any given level of the hierarchy, the processes will not only be specific to what occurs there, but will also include everything that happens in the levels below it. Therefore, an ecosystems and a cell biologist have in common a genetic and molecular scale for comparison; genetic adaptation due to evolutionary drivers will be directly comparable, for example.

A fundamental part of what we are trying to suggest in looking for common processes lies outside the science entirely. It goes to the mindset of us, the scientists. In order to finally resolve the issues of scale and reintegrate biology, we need to be able to see beyond our specific system to the whole picture. Imagine, if you will, being in a mountain range. When you are in a valley, you only see the valley, the enclosing mountains and perhaps some other valleys through passes. When you stand atop the ridge, however, you can see many valleys and the mountains blocking them from each other; you see the whole range system. The valleys are connected not only by the passes, but also waterways, forests, trails, the very mountains themselves. This is no different from how we have sectioned off biology. We all live in our little valley, surrounded by what we perceive as mountainous barriers blocking us from other levels and fields of study. If we could simply hike up to the ridge, take in the whole landscape of what we are trying to do and see the connections between all the scientific valleys, we would be able to bring the stories together in a cohesive way. Some of these connections might be surprising or hidden.

Take the case of the elephant from Figure 1. The elephant will have organismal level processes. It also contains tissue, cell and molecular level processes, as well as microbial ones. Scaling down allows us to identify individual processes taking place at each level, heightening our understanding of the whole organism. If we scale up to the savannah, we add in population, community and ecosystem processes, but we do not lose anything we had with the elephant. Suddenly, we have the ability to move from genetic processes in the microbial community of an elephant to the savannah system as a whole.

Moving from organisms to ecosystems is a common scaling idea in ecology (Hofmann et al. 2010; MacMahon et al. 1981; Schramski et al. 2015). While it may seem a daunting task, we can take our organismal understanding and scale molecular processes up to the ecosystem level. This is not without precedent in biology (Andrén et al. 1999; Elser et al. 1996; Weigel 2012). There have been many arguments that this is too simplistic. Perhaps that is exactly what we need to bring biology back together, a simple model from which we can explore the details that make each of our systems beautiful and unique. There is an opportunity for new and fundamental work in finding the processes which bind scales together, either through data mining or new multi-scale experimental work. We may also be able to identify indirect

connections--emergent properties--between levels which can be easily navigated to give a bigger picture than any one of our disciplines creates on its own.

Emergent properties

Emergent properties can provide a common understanding regardless of the currency (see below) used in the process. They arise when you integrate movement rates of currencies between sources and sink.

- The heart passes information along in the movement of ions to send a signal to trigger a contraction or dilation. The rate and flow of this movement is vital to the transmission of the signal. The emergent property is incorporating the effect of all the chemicals and their results.
- Food webs are transfers of energy through the interactions of many species-species interactions. Looking at the transfer of one pair gives you information, the emergent property comes when you scale up to the entire ecosystem.
- Protein-protein or protein-nucleic acid interactions and networks within a cell will determine its emergent property: what type of cell it is. With changes in the networks, the cell type will change. This pattern is repeated as you scale up to the entire organism. Cells will come together to form the tissues (emergent property) and the tissues will organize to define the organism (emergent property).

Are there Guiding principles for these common processes?

- Guiding principles are mechanisms by which we can find emergent phenomena. For instance, one principle is that all systems evolve by natural selection. Variation and potentially stochasticity are therefore important commonalities for natural selection.
- As guiding principles can be specific to a system, questions arise.
    - When do guiding principles change over? Does this depend on the scale? Does this depend on the system?
    - When do the relative importance of currencies change over? That is, when is energy more important than matter and when is matter more important than energy as a currency? (For a fuller description of currencies, see section below.)
- One thing is clear, we need to know and define the scope of our experiments.
- Additionally, we apply guiding principles to how we conduct scientific research. The problem arises when the guiding principles of a scientific community tell us what we are doing research on, pushing new discoveries in one direction and leaving behind others. This further fractures biology.
    - Is there room to explore systems that do not exhibit the guiding principles? When do we put emphasis on these explorations?

Are there other, perhaps more nebulous, emergent phenomena?

How do we think about stochasticity?

- Stochasticity arises from noise in the process

- It impacts that exchange of a currency and makes it more difficult to see patterns if you don't account for it, but it doesn't mean that the underlying patterns aren't there.
- Does the importance of stochasticity change with level or organization?
    - Stochasticity is certainly important at some levels, but might be less relevant in some cases in which individual variation does not influence the outcome of a system.
    - Hypothesis: narrow variance around the mean at lower levels of organization and large variance around means at upper levels of organization.
        - Our goal in presenting this hypothesis is to illustrate how stochasticity is an important, although sometimes overlooked, component of the structure of a system.
    - Integrating this into models becomes really important and might be an important part of advances in models.
    - 

        - In the diagram above, the mean is represented as the vertical black bar and the variance around the mean is represented by the blue curve. Stochasticity and other mechanisms will contribute to the width of the curve. We use normal curves here to illustrate a general idea. The shape of the curve and the variance around the mean will be mechanism dependent.
        - In this case, stochasticity is the uncertainty in the system, demonstrated by variance around the mean. Higher stochasticity adds uncertainty into the system which makes it harder to model the system and harder to predict the outcome or a perturbation.
- Is stochasticity used as an excuse to avoid looking for rules due to love for the individuality of the systems?
    - Scheiner and Willig (2008) described seven principles of ecology and highlighted the value of looking for generalizations in complex biology systems.
- More research is needed to integrate stocascity into models of currency exchange and use (processes that use currency)
- Emergent phenomena will be constant, stochasticity will relate to the details of the system. Examples below.
    - Precisely what organisms occupy the different levels of an emerging food web is the stochasticity.
    - Vision papers from this workshop are all transfers energy and information resulting in the big questions for reintegrating biology, but each one varies based on the topic

- Genotype + environment to phenotype. As an example, the environment and maternal effects and variation in cellular processes affected how a genome is translated into a phenome because they introduce stochasticity into the developmental process. This inhibits genetic determinism.

Complexity, is it emergent?

- Much depends on how you define your system. At a starting point, the complexity is a characteristic of the system. In this case, Mazzocchi (2008) argued that emergence was a part of what defines the complexity of the system.
- However, while we tend to do biology through one lens, we try to put the results into a larger context. When we do this, the understanding becomes more complex. In this way, complexity is an emergent property.
- Example: the heart
    - Looking at the signalling pathways, there is a given amount of complexity in how they react to the presence of certain ions. The emergence of a response in a non-linear way, which leads to a reorganization of the components fits a complexity theory definition well (Mazzocchi 2008). This is not an emergent property.
    - As you scale up, you not only retain that level of complexity, but add more as you go: the transmission of signals, response of tissues, effect on whole organism. You shift one part of the system and the other parts of the system also shift. In this case, complexity is an emergent property.
- Where complexity falls on the emergence spectrum requires a researcher to be clear about what they are looking at and whether it crosses scales or not.

Is there a common currency across levels of biological organization?

To date, the idea of a biological currency has focused on the exchange of goods between two different systems. The export of matter or energy within and between ecosystems, for example (Brown et al. 2004; Teal 1962). Common currencies such as energy can help us to identify, understand and even bridge across processes to find ones common to many systems, leading to theories which form the foundations of modern biology. However, focusing on the exchange can limit our ability to work across scales. While information is being integrated across levels of organization and into physics (O'Connor et al. 2019), it is difficult for many biologists to consider how individual proteins might convey information to an entire ecosystem. If we alter our ideas of what a currency is in biological systems, we can begin to bridge those gaps. Information processing requires energy and materials, so it is possible for these currencies to be turned into each other, we simply need a way to do it.

At their core, currencies are a way to compare across levels of biology. For our purposes, we are expanding the definition to include currencies as things with value in a system. In this way, currencies do not have to be exchanged, and can be evaluated across scales. Some currencies are universal, energy, matter, information. Others may not be, but can still be related to each other and give us similar information via our identified, indirect pathways.

To illustrate, consider for a moment the basic idea of buying a good from overseas. The monetary currency used in the buyer's country is not the same as that of the seller; however, it has an exchange rate which allows the good to be accurately and precisely valued in two different ways and traded between the two parties. Taking this view, such an "exchange rate" would allow biologists to consider how replication at the protein scale directly relates to growth of populations.

The idea of finding common, value-based currencies and exchange rates is not one which intends to forge something completely new to biology, at least not in all cases. Many of the established theories can be used as the building blocks for finding universal currencies or ways to link currencies over multiple scales. For example, using metabolic theory we may be able to link the replication of proteins to that of cells or tissues, allowing biologists working on population growth dynamics to use information from those studying gene expression. This flow of currencies between entities helps us to link, either directly or indirectly, the different levels of biology.

One method which may be especially useful in identifying common processes and connecting scales is network analysis. Network analysis seeks to find connections between system elements. It has successfully been applied within individual levels, finding pathways and interactions in both cell and molecular biology (Albert 2005; Deng et al. 2012). The application of this method for the biological hierarchy could lead to breakthroughs in linking different scales. Along similar lines, identifying common processes in scientific research can establish best practices that cover multiple scales, making data directly comparable. Even for those scales which best practices cannot connect, standardizing units and formats carries great value on its own.

Summary Statement:
With a proper framework of principles and currencies, accounting for stochasticity and considering complexity, we can link the various scales of biology, providing greater understanding of how systems will respond.

What happens when you do this poorly?

There are dangers to this approach, as there are for all sciences. We must be sure that our methods are the most appropriate for the questions we are trying to answer, including all necessary processes which link our levels. We must also be sure that we are clear in our definitions of our processes and currencies so as not to lead to type I errors. A recent paper suggesting a small causal link between heritable intelligence traits and income in white citizens of the United Kingdom, for example, equated the currency of "intelligence" with a completely different currency, one's score on a cognitive test (Hill et al. 2019). This was done even though it has been shown elsewhere that such exams truly only score a person's ability to take certain kinds of tests. It would have been more accurate to say a person's heritable ability to take a cognitive exam has a small positive effect on their income. Even then, your ability to take such tests is dependent upon the socio-economic background from which you come; richer families are more likely to be able to hire test tutors, and have greater opportunities which lead their

children to be more prepared. These processes were, for the most part, unaccounted for though the authors readily acknowledge the impact external factors have on a person's success as viewed through income.

Barriers

Scientific barriers (the flip side of the coin "problem we can solve by addressing the paper aim):

Most of the problem of scale can be considered a scientific barrier, and a number of barriers have been mentioned so far. Here we identify three common problems faced in the gathering of data which can bridge scales.

- *In vitro* vs *in vivo*: how to cross the bridge between the two.

  Many questions in biology are approached using in vitro methods. However, this begs the question of how applicable the results are in vivo/in hospite. In the coral literature, for example, the algal symbionts are often studied using cultures. This is done despite the fact that it is well known the cells behave in a vastly different way within the host. Tradition dictates that all results are qualified as coming from cultures, and only have implications for what happens in the symbiosis. However, in asking fundamental questions concerning the response of corals to perturbations to their system, understanding the symbiosis is essential. Instead of minimizing any applicability, finding common processes and currencies could allow researchers to bridge the gap between in vitro and in hospite. Likewise, collecting data in the same format can enhance the comparability of the information. All of this will require a better understanding of these systems.

- The use of biological models. Knockout mice, for example, are not necessarily good representations of humans.

  Selection of an appropriate model system for analysis depends greatly on the focus of the research. In plants for instance, *Arabidopsis thaliana* is frequently used due to it short generation time, ease of maintenance, and well-characterized genome. While studies in *Arabidopsis* are highly useful for uncovering or describing plant biological processes in general, *Arabidopsis* has little value as an agricultural model. Homology between *Arabidopsis* and crop species such as *Solanum lycopersicum*, *Glycine max*, and *Zea mays* allows for application of discoveries made in *Arabidopsis*, but studies are also required in these model crop species before the impact of these discoveries on agriculture can be determined. Similarly, *Arabidopsis* is not a good model for population-level studies in an ecological sense as it's almost exclusively cultured in a lab or greenhouse setting.

  Rather than immediately selecting a model species because of its convenience, consideration must be given to what type of data that system provides. A second consideration is how distant, evolutionarily, the model system is to the system we wish to apply the results to. Intelligent phylogenetic sampling is needed to ensure we can find commonalities. However, we can run into ethical problems when finding a good model system and depending on what we are testing.

- Data collection needs to be done in common formats. This will increase the applicability of information from one system to another.

What structural/institutional barriers exist?

It would be very valuable to attempt projects that integrate multiple levels of organization, as depicted in Figure 1. However, there are multiple barriers to accomplish this goal:

- **Cultural**: Current trends incentivize the formation of interdisciplinary groups, in which each disciplinary expert passes information at the disciplinary boundary. It would be more efficient for scientific discovery to have transdisciplinary scientists who are experts in one are and have proficiency in other areas (e.g. an expert in biology who could understand non-trivial programming, mathematics, and statistics). But, how to train them?
- **Technological**: There are many standards for data dissemination that increase the labor required to integrate data. There is a lack of mathematical and computational techniques that can model at various scales to understand when emergent behavior will result.
- **Infrastructure**: Large and complex projects that integrate multiple traditional areas of biology face the problem that the data generated has to be distributed across a constellation of databases. Traditionally there is space for "metadata" in individual database submissions, which could be used to document an entire experiment. This would lead to much duplication, and potential inconsistencies.
- **Funding**: There is little cross funding across NSF and NIH, leading to "NIH communities"and "NSF Communities" that are not fully aware of each other, with NIH having many more resources to fund database efforts. This is particularly detrimental for informatics resources, data sharing and quantitative analysis geared toward integrative efforts. The solution necessarily passes through standardization of data (another thread in itself). It would be nice to have structured mechanisms to ensure that data and analysis pipelines cross disciplinary and funding barriers. That is, multi-funding agencies working together.

Proposed concrete solutions to barriers

- **Cultural:** 1) Train transdisciplinary scientists early in careers, for instance by expanding RCN training offers, REU opportunities, etc. 2) Have transdisciplinary review panels for grants
- **Technological:** 1) Enhance machine learning capabilities for modeling; 2) Increase access to needed computational resources; 3) Standardize data dissemination requirements
- **Scientific:** 1) Determine relationships between in vivo vs in vitro; 2) Detailed phylogenetic work; 3) Use the appropriate organismal model system for the question, regardless of whether it is the “traditional” model system used
    - Thing that helps:
        - Using non-traditional models in the field
        - Understanding the limits of models and where those limits exist

        - Intelligently choosing systems. -> comparative approach can help identify (See Intelligent Phylogenetic sampling above)
        - Using similar techniques to collect data in different systems to allow use of a comparative framework.
        - Krogh's Principle

    4) Apply network analysis to working across hierarchical scales; 5) Dynamical/control systems (https://www.nature.com/articles/s41467-019-08890-y, https://royalsocietypublishing.org/doi/10.1098/rsif.2016.0380, https://academic.oup.com/femsre/article/42/3/273/4794944); 6) Further develop process based models to include new common processes and currencies; 7) Employ computer models to assist in finding common processes and currencies; 8) Develop Best Practices across broad topics bridging scales
- **Infrastructure:** 1) Funding agencies could catalyze the formation of "experiment databases", where the integration of multiple disjointed components is documented, including results of quantitative analysis; 2) Database of databases to allow scientists to see what information is already available across disciplines; 3) Create more challenges and contests, such as CAFA and CAMI, which address database shortcomings
- **Funding:** 1) Better communication and collaboration between funding agencies; 2) Another idea is to form a "Rules of Life" Division within NSF to focus on multidisciplinary biology; 3) Continue support for the current multi-scale programs

References not otherwise linked

Albert, R. 2005. Scale-free networks in cell biology. *J. Cell Sci* 118, 4947-4957. doi: 10.1242/jcs.02714

Andrén, O., L. Brussaard and M. Clarholm. 1999. Soil organism influence on ecosystem-level processes--bypassing the ecological hierarchy? Applied Soil Ecology, 11(2–3), 177-188 doi: 10.1016/S0929-1393(98)00144-9

Deng, Y., Jiang, Y., Yang, Y. *et al.* 2012. Molecular ecological network analyses. *BMC Bioinformatics* 13, 113 doi:10.1186/1471-2105-13-113

Elser, J.J., D.R. Dobberfuhl, N.A MacKay and J.H. Schampel. 1996. Organism size, life history and N:P stoichiometry: Toward a unified view of cellular and ecosystem processes. *BioSci* 46(9), 674-684. doi: 10.2307/1312897

Hill, W.D., Davies, N.M., Ritchie, S.J. *et al.* 2019. Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income. *Nat Commun* 10, 5741. doi:10.1038/s41467-019-13585-5

Hofmann, G.E., J.P. Berry, P.J. Edmunds *et al*. 2010. The effect of ocean acidification on calcifying organisms in marine ecosystems: An organism-to-ecosystem perspective. *Ann. Rev. Eco Evo Sys* 141, 127-147. doi: 10.1146/annurev.ecolsys.110308.120227

MacMahon, J.A., D.J. Schimpf, D.C. Andersen, K.G. Smith and R.L. Bayn Jr. 1981. An organism-centered approach to some community and ecosystem concepts. *J. Theoretical Bio* 88(2), 287-307. doi: 10.1016/0022-5193(81)90077-1

Mazzocchi, F. 2008. Complexity in Biology. *EMBO Rep* 9, 10-14. doi: 10.1038/sj.embor.7401147

O'Connor, M. I., Pennell, M., Altermatt, F., Matthews, B., Melian, C., & Gonzalez, A. 2019. Principles of ecology revisited: integrating information and ecological theories for a more unified science. *Frontiers in Ecology and Evolution,* 7, 219.

Scheiner, S. M., and Willig, M. R. 2008. A general theory of ecology. *Theor. Ecol.*1, 21–28. doi: 10.1007/s12080-007-0002-0

Schramski, J.R., A.I. Dell, J.M. Grady, R.M. Sibly and J.H. Brown. 2015. Metabolic theory predicts ecosystem properties. *PNAS* 112 (8) 2617-2622; DOI: 10.1073/pnas.1423502112

Weigel, D. 2012. Natural Variation in Arabidopsis: From Molecular Genetics to Ecological Genomics. *Plant Physiology*, 158 (1) 2-22; DOI: 10.1104/pp.111.189845