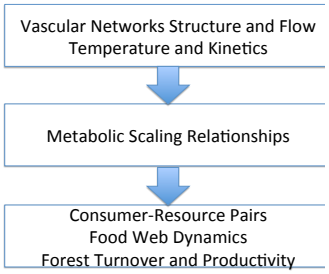
Research Statement

I work at the interface of ecology, evolution, medicine, mathematics, and computation. I aim to quantify and to understand the possible functions and forms of biological systems that result in the extraordinary diversity in nature. I have long studied and continue to research impacts of climate change on ecosystems, trait diversity and trait driver theory, rates of evolution, branching networks, and metabolic scaling theory—the effects of size and temperature on organismic physiology, evolution, and ecology. These research directions have also led to projects and progress on the structure of vascular systems, neuronal branching, rates of tumor growth, and the functions of sleep.

My overall research goal is to combine novel mathematical models with newly collected or analyzed empirical data to understand how this diversity is organized, constrained, and controlled across multiple levels of biological organization. My research program typically involves a few key steps. First, I analyze large empirical data sets to discover recurring, statistically significant patterns that hold across multiple scales. Examples of patterns I have found are allometric scaling curves for metabolic rate, sleep time, and population growth. Second, I construct mathematical models based on hypotheses for the underlying causes of the patterns observed in step 1, and I then derive further predictions from these models. Third, I test these new predictions using extensive empirical data obtained either by compiling data from thousands of papers in the literature, using online public databases, or as part of ongoing experiments by my collaborators. Fourth, I revise my models based on deviations of the data from the original model's predictions and devise new experiments. Much of my research focuses on the construction of mathematical models and analysis of empirical data to discover how physiology influences biological structure and dynamics.



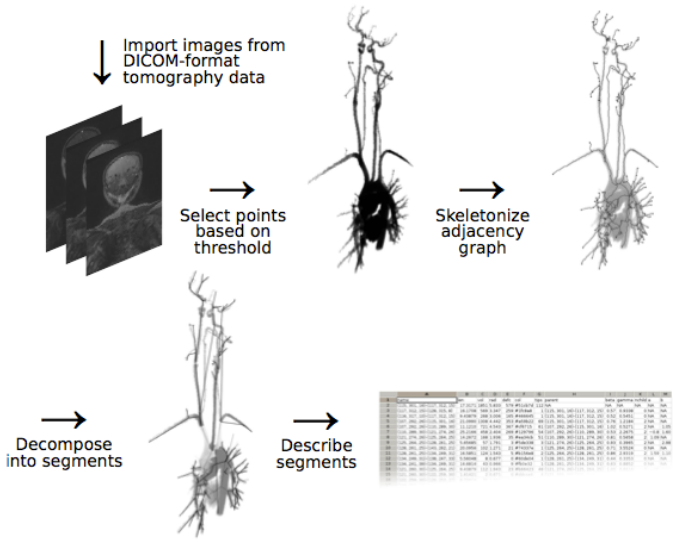
In my previous work, I constructed mathematical models and compiled and analyzed extensive empirical data to explain a suite of allometric scaling relationships—how organismic properties (e.g., metabolic rate or lifespan) change with body size and body temperature across species. These relationships were proposed yet unexplained for nearly a century. The theory focuses on the structure and fluid dynamics of vascular networks and biochemical reaction kinetics in animals and plants. I have also developed the first theories that mechanistically link biological scaling theory to population growth, tumor growth, sleep time, and cell size.

What governs the diversity and evolution of branching for vascular networks and neurons?
Within an organism there can be billions of vessels that range in size from meters (whale aorta) to microns (capillaries). Within humans, blood flow rate can vary by a factor of 250 as the aorta branches down to the capillaries. The proportional change in vessel radii and lengths determines the rate of oxygen supply to tissue. In particular, changes in vessel radii have large effects on blood flow rates, and changes in vessel lengths and branching angles enable a vascular network to span the body and have capillaries close enough to feed all cells. These proportional changes in vessel radii and lengths can be mathematically encapsulated in scaling ratios and scaling exponents that are used to predict not only blood flow rate, but also metabolic, growth, and death rates.

Thus, the vascular structure has profound effects on physiological processes, and detailed knowledge of vascular structure can be used to model and quantify these effects. My group has developed software—Angicart—that we use to analyze angiographic images and automatically extract an unprecedented amount and diversity of data on vascular networks, including vessel radii, lengths, volumes, branching numbers, branching angles, scaling ratios,

and scaling exponents. Angicart can obtain data in about 20 minutes that previously took months using silicone or polymer casting techniques. Moreover, these data are obtained in vivo and non-invasively, so the vessels are not damaged during the measurement process and can potentially be measured as they develop and grow. All previous knowledge of vasculature before was based on limited data of questionable quality from cats, pigs, and humans. My results thus far have looked at vasculature from human head and torso (Magnetic Resonance Images (MRI), lung cancer (MRI and PET), mouse lung (micro-Computed Tomography (CT)), including wild type mice and mice that have been modified so that their vasculature is either over- or under-branched. We also have now extracted and analyzed branching of neural structure for a variety of neuronal cell types using existing public databases of neuron images.

I have illuminated principles of existing models that are wrong and others that seem well founded. For instance, a commonly used space-filling principle predicts changes in vessel lengths that differ dramatically from our data. We aim to explain these differences by identifying and understanding systematic patterns in the branching angles between vessels and in the distribution of asymmetric branching—a vessel that branches into two children vessels of different sizes. Current models of whole vascular systems ignore branching angles and assume symmetric branching—children vessels within the same level always possess identical sizes and flow rates. Using our newly collected data, we measure asymmetry by calculating the ratio of the smaller child vessel radius or length to the larger child vessel radius or length, resulting in a value of 1 for symmetry and values near 0 for highly asymmetric branching. For vessel radii, the distribution of this asymmetry ratio shows a peak near 1 with a strongly skewed shape going away from the peak, implying a tendency towards symmetric branching. For vessel lengths the distribution of this asymmetry ratio shows no clear peak and appears close to a uniform distribution, implying highly asymmetric branching. These results hold for both human head and torso as well as mouse lung data that span length scales down to mm for MRI and 10-100 micron for micro-CT. I have developed new models of whole vascular system that include asymmetric branching—a vessel branching into child vessels of different sizes or flow rates—and more detailed fluid mechanics that predict curvature in the direction observed in empirical data. Despite vascular models being studied for decades along with several books published on the subject, there had been very sparse data on vascular structure and very few direct tests of the assumptions made for mechanistic models.



Concurrent with this work on vascular networks, over the past few years my group and I have extracted branching data for axons and dendrites in neural systems to characterize empirical branching patterns. Because the optimization principles for neurons likely differ substantially from vascular networks and even between neuronal cell types, we developed models that allow us to include different principles and even vary the weightings/tradeoffs between different principles—for instance, minimization of power versus response time versus maximal conduction/information processing versus material limitations and connectivity. Consequently, we used Lagrange multiplier calculations to derive predictions for how neuronal branching properties change depending on whether electrical conduction energy, electrical conduction time, charge conservation, or some combination of these is the fundamental driver of neural structures. We find that both time and energy play important roles, and their relative roles depend on the

types of neural projection (axon versus dendrite), whether the projection is myelinated or unmyelinated, and the function of the neuron—such as peripheral nervous system neurons for which quick responses are needed and under strong evolutionary pressure. Taken together, this work helps provide a more general framework for understanding and deriving the biological and physical principles that determine the structure of neuronal branching, a problem that dates back to the beginnings of neuroscience and Santiago Ramon y Cajal.

Crucially, we are now implementing machine-learning methods to distinguish among branching patterns—plants versus animals, brain versus head and torso, angiosperms versus gymnosperms, peripheral nervous system neurons versus astrocytes, and tumor vasculature versus healthy vasculature (see more in sections below). By processing the data to calculate scaling ratios at branching junctions—which from theory should drive fluid flow, power loss, and space filling—our theory-informed machine-learning methods perform much better than traditional methods applied to just the raw data with no input from theory. Our new models are essential to investigate the selective and developmental forces that mold vascular networks and will help reveal intriguing developmental and environmental influences.

How does tumor vasculature differ from and interact with healthy host vasculature?
There is a need for testable mathematical models that describe and predict the vast amount of experimental data generated by cancer labs, particularly if these results are to reach clinical significance. To model tumor growth, I extended the scaling theory for the cardiovascular system to model the abnormal vascular development in tumors and how that interfaces with the host vasculature. I used this extended model to predict the growth dynamics of a variety of tumors. This theory provides the first predictions for the scaling of tumor metabolic rate with tumor and host size, and thus, how these influence tumor growth trajectories. Predictions are consistent with available data. This theory represents a general framework for understanding detailed vascular properties of tumors, including degree of necrosis, rates of nutrient supply, and dependence on host size and cellular properties. Moreover, the model enables quantitative comparisons of tumor growth across species and connects whole tumor phenomena from cellular to organismal levels. As a result, it is possible to identify stages when tumor growth is most rapid and to shed light on Peto's paradox—why whales and humans are less cancer prone than mice. These findings have potential impacts on drug discovery and intervention, and my group and I worked to extend the model to predict the number of proliferative versus quiescent cells and how this affects tumor treatment by drugs that target dividing (proliferative) cells as well as recurrence of tumors.

Using our imaging software and branching classification techniques from sections above, we extracted and compared data with theoretical predictions for vascular branching to distinguish lung tumor vasculature from healthy vasculature and with a proof-of-principle and long-term goal of using these methods to categorize tumors according to stage, aggressiveness, primary tumor tissue, and potential treatments.

How do consumer-resource interactions and invasions respond to temperature change?
Invasive species can quickly alter biodiversity levels and community composition, so management strategies for biodiversity critically depend on the anticipation of invasions. One of my goals is to understand how global warming will affect the control of population abundance and the ability of species to invade. More generally, I want to use models of consumer-resource interactions to understand when invasions can occur and why certain consumer-resource links are over-represented in real food webs. I have worked to develop theory to realistically combine biological scaling with consumer-resource interactions. The theory predicts how encounter rate depends on body size and temperature in a way that accounts for differences in behavior (e.g., active capture or sit-and-wait) and environmental influence on organismal physiology (e.g., ectotherm-ectotherm versus endotherm-

ectotherm interactions). These predictions can then be used as input to Lotka-Volterra-type equations that now naturally and explicitly include the scaling of body size and body temperature. Using these modified equations, I can in turn predict the temperature and body size dependence of consumption rate, equilibrium population size, optimal consumer-resource relationships, and conditions for species coexistence.

To test these assumptions and predictions, my group and I surveyed the literature and compiled empirical data—tens of thousands of points—that represent a huge diversity of consumer-resource functional types, species, habitats, and trophic groups in order to determine the effects of body temperature and body size. We showed that there is a central tendency of approximately 0.65 eV for activation energies and made the first observation that those distributions are strongly right-skewed. I intend to use deviations from these patterns to gain a deeper understanding of how consumer-resource pairs are impacted by physiology, ecology, and evolution.

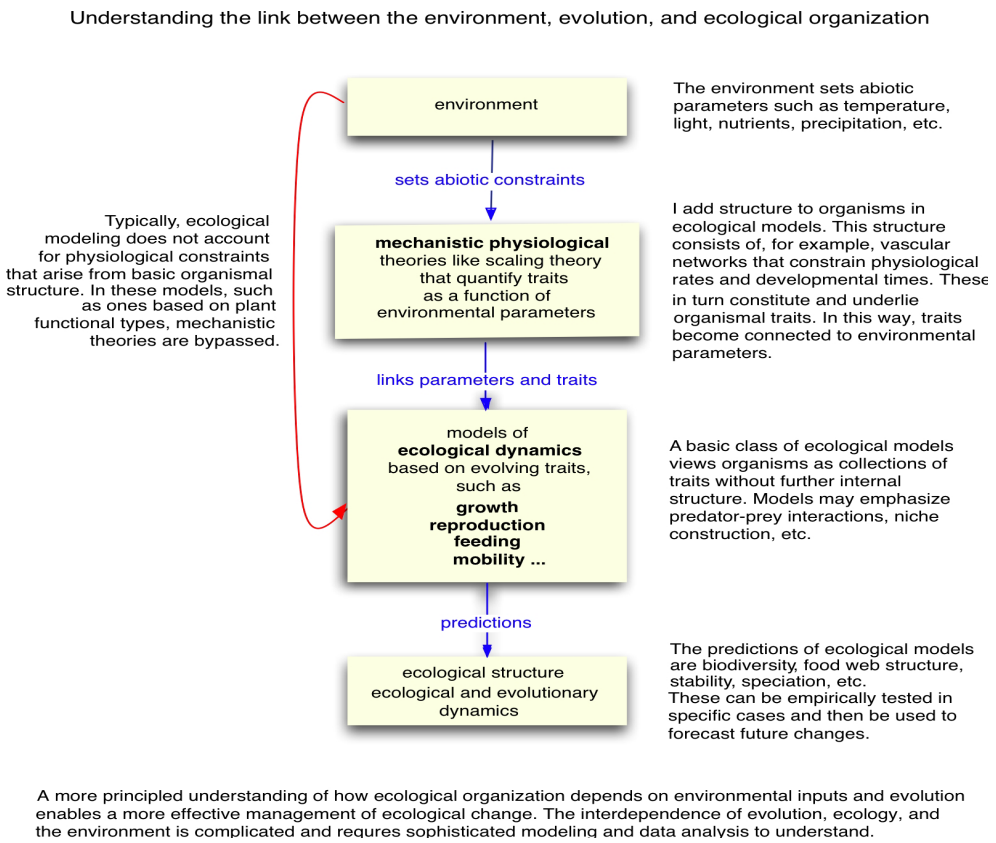
In particular, we observed stronger selection on traits related to negative motivation and an increase in the variance of activation energies across levels of biological organization—from cells to communities. These results led to a thermal formulation of the life-dinner principle—stronger selection to run for your life than to catch your dinner. We also identified strong differences between interactions in two (terrestrial and benthic) versus three (pelagic) dimensions. This theory can also be incorporated into food webs that have a mixture of these dimensionalities, and predictions for different dynamical patterns can also now be investigated using recent, massive online databases for species abundance time series. These future directions will further ground the theory in real empirical ecology and also undoubtedly lead to further refinements of the theory.

How does ecosystem structure respond to climate change?
I am helping to develop a framework for studying real ecosystems in which a changing environment affects a suite of interconnected traits. My fundamental goal with this work is to understand how species and ecosystems can respond to rapid environmental change. I partly developed a trait-based approach to explore the effects of a changing environment on biomass distributions and community structure for traits. Using this framework requires carefully constructed fitness/growth functions that couple physiological traits to the environment. With this model, trait distributions can be calculated numerically or using moment-closure approximations to follow the population dynamics and diversity through time. I have devised a framework that can be used to study multiple, correlated environmental variables (e.g., temperature and precipitation) and associated multiple, correlated traits (e.g., optimal temperature for growth and water use efficiency) to discover non-intuitive effects that arise from correlations.

My most recent work on this topic has been to construct new mathematical methods and numerical simulations to investigate when ecosystem size and complexity tend towards collapse as opposed to losing a single species to extinction that can be potentially reintroduced or that has little impact on the rest of the species in the system. Traditionally, work on questions related to the diversity-stability debate focuses on the mathematical notion of stability, which deems a system unstable if the abundance of even a single species is changed after a perturbation that then relaxes back to a steady state. I argue that this notion of stability is extremely stringent, which is why I have worked to mathematically discern between single-species extinctions and system chaos or collapse.

In this direction, collaborators and I numerically found a large intermediate range for food web complexity and size that allows single-species extinctions but does not tend towards collapse. We predict the size of this intermediate range using approximations from our mathematical methods based on extreme-value statistics. Further building on this, we showed that a constraint on the total consumption rate of a

population or species (as opposed to allowing unlimited consumption and intake as assumed in many current models) can be implemented by requiring a constant row-sum in the interaction matrix and that this leads to much more stable systems, a result that had not been previously appreciated either ecologically or mathematically.



Even more recently, I've developed new mathematical theory and eigenvalue spectral bounds to show these ideas extend to even small non-random matrices, and that it's easy and direct to compute my new bounds for an interaction matrix of any size. This new theory yields a much more ecologically and geometrically intuitive understanding of matrices, eigenvalues, and ecological constraints. Intriguingly, it also allows one to directly connect Damuth's rule and the Energetic Equivalence Rule to constant row sums of the interaction matrix that I now show leads to zero tradeoff or paradox between diversity and stability. Indeed, with this constraint, increasing diversity and complexity can sometimes improve stability of the food webs.

Finally, my group and I have also coded a numerical model for constructing food webs that includes self-regulation, trophic topology, trophic interaction strength, and a consumption rate constraint that can all be parameterized and jointly coordinated via body size effects tied to log-normal distributions. By comparing with empirical food webs, we showed that our new model does a better job of predicting food web stability than any previous models and any subset of possible models that only implement some of these constraints. We also correctly predicted how the empirical connectance of food webs changes with food web size, which to our knowledge is the first method to correctly predict this important property.

My next step is to combine this new theory on stability with the work in the previous section that tied temperature to consumer-resource interactions in order to predict and understand how changes in temperature will impact diversity and stability of food webs in light of climate change. This has been an abiding and primary goal of mine for two decades, and I feel I am finally perfectly poised to tackle it.

What is the frequency and impact of higher-order interactions among drugs or ecological stressors?
I have helped to build a general framework for expanding beyond the pairwise interactions that are often measured and displayed for networks. This general framework derives how to measure emergent interactions among more than two components, with a key part of that being to disentangle emergent interactions from interactions that arise purely due to combinations of pairwise parts. Another key part is to explain methods for rescaling the emergent interaction measure that normalizes for the strength of lower-order effects so that the magnitude of these interaction measures can be easily interpreted in terms of information about the intrinsic interactions and not the single-drug or single-stressor magnitudes. These general principles and methods can be applied to derive measures based on any starting definition of what is an interaction (i.e., covariance, mutual information, ANOVA, etc.)

I have applied this framework to experimental data on antibiotic interactions to show that emergent higher-order interactions are common and can be seen to be shifting to become more antagonistic such that successive antibiotics weaken the effects of previous antibiotics. Moreover, I have worked on how these types of interactions affect the shapes and roughness of fitness landscapes such that some drug combinations can be used to slow the evolution of drug resistance by bacteria.

My group, collaborators, and I further showed that antibiotics substantially alter the temperature responses of *E. coli* and that temperature adaptation of *E. coli* can affect evolution of antibiotic resistance. At a larger level of clustering analysis, this can be leveraged to provide general evidence that stress responses of *E. coli* to certain classes of antibiotics can be mapped onto stress responses to extreme cold or high temperatures, suggesting that some antibiotic stress responses were co-opted from ancient stress responses to temperature. Following up on this, I helped to write a review of the literature that demonstrates how climate change and specifically temperature changes could impact the rate of emergence of antibiotic resistance and also possibly help to identify geographic locations for emergence of specific types of pathogens, depending on what temperature stressors are created by the changing environment.

More generally, these findings can apply to genetic epistasis, predator-prey interactions, and social systems. I have shown that this newly derived interaction measure performs better than ANOVA, and that applying ANOVA to data with a small number of replicates necessarily assumes no interactions (via its implicit assumption of pseudo-replication and thus constant variance—across the range of the axes—that violates the actual systematic differences in variances when interactions are present). This misapplication of ANOVA often leads to missing cases of antagonistic interactions and overreporting synergistic interactions in ecology. Taken together, these misclassifications could have important consequences for studies of the combined impacts of climate change, human impacts, ecotoxicity, and more.

What is the Function of Sleep?
Sleep—like eating and breathing—is one of the most ubiquitous phenomena in all of biology, yet its function remains hotly contested. Many hypotheses exist for the function of sleep, but these are mostly based on qualitative models and possibly tests of correlations using empirical data. I developed some of the first mathematical models and truly quantitative, comparative tests for sleep function. Different models correspond to different predictions for the scaling exponent for how sleep times vary with body and brain size. Using empirical data for mammals (97 species and 79 genera), I showed that sleep is driven by rates and processes in the brain, not the body, and that hypotheses for sleep related to neuronal repair are the ones that are most consistent with empirical data. I continued this work by analyzing developmental sleep data in humans. As humans grow from birth to adulthood, their sleep times, brain size, brain metabolic rate,

and neural reorganization all change, providing a further test of sleep theories and a potential means for quantifying the relative importance of neural reorganization versus repair. With this in mind, I developed theory that helps to disentangle the necessary evolutionary and developmental functions that sleep currently serves in mammals and other organisms. In short, during early development (about the first 3 years of life for humans) REM sleep duration is primarily driven by neural reorganization and synapse formation, whereas across adult mammals, sleep duration is primarily driven by the need for metabolic clearance and repair. My current work on sleep takes a two-pronged approach. One prong is working with a lab that has many species of *Drosophila* and manipulates temperature conditions to measure how it affects sleep frequency and bout length. In this way temperature can be used to tune metabolic rate, to examine how sleep times, damage rate, and synapse formation change with temperature, and to then test my theory both across development and evolution to discover how temperature and metabolism relate to the function of sleep. The second prong is looking at sleep times for human fetuses before birth compared with sleep times of premature babies—ranging from more than 3 months to about 3 weeks premature. By contrasting these cases, we can determine the influence of the mother's metabolism and integrated vascular network on the fetus versus that of the premature babies. Moreover, we plan to study how differences in sleep times and sleep patterns for premature babies can affect their long-term prognosis and catchup in terms of brain development and cognitive abilities.

Summary

I want to derive novel mathematical models that I test with empirical data to understand how diversity is organized, constrained, and controlled in ecological systems, how it is influenced by organismic physiology, and in turn, how climate change will affect ecosystem dynamics and stability. Because of my interests, I feel Cornell University—with its exceptional core faculty in Computational Biology as well as outstanding faculty across biology, mathematics, physics, and computer science—would be an ideal place to achieve my research goals. Indeed, I've previously collaborated with Courtney Murdock (Dept. of Entomology) on effects of temperature on mosquito-borne disease, and I've had deep and rewarding discussions with Steven Strogatz (Dept. of Mathematics) about my research on sleep and complex systems more generally. I would contribute to the department by bringing an array of mathematical methods, diverse experience constructing quantitative models from conceptual ideas, a suite of techniques for analyses of large empirical and comparative data sets, knowledge of how to cull the literature for data, ideas for experiments, and a novel perspective on many problems across biological fields, from ecology to evolution to neuroscience to physiology.