

Algebraic and Combinatorial Approaches in Systems Biology 2015

ACSB 2015

May 22 – 24, 2015

Center for Quantitative Medicine

University of Connecticut (UConn) Health

Farmington, CT

Organized by

Seda Arat

Cory Brunson

Elena Dimitrova

Luis David Garcia-Puente

Brandilyn Stigler

Paola Vera-Licona

Supported by

Center for Quantitative Medicine

National Science Foundation

- 08:30–09:15 **Registration**
- 09:15–09:30 **Opening of the Conference**
- 09:30–10:20 **Invited Talk** Algebraic Reverse Engineering of Biological Networks
Michael Stillman, Cornell University
- 10:30–11:00 **Coffee Break**
- 11:00–11:50 **Invited Talk** The Art of Robust Networking
Karen Schlauch, University of Nevada, Reno
- 12:00–14:00 **Lunch Break**
- 14:00–14:50 **Poster Blitz**
- 15:00–15:50 **Invited Talk** Reducing Ambiguity in Gene Regulatory Network Inference via Grobner Bases
Brandilyn Stigler, Southern Methodist University
- 16:00–16:30 **Coffee Break**
- 16:30–16:50 **Contributed Talk** Algebraic Systems Biology and Experimental Design for the Wnt Signaling Pathway
Heather A. Harrington, University of Oxford
- 17:00–17:20 **Contributed Talk** Topological RNA Structures
Christian Reidys, Virginia Tech
- 17:30–19:00 **Poster Session**
- Some Combinatorics of the SIR Disease Spreading Dynamics on a Random Graph
Andrei Bura, Virginia Tech
 - Geometric Combinatorics and RNA Secondary Structure Prediction
Andrew Gainer-Dewar, UConn Health
 - The Regulatory Network of DNA Mismatch Repair
Claus Kadelka, Virginia Tech
 - A Difference Equation for Tracking Perturbations in Systems of Boolean Nested Canalizing Functions
Elena Dimitrova, Clemson University
 - Stratification and Enumeration of Boolean Functions by Canalizing Depth
Qijun He, Clemson University
 - Emergent Node Tree Structures for Multi-scale Modeling of the Dynamics of Interacting Agents
Ruchira Datta, Mathematical Biosciences Institute
 - Investigation of Parameter Space for Chaos
Sherli Koshy-Chenthittayil, Clemson University

- 08:30–09:20 **Invited Talk** Combinatorial Interventions for Control Tasks in Large-Scale Signaling Networks
Paola Vera-Licona, UConn Health
- 09:30–10:20 **Invited Talk** Discrete Dynamic Modeling Elucidates the Outcomes of Signal Transduction Networks and Helps to Control them
Reka Albert, Pennsylvania State University
- 10:30–11:00 **Coffee Break**
- 11:00–11:20 **Contributed Talk** Optimal Control Methods for Stochastic Boolean Networks
David Murrugarra, University of Kentucky
- 11:30–11:50 **Contributed Talk** Multibranch Loop Parameters in RNA Folding Algorithm
Qijun He, Clemson University
- 12:00–14:00 **Lunch Break**
- 14:00–14:50 **Invited Talk** Algebraic and Discrete Mathematical Biology for Undergraduates
Raina Robeva, Sweet Briar College
- 15:00–15:20 **Contributed Talk** Distribution of Motifs Generated by a Stochastic Grammar for RNA Folding
Svetlana Poznanovikj, Clemson University
- 15:30–15:50 **Contributed Talk** Reduction of Combinatoric Calcium Release Site Models via Fast/Slow Analysis and Genetic Algorithm
Yan Hao, Hobart and William Smith Colleges
- 16:00–16:30 **Coffee Break**
- 16:30–17:20 **Invited Talk** Discrete Homotopy and Homology Theory for Metric Spaces
Helene Barcelo, Mathematical Sciences Research Institute

- 08:30–09:20 **Invited Talk** Rational Design of Antibiotic Treatment Plans
Bernd Sturmfels, University of California, Berkeley
- 09:30–10:20 **Invited Talk** Airway Inflammation: Does it Add Up?
Shernita Lee, University of North Carolina Chapel Hill
- 10:30–11:00 **Coffee Break**
- 11:00–11:20 **Contributed Talk** Combinatorial Optimization Methods for Agent-based Modeling
Matthew Oremland, Mathematical Biosciences Institute
- 11:30–11:50 **Contributed Talk** Network Architectural Conditions for Noise-induced Biochemical Oscillators
Jaewook Joo, University of Tennessee
- 12:00–12:15 **Closing of the Conference**

Talks

Algebraic Reverse Engineering of Biological Networks

Michael Stillman

Cornell University; mike@math.cornell.edu

In this talk, we present the state of the art in algebraic methods for reverse engineering of biological networks, emphasizing the mathematical algorithms used (many devised by members of Laubenbacher's group). We describe the current state of software for implementing these ideas. Finally, we discuss open problems in computational algebra and in modelling that arise from reverse engineering, and consider possible future directions.

The Art of Robust Networking

Karen Schlauch

University of Nevada, Reno; schlauch@unr.edu

Networks provide classic, easy, and effective models to study complex biological systems. With the introduction of microarray technology in the late 1990's and in the advent of next-generation whole-genome and whole-transcriptome sequencing technology, fast and robust tools are needed to examine organisms at the systems level.

Using networks to map an organism's gene (or protein) relations provides researchers with a bird's eye view of possible associations and regulations in whole-omics experiments. The network's topological properties such as its average cluster coefficient, diameter, average path length, and node connectivity distribution may point to specific functional groupings or biological pathways of interest.

Here a co-expression network analysis method is presented that offers approaches not provided by other current co-expression analysis tools. First, most high-throughput expression data are not normally distributed; the network models generated here are based on parametric or non-parametric metrics to provide statistically sound network construction. Secondly, most biological networks are known to have approximate scale-free and small-world structure; the construction of the networks generated here is governed by these two properties, guaranteeing biologically meaningful network models. Thirdly, as the properties of scale-free and small-world drive the network construction, user input is minimal, thereby leading to reproducible results. Fourthly, this method can be used to infer a first-pass gene regulatory model when based on temporal expression data. Lastly, the approach is applicable to whole-omics data while running in acceptable compute time.

This method is implemented in R and has been applied to several recent RNA-seq and microarray experiments to identify putative groupings of gene functions and unknown pathways in plant studies.

Reducing Ambiguity in Gene Regulatory Network Inference via Grobner Bases*Brandilyn Stigler*Southern Methodist University; bstigler@smu.edu

Network inference, or reverse engineering, is sensitive to the amount of data used as input. When there are too few data, the ability to construct reliable models is greatly hindered: the number of models that explain the data are too numerous which diminishes their predictability. In the context of systems biology where substantial costs are incurred in laboratory experiments, it is of value to have an estimate of the amount of data required to infer the network. Furthermore, knowledge of specific data which uniquely identify the network is beneficial for minimizing wasted resources.

In this talk, we consider the class of minimal polynomial dynamical systems (PDSs) over a finite field as models of gene networks and establish a connection between model ambiguity and existence of multiple Grobner bases of ideals associated to input data. We provide criteria for determining whether a set of data points uniquely identifies a PDS and a strategy for selecting candidate data points for unique model identification.

Algebraic systems biology and experimental design for the Wnt signaling pathway*Heather A. Harrington*University of Oxford; harrington@maths.ox.ac.uk

In this talk I will focus on a specific signaling pathway of gene regulation and present different approaches for analysing models from systems biology. The focus of these techniques will be a set of models of the Wnt signaling pathway, which is involved in development, adult tissue cells and cancer. We propose a new mechanistic model that includes spatial localization, compare this model and existing models from the literature to data. We apply statistical and algebraic approaches to compare and reject models. We use ideas from chemical reaction network theory, computational algebraic geometry, combinatorics and statistics to study the Wnt signaling system, enabling us to inform design of experiments and reject a competing Wnt model. The techniques are applicable to other problems in systems biology.

Topological RNA structures

Christian Reidys

Virginia Tech; duck@santafe.edu

The folding of coarse grained RNA structures is of interest for a variety of questions. Mike Waterman pioneered the folding of RNA secondary structures producing recursive bijections for constructing longer secondary structures from shorter ones and loop-based polynomial time folding algorithms. It is known that RNA structures exhibit cross-serial interactions and in addition that their folding is in general extremely complex. This calls for suitable filtrations of these pseudoknotted configurations, i.e. identification of classes where polynomial folding is feasible.

Such filtrations come naturally from a fatgraph approach. That means via an enrichment of the contact-graph to a ribbon graph. As such configurations can be viewed as drawings on an (orientable) surface. In this talk I present the ideas behind topological RNA structures. Novel recursions, not in length but topological genus will be shown and the concept of shapes discussed. An outlook towards non orientable fatgraphs will be given, i.e. structures in which the ribbons can be twisted. Finally we discuss briefly a folding algorithm based on this framework.

Posters

Some Combinatorics of the SIR Disease Spreading Dynamics on a Random Graph

Andrei Bura

Virginia Tech; evosavage@yahoo.com

In this presentation we study a simple SIR (susceptible/infected/recovered) process over the random graph $G(n,p)$, in which each edge is selected with independent probability p . A susceptible vertex can be infected via an infected vertex with probability p and is recovered thereafter. Recovered vertices remain recovered. The probability space $G(n,p)$ exhibits a threshold function for the existence of a giant component which can be constructed and analyzed by the Karp-process. The giant component represents the infected entities after the process stops.

We assume the disease starts at a distinguished vertex, S . For fixed p we compute the probability distribution of the random variables $X(d)$ counting the number of infected individuals at distance d from S . These rvs describe the wave of infection as a function of time and are closely related to the paths in $G(n,p)$ connecting S with a vertex at distance d . We present a combinatorial framework which avoids to estimate/compute independent paths using spanning trees.

We finally outline ongoing work to extend this framework to other graphs.

Emergent Node Tree Structures for Multi-scale Modeling of the Dynamics of Interacting Agents

Ruchira Datta

Ohio State University; Ruchira.Datta@gmail.com

We describe ENTs: emergent node tree structures to decompose the dynamics of systems of interacting agents. These decompose the systems into groups; however this notion of grouping is independent of cooperation. Indeed members of a group can be entirely in conflict; more generally, members of a group can have a combination of shared and competing interests. Rather, roughly speaking, players are grouped if the effect of their actions outside the group is of lower dimension than the effects of their actions on one another. This condition is less stringent than subgraph decomposition in graphical games; all players can interact with all other players, including those outside of their group. The reduction in dimensionality, which is due to certain rank conditions and degeneracy conditions among tensors, has the potential to support simplified solution concepts and methods, allowing us to handle the dynamics of large-scale systems. This research is in progress.

A Difference Equation for Tracking Perturbations in Systems of Boolean Nested Canalizing Functions

Elena Dimitrova

Clemson University; edimit@clemson.edu

We develop a difference equation model of the probability that a Boolean network node's value is affected by an initial perturbation over time. The model is analyzed and validated numerically. It is shown that the effect of a perturbation decreases towards zero over time if the Boolean functions are canalizing in sufficiently many variables. The maximum dynamical impact of a perturbation is shown to be comparable to the average impact for a wide range of values of the average sensitivity of the network. Percolation limits are also explored; these are parameter values which generate a transition of the expected perturbation effect to zero as other parameters are varied, so that the initial perturbation does not scale up with the parameters once the percolation limits are reached.

Geometric Combinatorics and RNA Secondary Structure Prediction*Andrew Gainer-Dewar*University of Connecticut; andrew.gainer.dewar@gmail.com

In the last few decades, it has become apparent that the functional interactions of an RNA molecule are governed by its physical structure as well as its sequence of nucleotide "bases". Predicting how a given structure will fold is therefore an important problem in computational biology. The widely-used NNTM approach has a simple combinatorial structure, but relies on thousands of numerical parameters. Numerous authors have used experiments and statistical regressions to determine these parameter values, but at present little is known about how variations in the parameter values effect the model. We apply computational and combinatorial techniques from polyhedral geometry to investigate this issue of parametric sensitivity from a mathematical perspective.

Stratification and Enumeration of Boolean Functions by Canalizing Depth*Qijun He*Clemson University; qhe@g.clemson.edu

Boolean network models have gained popularity in computational systems biology over the last fifteen years. Many of these networks use canalizing Boolean functions, which has led to increased interest in the the study of these functions. The canalizing depth of a function describes how many canalizing variables can be recursively "picked off", until a non-canalizing "core function" remains. In this paper, we show how every Boolean function has a unique algebraic form involving extended monomial layers and a well-defined core function. This generalizes recent work on the algebraic structure of nested canalizing functions, and it yields a stratification of all Boolean functions by their canalizing depth. As a result, we obtain a closed formula for the number of n-variable Boolean functions with depth k, which simultaneously generalizes enumeration formulas for canalizing, and nested canalizing functions.

The Regulatory Network of DNA Mismatch Repair

Claus Kadelka

Virginia Tech; claus89@vt.edu

Although failure of DNA Mismatch Repair (MMR) is associated with microsatellite instability and colorectal cancer, little is known about MMR except for its biochemical pathway. By assembling known regulatory interactions, we build a novel gene regulatory network of MMR, based on a time- and state-discrete modeling framework, which accounts for the cell's inherent stochasticity. This model provides phenotypic predictions for MMR's response to hypoxia and DNA damage. We also substantiate the hypothesis that microRNAs can stabilize network dynamics, thus enhancing genomic stability, by showing that overexpressing microRNAs increases robustness while knocking them out seems to have the opposite effect. In addition to providing a gene regulatory network of MMR, the model yields experimentally verifiable predictions and enables further analysis of the potential stabilizing effect of microRNAs on dynamics of biological networks.

Investigation of Parameter Space for Chaos

Sherli Koshy-Chenthittayil

Clemson University; sherlikoshyc@yahoo.co.in

A popular way of measuring chaos is through the Lyapunov exponents. My research is on using the Lyapunov exponents and the Metropolis-Hastings sampling algorithm to investigate the parameter space of some biological models. The sampling algorithm coupled with a parallel coordinates plot visualization will help us understand the parameters/initial conditions for which we may obtain chaos in a model based on a biological system.

Talks

Combinatorial Interventions for Control Tasks in Large-Scale Signaling Networks

Paola Vera-Licona

UConn Health; veralicona@uchc.edu

Many natural, social and engineered complex systems are vital to human life, society and technology. Undesired behavior of complex systems has produced a great interest in being able to control them. Many diseases, for instance, are the result of abnormal behavior in cellular signaling and regulatory networks. Thus, the design of targeted therapies from a systems biology perspective aims to identify combinations of components to intervene in these networks, to control a pathological behavior to repress a disease while minimizing side effects.

In this talk I will introduce some approaches geared towards the identification of optimal combinations of interventions for large-scale signaling networks, based on the analysis of the network's structure. I will provide a systems medicine example based on the analysis of a signal transduction network built from a cohort of HER2+ breast cancer patients to identify and prioritize combinatorial-targeted treatments. I will conclude with a summary of some of the challenges and future directions to enhance our current approaches.

Discrete Dynamic Modeling Elucidates the Outcomes of Signal Transduction Networks and Helps to Control them

Reka Albert

Pennsylvania State University; rzal@psu.edu

Interaction networks formed by gene products form the basis of cell behavior (growth, survival, apoptosis, movement). Experimental advances in the last decade mark the first steps toward understanding and modeling the properties of the various networks that control the behavior of the cell. Such an understanding would allow the development of comprehensive and effective therapeutic strategies. This talk will present my group's recent work on discrete dynamic modeling of signal transduction networks related to cancer. These models can be developed from qualitative information yet show a dynamic repertoire that can be directly related to the real system's outcomes. For example, our model of the signaling network related to a cell fate change called epithelial to mesenchymal transition indicated effective combinatorial interventions that can block this transition, several of which were then validated experimentally. I will then present a method to integrate network structure and regulatory logic into an expanded network. Identification of specific strongly connected components of this expanded network, called stable motifs, allows an efficient way to determine the dynamic repertoire of the network. We can use the knowledge of these stable motifs to predict manipulations that can stabilize or, conversely, block certain outcomes.

Optimal Control Methods for Stochastic Boolean Networks

David Murrugarra

University of Kentucky; murrugarra@uky.edu

One of the main goals of computational biology after the post-genomic era is to develop optimal control strategies to find efficient medical treatments for changing the state condition of a cell into a new desirable state. The state of a cell is commonly modeled as a Gene Regulatory Network (GRN) that determines the interaction between the different genes involved in a specific cell function. In this context, given a GRN, the possible control actions can be represented as manipulation of edges and nodes of the GRN. Node manipulation requires technology to completely repress or fully activate a particular gene product while edge manipulations only requires a drug that inactivate the interaction between two gene products, which is a realistic control action for medical treatment. The combination of the two possible actions, node and edge manipulations, may produce more effective control strategies for realistic GRNs currently available. This talk will present optimal control algorithms for the identification of the best combination of control actions that are represented by control edges and nodes. Additionally, a discussion about how to use algebraic techniques for the identification of potential intervention targets will be provided.

Multibranch Loop Parameters in RNA Folding Algorithm*Qijun He*Clemson University; qhe@g.clemson.edu

The degree of branching is an important characteristic of an RNA secondary structure, with clear differences between predicted and native structures. The free energy for a multi-branch loop under the NNTM is a simple linear function with parameters for the offset, free base, and helix penalties. Previous work used rooted plane trees as a basic model of RNA folding to analyze trade-offs among the different types of loop structures via a polytope and its normal fan. We extend the parametric analysis to actual RNA sequences using an incremental approach based on the Beneath-and-Beyond algorithm. We also developed an appropriate branching signature and integrated the dynamic programming code.

Algebraic and Discrete Mathematical Biology for Undergraduates*Raina Robeva*Sweet Briar College; robeva@sbc.edu

Over the past 15 years, many advances in biology have utilized algebraic and discrete mathematical approaches. In contrast, the standard basic material at the intersection of mathematics and biology included in the undergraduate curriculum has remained largely unchanged and is still dominated by difference and differential equation models. This troubling difference has been attributed to a lack of appropriate educational resources and to a lack of awareness among the general mathematics faculty of the important role played by algebraic methods in modern biology. The talk will outline some ideas for successfully bridging the gap between what is now common practice in research and the use of “non-calculus” methods for biology at the undergraduate level. Examples will include gene regulatory and signal transduction networks, infectious disease transmission networks, and CpG island identification.

Distribution of Motifs Generated by a Stochastic Grammar for RNA Folding*Svetlana Poznanovikj*Clemson University; spoznan@clemsun.edu

Some of the methods for predicting RNA secondary structures are based on stochastic context-free grammars (SCFGs). We will analyze one of the most notable SCFGs which is used in the prediction program Pfold. In particular, we will show that the distribution of base pairs, helices and various types of loops in RNA secondary structures generated by this SCFG is asymptotically Gaussian, for a generic choice of the grammar probabilities. This analysis reveals that some interesting properties of the distribution are invariant under parameter change. We will also discuss how these properties are affected by slight modifications of the grammar.

Reduction of Combinatoric Calcium Release Site Models via Fast/Slow Analysis and Genetic Algorithm*Yan Hao*Hobart and William Smith Colleges; hao@hws.edu

Mathematical models of calcium release sites derived from Markov chain models of intracellular calcium channels exhibit collective gating reminiscent of the experimentally observed phenomenon of calcium puffs and sparks. Such models often take the form of stochastic automata networks in which the transition probabilities of each channel depend on the local calcium concentration and thus the state of the other channels. These compositionally defined calcium release site models have been proven to be superior to traditional differential equations models, yet they suffer combinatoric state-space explosion thus are computationally expensive to be applied to whole cell models. In order to overcome this problem, we have implemented several automated procedures for model reduction using fast/slow analysis. After categorizing rate constants in the single channel model as either fast or slow, groups of states in the expanded release site model that are connected by fast transitions are lumped, and transition rates between reduced states are chosen consistent with the conditional probability distribution among states within each group. In the case where the time scale difference is absent, genetic algorithm is applied to select the most efficient grouping of the states. For small problems these conditional probability distributions can be numerically calculated from the full model without approximation. For large problems the conditional probability distributions can be approximated without the construction of the full model by assuming rapid mixing of states connected by fast transitions. Alternatively, iterative aggregation/disaggregation may be employed to obtain reduced calcium release site models in a memory-efficient fashion. Benchmarking of several different iterative aggregation/disaggregation-based fast/slow reduction schemes establishes the effectiveness of automated calcium release site reduction utilizing the Koury-McAllister-Stewart method.

Discrete Homotopy and Homology Theory for Metric Spaces*Helene Barcelo*Mathematical Sciences Research Institute; hbarcelo@msri.org

Discrete homotopy theory is a discrete analogue of homotopy theory, associating a bigraded sequence of groups to a simplicial complex, capturing its combinatorial structure, rather than its topological structure. Discrete homotopy can be equivalently defined for finite connected graphs, resulting in an algebraic invariant of finite connected graphs and graphs homomorphisms, in the same way that classical homotopy theory gives invariants of topological spaces and continuous maps: it associates a sequence of groups to a finite connected graph. The notion of discrete homotopy can also be generalized to arbitrary metric spaces. One can also define a corresponding notion of discrete homology in the same way that classical homology is related to classical homotopy theory. We will review these notions and discuss several applications.

Talks

Rational Design of Antibiotic Treatment Plans

Bernd Sturmfels

University of California, Berkeley; bernd@berkeley.edu

We present work with Portia Mira, Kristina Crona, Devin Greene, Juan Meza and Miriam Barlow, aimed at developing antibiotic treatment plans that can reverse the evolution of antibiotic resistance. The Barlow lab at UC Merced generated adaptive landscapes for 16 genotypes of the TEM beta-lactamase that vary from the wild type genotype TEM-1 through all combinations of four amino acid substitutions, and determined the growth rate of each genotype when treated with each of 15 beta-lactam antibiotics. Using growth rates for fitness in two models from population genetics, we computed the probability of each amino acid substitution in each beta-lactam treatment, and we searched through the 15 treatments for substitution paths leading from each of the 16 genotypes back to TEM-1. We identified treatment paths with the highest probabilities of returning TEM to the wild type state, thus offering promise for reversing the evolution of resistance to antibiotics. This lecture highlights the mathematics in this project.

Airway Inflammation: Does it Add Up?

Shernita Lee

University of North Carolina Chapel Hill; shernita@email.unc.edu

In this talk, two projects are discussed which focus on inflammation in the airway epithelia: one to a fungal pathogen and the other from genetic mutations such as with cystic fibrosis (CF). The lung epithelium is responsible for pathogen recognition, ion transport, and cell recruitment. Here, a computational model of the host-fungal interaction is discussed along with supporting experimental validations. This model addresses changes in the immune response and downstream effects on iron metabolism. Additional work in the Tarran Lab at UNC Chapel Hill is also presented, which focuses on CF, an autosomal recessive disease, which minimizes lung functionality due to mucus obstruction, causing a reduction in survival and the overall quality of life. The goal of this research is use a computational model to capture cellular changes in CF epithelial cells. This work provides a global picture of the scientific knowledge regarding epithelial changes in CF and non-CF individuals, and discusses the relevant proteins for future research applications.

Combinatorial Optimization Methods for Agent-Based Modeling*Matthew Oremland*Ohio State University; oremland.2@osu.edu

Agent-based models (ABMs) are computer simulations in which discrete entities (agents) interact with each other and their local environment according to a set of rules. Localized interactions give rise to global dynamics that can be otherwise difficult to predict. ABMs provide a natural *in silico* laboratory in which experiments can be conducted; in particular, it is natural to pose optimization problems that can be investigated via simulation. While the rule-based framework is convenient for prescribing a variety of behaviors and for incorporating heterogeneity, rigorous methods for solving optimization problems for ABMs are not well-developed.

In this work, examples are given of two representative ABMs in biological contexts. Optimization problems are posed for each; in both cases, the solution of the problem depends on the determination of an optimal sequence of discrete control inputs to the model. In order to accomplish this, relevant ABM dynamics are approximated via a system of time-discrete difference equations. Heuristic combinatorial optimization methods are applied to the equation models, and a statistical metric for determining the accuracy of the equations with respect to ABM dynamics is presented. Results obtained are implemented back into the corresponding ABMs as validation of the approach.

In both cases, results from combinatorial optimization line up very well with the ABM dynamics; these results are obtained using systems of equations that are more rigorously formulated and much more computationally efficient than direct application of simulation optimization techniques. Additionally, the use of Pareto optimization provides a suite of solutions without need of a cost function. The framework presented here has applicability to a variety of biological models, particularly for those interested in solving discrete optimization problems.

Network Architectural Conditions for Noise-induced Biochemical Oscillators*Jaewook Joo*University of Tennessee; jjoo1@utk.edu

Biological rhythm is one of ubiquitous phenomena in nature. In biological signaling or regulatory networks, it is of great interest to understand the mechanisms governing the oscillatory behaviors and their biological implications, and elucidate the controllable conditions. When the systems are under the influence of stochastic fluctuations, the dynamical behaviors are highly unpredictable and uncontrollable. The noise can confer a new dynamical capability or destroy the existing dynamical properties of the systems. Here we investigate such a case where stochastic fluctuations can give rise to the new capability of noise-induced oscillation in a subset of biochemical reaction networks, the networks with only three biochemical species whose reactions are governed by mass action kinetics. We model the networks with master equations and approximate them by using linear noise approximation. First, we provide a proof that the existence of feedback loops in biochemical reaction networks is a necessary topological condition for noise-induced oscillation. Second, we prove that the biochemical reaction networks with two/three nodes and with a single negative feedback loop are indeed capable to admit a peaked power spectrum with an arbitrarily large signal-to-noise ratio value, an indicator of amplified and coherent stochastic oscillation. Third, sampling the values of model parameters from a biologically feasible range and calculating the average signal to noise ratio of each biochemical reaction network, we classify the ensemble of sixty-three biochemical reaction networks into three groups of noise-induced oscillation performance. From the common network architecture among the networks belonging to the same performance group, we learn that the networks equipped with the coupling of negative and positive feedback loops have the better noise-induced oscillation performance than the networks with only negative feedback loops. The performance of networks also depends on the relative size of the positive and negative feedback loops; the networks with the bigger positive and smaller negative feedbacks are much worse oscillators than the networks with only negative feedback loops. The results of our work elucidate the network architectural conditions for highly amplified and coherent noise-induced oscillations and can be used for controlling signaling pathways in natural organisms and/or for better designing new synthetic circuits with oscillatory functions.

SUPPORTED BY

UConn
HEALTH

