

# SAMSI 2008-09 Program on Algebraic Methods in Systems Biology and Statistics

## Final

Reinhard Laubenbacher  
Chair, Program Leaders  
On behalf of the Working Groups Leaders

May 2010

## 1 Program Overview

In recent years, methods from algebra, algebraic geometry, and discrete mathematics have found new and unexpected applications in systems biology as well as in statistics, leading to the emerging new fields of “algebraic biology” and “algebraic statistics.” Furthermore, there are emerging applications of algebraic statistics to problems in biology. This year-long program provided a focus for the further development and maturation of these two areas of research as well as their interconnections. The unifying theme is provided by the common mathematical tool set as well as the increasingly close interaction between biology and statistics. The program allowed researchers working in algebra, algebraic geometry, discrete mathematics, and mathematical logic to interact with statisticians and biologists and make fundamental advances in the development and application of algebraic methods to systems biology and statistics. The essential involvement of biologists and statisticians in the program provided the applied focus and a sounding board for theoretical research.

### 1.1 Research Foci

**Systems Biology:** The development of revolutionary new technologies for high-throughput data generation in molecular biology in the last decades has made it possible for the first time to obtain a system-level view of the molecular networks that govern cellular and organismal function. Whole genome sequencing is now commonplace, gene transcription can be observed at the system level and large-scale protein and metabolite measurements are maturing into a quantitative methodology. The field of systems biology has evolved to take advantage of this new type of data for the construction of large-scale mathematical models. System-level approaches to biochemical network analysis and modeling promise to have a major impact on biomedicine, in particular drug discovery.

**Statistics:** It has long been recognized that the geometry of the parameter spaces of statistical models determines in fundamental ways the behavior of procedures for statistical inference. This connection has in particular been the object of study in the field of information geometry, where differential geometric techniques are applied to obtain an improved understanding of inference procedures in smooth models. Many statistical models, however, have parameter spaces that are not smooth but have singularities. Typical examples include hidden variables models such as the phylogenetic tree models and the hidden Markov models that are ubiquitous in the analysis of biological

data. Algebraic geometry provides the necessary mathematical tools to study non-smooth models and is likely to be an influential ingredient in a general statistical theory for non-smooth models.

## Algebraic methods

*Algebraic biology* is emerging as a new approach to modeling and analysis of biological systems using tools from algebra, algebraic geometry, discrete mathematics, and mathematical logic. Application areas cover a wide range of molecular biology, from the analysis of DNA and protein sequence data to the study of secondary RNA structures, assembly of viruses, modeling of cellular biochemical networks, and algebraic model checking for metabolic networks, to name a few.

*Algebraic statistics* is a new field, less than a decade old, whose precise scope is still emerging. The term itself was coined by Giovanni Pistone, Eva Riccomagno and Henry Wynn. Their book explains how polynomial algebra arises in problems from experimental design and discrete probability, and it demonstrates how computational algebra techniques can be applied to statistics. The first of these applications have focused on categorical data and include the study of Markov bases and conditional inference, disclosure limitation, and parametric inference, to name a few.

The central idea underlying algebraic statistics is that the parameter spaces of many statistical models are (semi-)algebraic sets. The geometry of such possibly non-smooth sets can be studied using tools from algebraic geometry. Many problems in computational biology can be described within this framework. This is where algebraic statistics joins algebraic biology as a new methodology for solving problems in systems biology.

The unifying theme of the program was the development and use of a particular set of tools from algebra, algebraic geometry, and discrete mathematics to solve problems in statistics and biology.

## 1.2 Organization and Program Leadership

**Organizing Committee:** Peter Beerli (School of Computational Sciences and Department of Biological Sciences, Florida State University), Andreas Dress (Director, CAS-MPG Partner Institute for Computational Biology, Shanghai), Mathias Drton (Department of Statistics, University of Chicago), Ina Hoeschele (Department of Statistics, Virginia Tech, and Virginia Bioinformatics Institute), Christine Heitsch (School of Mathematics, Georgia Tech), Serkan Hosten (Department of Mathematics, San Francisco State University), Reinhard Laubenbacher, Committee Chair (Department of Mathematics, Virginia Tech, and Virginia Bioinformatics Institute), Bud Mishra (Departments of Computer Science, Mathematics, and Cell Biology, Courant Institute, NYU), Don Richards (Department of Statistics, Pennsylvania State University), Seth Sullivant (Department of Mathematics, NCSU), Brett Tyler (Department of Plant Pathology and Weed Science, Virginia Tech, and Virginia Bioinformatics Institute), Ruriko Yoshida (Department of Statistics, University of Kentucky).

## 1.3 Major Participants

**Long-Term Visitors:** Edward Allen (Wake Forest University), Elizabeth Allman (University of Alaska), James Degnan (University of Canterbury), Alicia Dickenstein (University of Buenos Aires), Luis Garcia-Puente (Sam Houston State University), Jeremy Gunawardena (Harvard Medical School), Chris Hillar (MSRI), Serkan Hoşten (San Francisco State University), Reinhard Laubenbacher (VA Tech), Catherine Mathias (Universite d'Evry), Uwe Nagel (University of Kentucky), Edwin O'Shea (UNAM, Mexico), Sonja Petrović (University of Illinois, Chicago), Giovanni Pistone (Politecnico di Torino), John Rhodes (University of Alaska), Eva Riccomagno (University of Genoa), Anne Shiu (UC Berkeley)

**Postdoctoral Fellows:** Megan Owen (Cornell University), Ahmad Saeid Yasamin (Indiana University)

**Graduate Students:** Wenjie Chen (UNC), Julia Chifman (University of Kentucky), Deidra Coleman (NCSU), Thomas Friedrich (FU- Berlin), Benjamin Wells (NCSU), Jason Yellick (NCSU)

**Faculty Releases:** Ian Dinwoodie (Duke), Scott Provan (UNC), Eric Stone (NCSU), Seth Sullivant (NCSU), Yung-Jing Tzeng (NCSU)

## 2 Description of Activities

### 2.1 Workshops

The SAMSI program on Algebraic Methods in Systems Biology and Statistics has been bolstered by a number of workshops and special sessions held throughout the year both at SAMSI and nearby locations.

#### 2.1.1 Opening Workshop:

The Kickoff Workshop and Tutorial was held September 14–17, 2008. The principal goal of the workshop was to engage a broadly representative segment of the mathematical, statistical, and life sciences communities to determine research directions to be pursued by working groups during the program. Four working groups were formed that eventually merged down to three working groups.

The workshop covered a very broad range of topics in the interactions of algebraic methods with systems biology and statistics. After introductory tutorials on Sunday, more focused talks directed towards the problem areas to be highlighted throughout the year were given. Highlighted topics included: algebraic statistical models, combinatorics of biological molecules, automata theory and finite dynamical systems, phylogenetics, causal models, and random graph models.

The workshop also contained a number of discussion sessions with the goal of identifying research areas that would be highlighted throughout the yearlong program. After these intended working group target areas were identified, the program members formed breakout sessions to begin discussions of topics that would be discussed in the working groups throughout the year. The particular areas of the working groups are described below.

The tutorial speakers were: Bernd Sturmfels, Reinhard Laubenbacher, and Elizabeth Allman. The keynote speakers were: Mathias Drton, Jeremy Gunawardena, Christine Heitsch, Bud Mishra, Abdul Jarrah, Chris Schardl, Michael Savageau, Gheorghe Craciun, Sumio Watanabe, Brandilyn Stigler, Meera Sitharam, Lior Pachter, Brett Tyler, Niko Beerenwinkel, Eva Riccomagno, and Steve Fienberg.

#### 2.1.2 Discrete Models in Systems Biology Workshop

The discrete models workshop was held December 3-5, 2008 and was organized by Elena Dimitrova (Clemson University), Ilya Shmulevich (Institute for Systems Biology), and Brandilyn Stigler (Southern Methodist University). The workshop focused on the use of discrete models in systems biology. Discrete modeling approaches have been applied to a wide variety of biological contexts, including gene regulatory networks, epidemiology, and ecosystem dynamics. Examples of topics of interest in the workshop were

1. discrete dynamical systems: multi-state models such as Boolean networks, logical models, and finite dynamical systems; random networks; and analytic tools including statistical-mechanical approaches

2. Bayesian networks, including dynamic Bayesian networks and graphical models
3. static networks: interaction networks and graph-theoretic approaches
4. simulation: finite-state machines, agent-based networks, process algebras

A chief goal was to stimulate the organization of working groups focused on addressing key challenges in discrete modeling in computational systems biology, particularly the establishment of unifying themes and principles.

As part of our commitment to foster a synergistic community, we organized three question-and-discussion sessions to encourage interactions among workshop participants and two poster sessions to showcase the work of junior researchers, including postdocs and graduate students.

### **2.1.3 Algebraic Statistical Models Workshop**

Many classical statistical models, in particular Gaussian models from multivariate statistics and models for discrete random variables, exhibit algebraic structure in their parameter spaces. This workshop focused on both algebraic and statistical aspects of such algebraic statistical models. It was intended to complement other mid-program workshops, which focused more on particular application areas.

The workshop was held at SAMSI, January 15–17, 2009 and featured topics by working group participants, as well as outside experts, whose opinions helped to provide new research directions in the program. The organizers were Mathias Drton, Eva Riccomagno, and Seth Sullivant. Focus topics of the workshop included Markov bases, graphical models, algebraic tools for maximum likelihood estimation, identifiability problems, and cumulant methods.

The workshop included talks by Steffen Lauritzen, Thomas Richardson, Donald Richards, Ruriko Yoshida, Elizabeth Allman, Sonja Petrović, Akimichi Takemura, Serkan Hoşten, Hugo Maruri, and Jason Morton and a poster session.

### **2.1.4 Miniworkshop on Systems Biology**

A subgroup of the systems biology working group was formed to focus on software development for parameter estimation for discrete models. As part of the subgroup activities a miniworkshop at SAMSI was conducted February 24-26, 2009. Participants included E. Dimitrova, L. Garcia, F. Hinkelmann, A. Jarrah, R. Laubenbacher, B. Stigler, and P. Vera-Licona. The workshop was focused on the design of the overall architecture of a software package for parameter estimation and simulation for Boolean network models. A significant part of the time was devoted to actual code development.

### **2.1.5 Evolutionary Biology and Phylogenetics Workshop**

Recently there has been a marked synergy between modern biology and higher mathematics. A number of important connections have been established between computational biology and the emerging field of “algebraic statistics,” which combines combinatorics, computational algebra, polyhedral geometry and statistical modeling. The primary objective of this workshop was to bring together new and established researchers in mathematics, biology, and statistics in order to discuss the crossover between algebraic statistics, molecular evolution, and phylogenetics.

As part of our commitment to foster a synergistic community, we organized several discussion sessions to encourage interactions among workshop participants to actively begin new collaborations, discuss new research directions, and make new connections. For example, we discussed phylogenetic

invariants on group-based models, such as Jukes-Cantor model and their applications to tree reconstruction as well as discussion on the phylogenetic mixture models. There were several discussion about coalescent theory, such as coalescent approach to approximate the distribution of gene trees. Also we discussed inferences about the impact of phenotype on genotype from the ancestral lineage and also some reviews of phylogenetic reconstructions, what are known and unknown.

We invited four researchers to give a one-hour keynote address, and six researchers to give contributed (invited) talks of 45 minutes length (those include 5 minutes for questions at the end) at the workshop. Invited speakers were: Jeff Thorne, Cecile Ané, Junhyong Kim, Tandy Warnow, Seth Sullivant, Eric Stone, Fumei Lam, Laura Kubatko, Jeremy Sumner, Sonja Petrović, and Jesus Fernandez-Sanchez.

### **2.1.6 AMS Special Session: Mathematics of Biochemical Reaction Networks**

The special session "Mathematics of Biochemical Reaction Networks" was held during the Southeastern Section meeting of the AMS at North Carolina State University during the weekend of April 4-5. The organizers were Gheorghe Craciun, Manoj Gopalkrishnan, and Anne Shiu. The idea to organize this workshop in proximity to SAMSI, and close to the evolutionary biology workshop was formed during the SAMSI opening workshop.

Our intent was to bring together individuals who study biochemical reaction networks, in order to share ideas that range from those building upon the classical Feinberg-Horn-Jackson deficiency theory, to those more recent algebraic techniques that highlight the rich algebraic structure inherent in these networks. For example, much work has focused on predicting dynamics and resolving questions of stability simply from the topological structure of the underlying reaction network. Another topic to be covered is the class of "monotone" systems. In particular, this session featured reports on collaborations that grew out of activities during the 2008-9 academic year at SAMSI in North Carolina and the Mathematical Biosciences Institute in Ohio.

Special 45-minute talks were given by Martin Feinberg, Jeremy Gunawardena, and Eduardo Sontag. Several talks in particular had a strong algebraic aspect. For example, Greg Rempala talked about an algebraic statistical model for inferring biochemical reaction networks. Ezra Miller discussed results on binomial primary decomposition and its connection to boundary steady states of a chemical reaction system. Luis Garcia connected Birch's Theorem and chemical reaction network theory to Bézier patches and their generalizations. Alicia Dickenstein shared results that compare detailed balancing to complex balancing in terms of their associated algebraic varieties. In addition, there were speakers from various backgrounds ranging from theoretical computer science to control theory to probability, and talks whose titles include the words "homotopy" and "number theory." This session led to fruitful discussions and further collaborations.

### **2.1.7 Transition Workshop**

The transition workshop was designed to synthesize the year's activities and provide a blueprint to go forward with research in this field. Talks covered a range of topics related to the working group topics. Talks by Reinhard Laubenbacher, Seth Sullivant, and Ruriko Yoshida indicated how we want to move forward beyond the program year. Some talks highlighted successes from the working groups, and some talks were chosen to increase the range of topics present in the program.

In addition to the talks, there was ample time for participants to discuss future plans for research, and for the research area. Each day included multiple "Second Chances" sessions, which gave participants a chance to further engage the speakers and to explore relations between talks, and future directions. One outgrowth of these discussions is that Laubenbacher, Sullivant, and Yoshida will

coorganize a workshop on Algebraic Methods during the MBI 2011-12 program year on “Stochastics in Biological Systems”.

Invited speakers were: Reinhard Laubenbacher, Heike Siebert, Luis Garcia-Puente, Katherine St. John, Megan Owen, Marcy Uyenoyama, Henry Wynn, Ruriko Yoshida, Peter Huggins, Giles Gnacadja, Anne Shiu, Olgica Milenkovic, Elena Dimitrova, Paul Kidwell, and Seth Sullivant.

### **2.1.8 Joint Statistics Meeting Mini-symposium on Algebraic Methods in Systems Biology and Statistics**

Ian Dinwoodie organized a mini-symposium, at the Joint Statistics Meeting in Washington D.C. in August 2009, that will highlight some of the results that have come out of the SAMSI program. Invited speakers were: Saeid Yasamin, Simon Lunagomez, Seth Sullivant, and Reinhard Laubenbacher.

## **2.2 Working Groups**

At the end of the opening workshop in September, the afternoon was devoted to the formation of working groups for the year. Based on participant interest and program themes, the topics that emerged were:

1. Systems biology: The relationship between the structure and dynamics of biological networks;
2. Algebraic statistics and experimental design;
3. Evolutionary biology and phylogenetics.

There was significant overlap between topics and membership of the different working groups. For instance, experimental design is a very important topic in systems biology as well.

### **2.2.1 Relationship between structure and dynamics of biological networks**

The working group leaders are R. Laubenbacher (VT) and Brandilyn Stigler (SMU).

One of the dominant themes at the opening workshop was the relationship between the structure of biological networks and the kind of dynamics this structure supports. The questions about this relationship is appropriate for a variety of biological networks, ranging from molecular pathways to social networks that support the spread of epidemics. For the working group, the focus was entirely on biochemical networks, encompassing two different modeling frameworks: polynomial dynamical systems over finite fields, in particular Boolean networks, and systems of polynomial differential equations. The primary structure of a network is given by a directed graph that indicates the dependence of the network variables on each other. In both modeling frameworks, one of the goals is to infer constraints on the dynamics of the network from properties of this graph. In the other direction, the goal is to infer the structure of the graph from a partial specification of the network dynamics, e.g., through a collection of time course experiments.

### **Working Group Activities.**

Since the background of the working group members varies considerably, a primary focus of the group activities in the fall and part of the spring were on presentations and discussions aimed at establishing a common background in systems biology and the different approaches to modeling and simulation. The group is now at a stage where first results are being presented, beginning with the work of the subgroup on software development, described below.

A major problem with the construction of large-scale algebraic models is that there are no sophisticated tools available, comparable to ODE tools. Most importantly, the tool of fitting ODE model parameters to available data is key in continuous model construction, but completely absent for algebraic models. Also, tools like bifurcation analysis, sensitivity analysis, stability analysis are all unavailable to the algebraic modeler. However, there exist several or all of these tools in the polynomial dynamical systems framework. The goal of this subgroup of the working group was to collect together available software and integrate it in a coherent package scheduled which was released in April 2009.

#### **Active participants.**

Carsten Conradi (MPI Magdeburg), Alicia Dickenstein (University of Buenos Aires), Elena Dimitrova (Clemson University), Ian Dinwoodie (Duke University), Lee Falin (Virginia Tech), Thomas Friedrich (TU Berlin), Gilles Gnacadja (Amgen), Richard Haney (Cellular Statistics), Franziska Hinkelmann (Virginia Tech), Serkan Hosten (San Francisco State University), Abdul Salam Jarrah (Virginia Tech), Reinhard Laubenbacher (Virginia Tech), Tong Lee (Virginia Tech), Shaowei Lin (Univ. of California-Berkeley), Megan Owen (SAMSI), Mercedes Soledad Perez Millan (University of Buenos Aires), Anne Shiu (Univ. of California-Berkeley), Heike Siebert (FU Berlin), Brandy Stigler (Southern Methodist University), Seth Sullivant (NCSU), Jung-Ying Tzeng (N.C. State University), Alan Veliz-Cuba (Virginia Tech), Benjamin Wells (N.C. State University), Henry Wynn (London School of Economics), Richard Yamada (University of Michigan), Shantia Yarahmadian (Indiana University), Saeid Yasamin (SAMSI),

### **2.2.2 Algebraic Statistics and Experimental Design**

The Algebraic Statistics and Experimental Design (ASED) working group in the 2008-2009 SAMSI program Algebraic Methods in Systems Biology and Statistics had approximately 30 members, including many remote participants in England, Italy, Japan, and throughout the U.S. The group was formed during the opening workshop in September, when all members were present and decided upon research themes and meeting times. The working group leaders were Serkan Hosten of San Francisco State University and Ian Dinwoodie of Duke University.

The ASED group worked on a range of statistical applications that use computational tools of commutative algebra. These applications include sampling and Monte Carlo methods for discrete data (tables and sequences), experiments (data gathering) and data analysis for reverse engineering of biological networks, disclosure limitation, and foundations of phylogenetic trees in evolutionary and population biology. Each application area emphasizes certain algebraic tools and each has roots in particular research groups with a wide international base. The different mathematical tools and research groups were well-represented in the group members. The working groups benefited from the active and steady participation of individuals with much depth and experience. In particular, Henry Wynn and Giovanni Pistone shepherded the experimental design part of ASED, and worked on collaborations with the Systems Biology working group, where design issues are important for finding wiring diagrams and network connections. Phylogenetic trees were supported by John Rhodes, Elizabeth Allman, and Seth Sullivant, who completed foundational work on identifiable tree models in the course of the year, and shared their work from the Evolutionary Biology working group. Ongoing work in high-dimensional tables (sampling and disclosure limitation) was well-represented as well, with presentations by researchers and practitioners Akimichi Takemura, Larry Cox, Edwin O'Shea, Adrian Dobra, and others. Also some connections were made with the Sequential Importance Sampling program through research on sequential Monte Carlo methods for statistical inference on Boolean dynamics in biological networks.

The ASED working group met formally on Mondays at noon, in addition to informal collabo-

rations. About half the participants logged-in from remote locations using the Webex networking application. The list of talks and speakers is at [www.samsi.info](http://www.samsi.info) under the ASED working group link, together with supporting materials and documents.

The complete list of ASED working group members is below:

Elizabeth Allman (University of Alaska-Fairbanks), Deidra Coleman (N.C. State University), Lawrence H. Cox (CDC), Elena Dimitrova (Clemson University), Ian Dinwoodie (Duke University), Luis David Garcia-Puente (Sam Houston State University), Hisayuki Hara (Tokyo), Serkan Hosten (San Francisco State University), Thomas Kahle (Leipzig), Imre Risi Kondor (University College London), Reinhard Laubenbacher (Virginia Bioinformatics Institute), Tong Lee (Virginia Tech), Hugo Maruri-Aguilar (London School of Economics), Catherine Matias (CNRS), Uwe Nagel (University of Kentucky), Edwin O'Shea (Avanzados del IPN), Vittorio Perduca, Mercedes Soledad Perez Millan (Universidad de Buenos Aires), Sonja Petrovic (University of Illinois at Chicago), Giovanni Pistone (Torino), Eva Riccomagno (Genoa), Seth Sullivant (NCSSU), Akimichi Takemura (Tokyo), Caroline Uhler (Univ. of California-Berkeley), Alan Veliz-Cuba (Virginia Tech), Benjamin Wells (N.C. State University), Henry Wynn (London School of Economics), Richard Yamada (University of Michigan), Ahmad S. Yasamin (SAMSI), Jason Yellick (N.C. State University), Ryo Yoshida (Japan), Yi Ming Zou (University of Wisconsin-Milwaukee), Or Zuk (M.I.T.), Piotr Zwiernik (University of Warwick)

### **2.2.3 Evolutionary Biology and Phylogenetics**

As part of SAMSI's 2008-09 program on Algebraic Methods in Systems Biology and Statistics a working group in 'Evolutionary Biology' was formed during the opening workshop in September. Broadly speaking, members of this group were interested in finding, understanding, and solving problems arising in evolutionary biology that might require sophisticated mathematical and statistical techniques that have yet to be developed.

The group was lead by Seth Sullivant, Elizabeth Allman, and John Rhodes, and during the opening workshop interested participants indicated that primary areas of common interest included phylogenetics, coalescent theory, population genetics, and comparative genomics.

#### **Working Group Activities.**

It was immediately clear that group members had widely diverse backgrounds in statistics, mathematics, and biology, and that participants needed a 'common language' and 'common background knowledge' in order to collaborate. During the first semester and spilling over into the beginning of the second semester, the working group met weekly. Each week a particular group member with expertise in one of the areas of common interest, gave a talk at an introductory level to familiarize other group members with the area, discuss his/her research, and suggest possible problems where algebraic techniques might yield results. Typically, after an hour or more of introduction by the speaker, group members discussed the problems and a question-and-answer period began.

The main topics included: the structure of tree space for phylogenetic trees (2 sessions), mixture models in phylogenetics and invariants (3 sessions), the coalescent model (4 sessions), geometry of cophylogeny (1 session), and comparative genomics (1 session).

- 9/23/08 Megan Owen on the space of phylogenetic trees and the geodesic distance: Space of Phylogenetic Trees
- 9/30/08 John Rhodes on phylogenetic invariants
- 10/07/08 Seth Sullivant: Some algebraic ideas for phylogenetic mixtures
- 10/14/08 Peter Beerli (Part 1 of introduction to coalescent theory series): Population genetic calculations that do not fit on the back of an envelope
- 10/21/08 Laura Salter Kubatko (Part 2 of introduction to coalescent theory series):
- 10/28/08 Peter Beerli (Part 3 of the introduction to coalescent theory series): Finding good trees - Simplifying Coalescent trees
- 11/04/08 Serkan Hosten: Extended UPGMA and phylogenetic tree reconstruction
- 11/11/08 Rudy Yoshida: Open Problems in Geometry of Cophylogeny
- 11/18/08 Julia Chifman: Group-based models
- 01/19/09 James Degnan: Gene tree distributions and coalescent histories
- 01/26/09 Or Zuk: Annotating the Human Genome Using Comparative Genomics

For a particularly successful ending to the fall semester, the evolutionary biology meeting consisted of a session in which individuals suggested open problems for the group to work on.

After an organizational meeting in mid-January, the working group decided to focus primarily on reading papers to acquire a deeper understanding of tree space, gene-tree/species-tree problems (coalescent theory), models of speciation, and ancestral recombination graphs. Several talks were also scheduled during the semester while researchers were visiting at SAMSI for collaborations and workshops.

Weekly working group meetings ran differently the Spring term. The idea was to have all group members read the papers for the week, and one person was assigned to lead a discussion. It was assumed that no one was an expert in the the area, so that the group could learn by reading up on a topic together. This worked reasonably well, but the best discussions took place when there were more group members physically present at SAMSI.

### **Group Membership.**

The official number of group members was quite high, around 35, though the number of active participants is closer to 20. The number of participants on a weekly basis (“the faithful”) was typically about eight to ten. The meetings on the coalescent model were particularly well-attended.

The names of the active participants have been included below:

Elizabeth Allman (University of Alaska-Fairbanks), Elisaveta Arnaudova (University of Kentucky), Peter Beerli (Florida State University), Julia Chifman (University of Kentucky), Luis Garcia-Puente (Sam Houston State University), Serkan Hosten (San Francisco State University), Laura Kubatko (Ohio State University), Jinze Liu (University of Kentucky), Catherine Matias (CNRS, Laboratoire Statistique et Genome), Uwe Nagel (University of Kentucky), Megan Owen (SAMSI), Sonja Petrovic (University of Illinois at Chicago) Scott Provan (University of North Carolina), John Rhodes (University of Alaska-Fairbanks), Chris Schardl (University of Kentucky), Seth Sullivant (N.C. State University), Amelia Taylor (Colorado College), Jason Yellick (N.C. State University), Ruriko Yoshida (University of Kentucky), Or Zuk (M.I.T. Broad Institute) Piotr Zwiernik (University of Warwick)

Several new collaborations were formed among evolutionary biology group members, and the working group format gave the opportunity for extended research interaction to both these new and pre-existing collaborations.

## 2.3 University Courses

**Title:** Algebraic Methods in Systems Biology and Statistics

**Instructors:** Seth Sullivant (NCSU) and Reinhard Laubenbacher (VA Tech)

**Course Day and Time:** Tuesday 4:30-7:00

**Course Description:** This course provided an introduction to the algebraic techniques that have emerged as useful tools in biology and statistics. This course was intended to bridge the gap between abstract algebra and the application areas covered in the year-long program.

After providing an introduction to polynomial rings, ideals, and Gröbner bases, we will survey a range of applications of these ideas. Possible topics include: Polynomial dynamical systems over finite fields and applications, graphical and hierarchical models, Markov bases for contingency table analysis, phylogenetic models and the space of trees, applications of tropical geometry in MAP estimation.