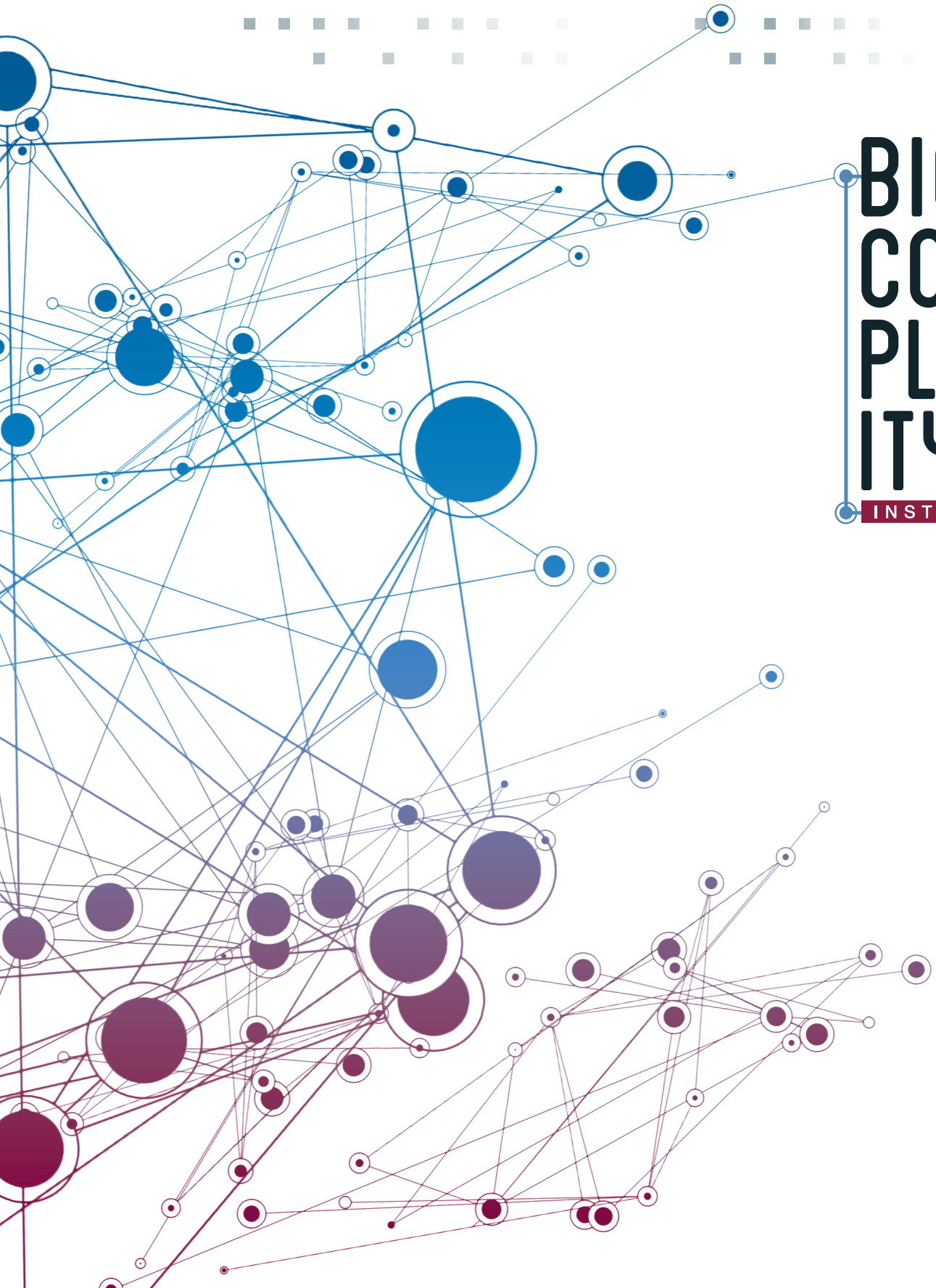

BIO COMPLEXITY INSTITUTE

RESEARCH SYMPOSIUM | 2017

Wednesday, November 8th

VIRGINIA TECH™

Research Symposium | 2017-11-08 | Abstracts

MOLECULAR INSIGHTS INTO CIRCADIAN REGULATION OF P53 TUMOR SUPPRESSOR ACTIVITY

Sarah Jachim, Kelsey O'Hern, Philip Stauffer, Carla V. Finkielstein

Mutations in the tumor suppressor p53 were identified in up to 80% of all cancer cases. In response to genotoxic stress, such as UV and ionizing radiation, p53-mediated transcription is responsible for either cell cycle arrest or apoptosis in cells with damaged DNA. The activity of p53 is negatively regulated by the oncogenic protein MDM2, an E3 ligase which ubiquitinates the C-terminus of p53 and targets it for proteasomal degradation.

Our laboratory has recently found the circadian protein Period 2 (PER2) directly binds p53, prevents MDM2-mediated ubiquitination, and favors p53 stability in unstressed cells. Furthermore, dissociation of PER2 from p53 is a pre-requisite for p53 oligomerization and transcriptional activity in response to genotoxic stress.

This project involves investigating the biochemical consequences of mutations (i.e., missense, nonsense, frame-shift, and in-frame deletions) within the PER2:p53 binding interface. Our work aims to identify critical hot-spot mutations which contribute to the instability of the PER2 and p53 complex, and favor an abnormal p53 transcriptional response. We expect findings from our work will contribute to our understanding of how signals are integrated to control the cellular response top stress.

Research Symposium | 2017-11-08 | Abstracts

GLOBAL PERTURBATIONS OF ANAPHASE CHROMOSOME MOVEMENT CAUSED BY A SINGLE MIS-ATTACHED KINETOCHORE

Alyssa Osimani, Bin He, Heather Bomberger, Gul Civelekoglu-Scholey, Daniela Cimini

Incorrect chromosome segregation leads to an abnormal number of chromosomes, or aneuploidy, a condition that is a leading cause of miscarriage in humans and a ubiquitous feature of cancer cells. Correct chromosome segregation during mitosis relies on correct kinetochore-microtubule attachment by which two sister kinetochores bind microtubules from opposite poles of the mitotic spindle.

A common error in mitosis is the attachment of single kinetochores to microtubules from both spindle poles instead of just one (or merotelic kinetochore attachment). This can cause the mis-attached chromosome to lag behind in anaphase, when all other chromosomes move to the poles. Preliminary observations indicate that when a lagging chromosome is present in anaphase, the rates of poleward movement for other chromosomes are decreased.

We hypothesize this is due to an inward force produced by the lagging chromosome that perturbs the force balance across the spindle. Here, we test this hypothesis by (1) assessing whether the rates of poleward movement are homogeneous across the mitotic spindle in cells without lagging chromosomes and (2) investigating the role of the microtubule motor Eg5 in the transmission of forces during anaphase.

Future experiments will determine whether the effect of the lagging chromosome on poleward chromosome movement is homogeneous across the spindle or is stronger on the nearest neighbors compared to the farthest chromosomes. In addition, we will use our previously established force-balance quantitative model of chromosome dynamics to understand the mechanism by which the presence of a lagging chromosome perturbs the distribution of forces within the spindle.

BIOCOMPLEXITY INSTITUTE

MODELING RESPONSE TIMES FOR STRUCTURE FIRES

Madison Arnsbarger, Joshua Goldstein, Claire Kelling, and Gizem Korkmaz

In order to improve communities' general safety, to make the allocation of emergency resources more efficient, and to improve situational awareness, it is important to reduce response time to incidents.

In this paper we identify which factors significantly affect turnout times and travel times for Arlington County Fire Department in Virginia by applying linear and spatial Gaussian Process models to U.S. National Fire Incident Reporting System (NFIRS) data.

Due to the granular nature of the results, we are able to specify policy recommendations and opportunities for improvement. The uniformity of NFIRS data makes this paper's methodological innovations applicable to every participating fire department in the United States, and advances the effort to incorporate scientific evidence into government-level policymaking.

ANALYZING ONLINE SOCIAL EXPERIMENTS OF COLLECTIVE IDENTITY

Vanessa Cedeno, Christopher Kuhlman, Yihui Ren, Xinwei Deng, Zhihao Hu

Collective identity (CI) is an individual's cognitive, moral, and emotional connection with a broader community, category, practice, or institution. It is a perception of shared status or relation. CI reflects the willingness to place the needs of the group above own personal needs. A person without feelings of belonging to a group can develop this sense of belonging. The formation of CI involves collaboration, yet it's difficult to measure.

We conduct and analyze two-phase experiments, where in Phase 1, group members cooperate to form words from given letters, and in Phase 2, a public goods game and a DIFI game are played. Players are recruited from Amazon Mechanical Turk. The purpose of Phase 1 is to form CI among team members, and the purpose of Phase 2 is to measure the amount of CI that is formed.

We analyze the experimental data to determine what (experimental) parameters are and are not associated with CI formation. Then, following exploratory and statistical analyses, we analyze player behavior data to inform agent-based models that will provide insight for different contexts besides the experiment itself, and at scales well beyond those of the experiments. Our analysis includes quantifying dynamics of the game. This includes initial game parameters, word formation rate, sharing information through requests and replies of letters and temporal dynamics of the game.

BIOCOMPLEXITY INSTITUTE

BEHAVIOR MODEL CALIBRATION FOR EPIDEMIC SIMULATIONS

Meghendra Singh, Achla Marathe, Samarth Swarup, Madhav Marathe

Computational epidemiologists frequently employ large-scale simulations of human populations to study disease outbreaks and assess intervention strategies. These agent-based simulations function at some level of abstraction and make simplifying assumptions to make implementation feasible.

However, for these simulations to be reliable, it is essential that the underlying models of agent behavior capture the real-world decision-making process of human beings. This aspect becomes critical if such simulations are used to inform policy making, especially in the domain of public health.

In this work, we present a method for calibrating an agent behavior model using data obtained from an epidemiological survey. The epidemiological survey captures people's precautionary behaviors in response to a hypothetical influenza outbreak. We begin by identifying 350 precautionary behaviors (combinations of actions) exhibited by survey respondents during influenza outbreaks.

We formulate the behavior calibration problem as an inverse reinforcement learning problem, where we have to infer the costs of behaviors that result in a distribution of simulated behaviors that matches the observed distribution of behaviors in the survey. We compare the performance of various optimization algorithms for the behavior calibration problem.

Research Symposium | 2017-11-08 | Abstracts

THE ROLE OF PHAFIN2 IN AUTOPHAGY

Tuo-Xian Tang and Daniel G. S. Capelluto

Autophagy is a highly conserved cellular pathway in eukaryotic cells. A portion of the cytosol, which contains invading pathogens and long-lived proteins, is taken up by an autophagosome. This double-membrane organelle fuses with lysosomes, where the contents were digested by the lysosomal enzymes.

Previous data showed that Phafin2 was involved in the induction of autophagy. Phafin2 has two domains, N-terminal PH (Pleckstrin Homology) domain and C-terminal FYVE (Fab 1, YOTB, Vac 1, and EEA 1) domain. Both domains can bind the phospholipid phosphatidylinositol 3-phosphate (PtdIns(3)P). In this study, we aimed to investigate why Phafin2 has two PtdIns(3)P binding domains and how it facilitates the induction of autophagy. The binding affinity between PtdIns(3) P and Phafin2 was studied by surface plasmon resonance.

Results showed that PtdIns(3)P and Phafin2 had a strong binding, triggering minor conformational changes in the protein. Another interesting finding is that Phafin2 can cause membrane curvature, which may be required for tethering of lysosomes to autophagosomes, and consequently initiating autophagy.

Research Symposium | 2017-11-08 | Abstracts

THE PROTECTIVE ROLE OF NLRX1 DURING INVASIVE PULMONARY ASPERGILLOSIS

Tariq Ayubi, Andrew Leber, Nuria Tubau-Juni, Kelsey Simmons, Victoria Godfrey, Raquel Hontecillas, Josep Bassaganya-Riera, Shiv D. Kale

Aspergillus fumigatus is a ubiquitous opportunistic fungal pathogen afflicting diverse immunocompromised populations. Invasive pulmonary aspergillosis (IPA), a severe infection of the lungs by A. fumigatus, is associated with elevated mortality rates and progresses rapidly. With over 1.5 million deaths globally from fungal diseases, there is a need to develop novel antifungals.

The NLR family of proteins are critical for an appropriate immune response towards viruses, bacteria, and tumors. Here we describe the importance of host Nlrx1 in immunosuppressed models of IPA. Enhanced fungal burden was observed in Nlrx1-/- mice compared to wild-type (WT) using steroid, leukopenic, and neutropenic models of IPA. Survival studies using neutropenic mice resulted in significantly increased mortality in Nlrx1-/- mice in comparison to the wildtype. Nlrx1-/- mice were also noted to have an overall ablated recruitment of leukocytes populations as well as varied production of intracellular cytokines in a cell type specific manner.

In vitro fungal challenges of mouse bone marrow-derived macrophages (BMDMs), dendritic cells (BMDCs), and human BEAS-2B airway epithelial cells revealed cell type-specific patterns of Nlrx1 gene expression. Nlrx1-/- BMDCs processed significant numbers of conidia compared to WT, but were more prone to exhaustion. Nlrx1-/- BMDCs also secreted relatively lower quantities of chemokines and cytokines at multiple times post fungal challenge.

In conclusion, NLRX1 regulates immune signaling in a cell-type specific manner, is essential for leukocyte recruitment, and T-helper cell mediated responses during IPA. Further we provide a proof of concept for therapeutic development centered around NLRX1 for fungal infection.

Research Symposium | 2017-11-08 | Abstracts

MOLECULAR BASIS OF LIGAND BINDING BY THE ENDOSOMAL ADAPTOR PROTEIN TOM1

Wen Xiong, Phillip Choi, Jeff F. Ellena, Anne Brown, and Daniel G. S. Capelluto

Tom1 (target of Myb 1) plays a role in membrane trafficking by serving as an alternative endosomal sorting complex required for transport (ESCRT)-0 component. Tom1 possesses a N-terminal VHS domain followed by a central GAT domain. Tom1 has been shown to serve as a new phosphatidylinositol 5-phosphate (PI(5)P) effector at signaling endosomes through its VHS domain, delaying cargo degradation in a bacterial infection model.

The Tom1 VHS domain also binds ubiquitin moieties in cargo for endosomal transport and degradation; therefore, we hypothesize that the ubiquitin and PI(5)P compete each other for Tom1 VHS binding. In order to address this question, the backbone NMR resonances of Tom1 VHS were assigned.

Our heteronuclear single quantum coherence (HSQC) data, lipid binding overlay assay, and molecular dynamic simulations revealed that Tom1 VHS interacts with PI(5)P following a fast-exchange regime, with the PI(5)P binding site predicted to be at a region of the α-helix 8. HSQC analysis showed that the ubiquitin-binding site in Tom1 VHS is located at the α-helices 2, 5 and 7 of Tom1 VHS. Despite the binding sites are not overlapped, the ubiquitin and PI(5)P may compete each other by inducing conformational changes in the Tom1 VHS domain upon binding.

We will address ligand-dependent conformational changes through NMR dynamic studies of Tom VHS. By providing the molecular basis of the Tom1 interactions, we will generate cargo sorting mechanistic insights, create functionally specific mutations, and precisely manipulate alternative ESCRT-0 proteins.

Research Symposium | 2017-11-08 | Abstracts

EXPLORING THE SCOPE AND IMPACT OF OPEN SOURCE SOFTWARE

Eirik Iversen, Ben Swartz, Claire Kelling, Sayali Phadke, with Gizem Korkmaz and Stephanie Shipp

Open source software (OSS) plays an important but not well understood role in providing tools for the creation of useful goods and services, as well as tools for innovative activity. This type of software is developed, maintained, and extended both within the private sector and outside of it, through the contribution of people from universities, government research institutions, nonprofits, and individuals. Examples include Linux, Apache, Python, and R.

Current NCSES and other economic indicators do not measure the value of goods and services that do not have market transactions, i.e., Development of OSS is not captured in surveys nor are in economics measures such as the Gross Domestic Product (GDP).

This project aims to test the feasibility and to develop methods to measure the scope and impact (e.g., cost and benefit) of OSS. Our goal is to understand how much OSS is in use (stock) and how much is created (flow). We use two data sources, OpenHub and SourceForge, to characterize the development and the categories (users) of OSS, respectively.

BIOCOMPLEXITY INSTITUTE

MODELING THE IMPACT OF OPEN SOURCE SOFTWARE: A STUDY OF R PACKAGES

Eirik Iversen, Ronnie Fesco, Claire Kelling, Gizem Korkmaz

This project aims to identify the factors that affect the impact of Open Source Software (OSS), measured by the number of downloads and citations, focusing on a case study of R packages. We have collected all the R packages hosted on the repository CRAN, and on depsy.org, which is a website that compiles R and Python packages.

To date, information about 9,810 R packages have been scraped from Depsy.org (last update in Sept. 2015). The information for each package includes its contributors, number of commits, number of downloads, number of citations, and stars (identifying active development). We generate the dependency network of the packages collected from depsy.org, and develop statistical models that use the network characteristics and the package attributes.

To estimate the impact of R packages, we develop two Quasi-Poisson models[1] with the number of downloads (Model I) and the number of citations (Model II) as the dependent variables, and use the network characteristics and the package attributes as independent variables.

We find that the network centrality of a package (the number of packages that depend on it, and how they are connected) as well as package attributes (e.g., the number of contributors, the number of commits) are important factors in showing the impact of open source software, measured by the number of downloads and the number of citations.

Research Symposium | 2017-11-08 | Abstracts

ESTIMATING THE COST OF OPEN SOURCE SOFTWARE

Matthew Evans, Joey Ghebremichael, Yusheng Zhang, with Gizem Korkmaz and Stephanie Shipp

The goal of this project is to quantify the cost of creating Open Source Software (OSS) projects by developing new methods and models that use OSS features and developer information.

The project involves identifying existing cost models used for proprietary software, and adapting them to apply to creation of OSS projects. The cost method estimates the total effort to develop a software project using information embedded within the code about complexity.

Research Symposium | 2017-11-08 | Abstracts

STEM EDUCATION PATHWAYS

Ben Swartz, Madison Arnsbarger, Chanida Lerdritsomboon, Hannah Brinkley, Bianica Pires, Stephanie Ship

Data on traditional STEM (Science, Technology, Engineering, and Mathematics) pathways are well-documented and explored. Traditional pathways include obtaining a bachelors or PHD in a STEM field. The Census Bureau, for example, provides an interactive pathway map of bachelor's degrees to occupation on its website.

These tools shows that half of STEM occupations do not require a bachelor's degree. Additionally, women and minorities are more likely to shift to non-STEM fields on average. We aim to understand STEM pathways, particularly non-traditional pathways with a focus on women and minorities as they enter and leave the STEM field, as well as employer expectations of the STEM workforce.

BIOCOMPLEXITY INSTITUTE

DISCOVERING NON-TRADITIONAL DATA SOURCES FOR BUSINESS INNOVATION

Daniel Wilkin, Harsimrat Pandher, David Park, Joseph Kim with Gizem Korkmaz and Stephanie Shipp

The National Center for Science & Engineering Statistics at NSF traditionally measures innovation through surveys of selected companies (e.g., Business R&D and Innovation Survey (BRDIS)). While BRDIS measures innovation incidence, i.e., the number of innovating firms, NCSES is interested in exploring the possibility of using non-traditional data to richer and complementary innovation measures.

This project focuses on identifying non-traditional data sources to measure product innovation. Non-traditional data sources include administrative and opportunity data (e.g., product announcements, press releases, social media) that are obtained through databases, web-scraping or queries of selected company websites. The goal is to assess the feasibility of using these data sources and to develop methods to measure business innovation.

Research Symposium | 2017-11-08 | Abstracts

COMPARATIVE TRANSCRIPTOME ANALYSIS BETWEEN SPECIES USING NEXT GENERATION SEQUENCING DATA

Jiyoung Lee, Lenwood Heath, Ruth Grene, Song Li

Comparative transcriptome analysis is the comparison of expression patterns between homologous genes in different species. Because most gene functions in plants have been identified in the model species, Arabidopsis, comparative transcriptome analysis is particularly important for functional annotation of genes in other plant species.

One major challenge of comparative transcriptome analysis is to establish one-to-one mapping of the developmental stages between two species. To overcome this challenge, we converted the gene expression patterns into a co-expression network and then applied network module-finding algorithms to the cross-species co-expression network.

We predicted modules with homologous genes from both species with converged energy levels using a simulated annealing method. We tested this method using data from embryo developmental transcriptome in Arabidopsis and Soybean. We investigated the effect of using different co-expression thresholds and homologous weights on the number of modules identified in both species.

We found that including homologous edges in the co-expression network significantly increased the number of modules identified in both species, suggesting homologous genes form smaller co-expression clusters. We also found that, in some clusters, more than half of Arabidopsis genes have homologous genes that are expressed in similar patterns in soybean, suggesting these genes play conserved functions in both species.

BIOCOMPLEXITY INSTITUTE

SYNTHETIC HUMAN MOBILITY MODELING RESEARCH: GENERATIVE SYNTHESIS OF WIFI ASSOCIATION TRACES TO CREATE SYNTHETIC HUMAN MOBILITY PATTERNS

Mark E. DeYoung, Bryan Jonas, Ryan Kingery

The intent of this research is to support privacy preserving publication of WiFi association data that is representative of human motion through a WiFi network without the problems of identifiability.

We produce models from de-identified observational data and then generate synthetic data from inferred models. Because no observational data is released we hypothesize that generative synthesis will provide sufficient utility and privacy protection such that sanitized data can be published for study replication purposes.

We propose that generative synthesis can obscure high-order relational information (spatial, spatial-temporal) or other 'semantic' relationship that can be revealed by non-classical statistical learning techniques. We design, test, and evaluate direct generation of synthetic trajectories using machine learning models (e.g. generative adversarial networks, long short-term memories).

Research Symposium | 2017-11-08 | Abstracts

AESTHETIC RENDERINGS OF PUBLICATION MENTORSHIP NETWORKS

James Schlitt, Bryan Lewis, Stephen Eubank

Mentorship presents a profound part of the experience of academic research. A mentor whom provides guidance and wisdom may earn our lifelong admiration. A mentee who struggles may earn our compassion; one who fails, our empathy; and one whom succeeds, our greatest pride. While most researchers will work closely and develop strong relationships with their PIs, it's important to recognize those interstitial mentors. Those who may have helped us to fix a thorny bit of code, to procure a prized analyte, or to see us through a difficult gatekeeper.

In this project we sought to identify these networks of proximal mentors, providing a means of charting the flow of knowledge and wisdom through an academic organization. The presented visualization draws from a coauthorship network of the Network Dynamics and Simulation Science Laboratory (NDSSL) of the Biocomplexity Institute of Virginia Tech. Proximal mentorship was estimated by seeking neighbors of a given author-node with incrementally greater centrality. A tree structure then forms starting with the author with the highest centrality and fanning out across disparate disciplines and researchers.

Abstracts were queried from PubMed for each mentor/ mentee pairing, categorized and color coded by latent dirichlet allocation, and rendered on the tree in chronological order by date of publication. The visualization uses a natural metaphor to illustrate the branching flow of knowledge across disciplines originating within the Biocomplexity Institute.

Research Symposium | 2017-11-08 | Abstracts

A MACHINE LEARNING APPROACH TO IMPROVE TRANSCRIPTOME ASSEMBLY USING COMBINATION OF LONG READ AND SHORT READ TECHNOLOGIES

Qi Song, Song Li

Short reads RNA-seq followed by de novo transcriptome assembly has been widely used to investigate the transcriptome of species without reference genome. How to remove false positive transcripts from an assembly is still an open challenge for computational biology. Recently, single molecule sequencing (PacBio sequencing) has emerged as a promising method, as it directly sequences individual RNA molecules, with read length much longer than short reads. However, long reads from PacBio sequencing are known to have higher error rates than short reads and often cover only a small portion of all expressed genes.

In this study, we propose to combine both short reads and long reads sequencing data to produce high-quality transcriptome assembly. We used published short read and long read data generated from the model organism Arabidopsis as gold standard to train several machine learning methods including K-Nearest Neighbors, Random Forest, Support Vector Machine, Logistic Regression and Naïve Bayes to identify high-quality contigs. This model was then used to predict the quality of the contigs that are of intermediate quality. We further compared the performance of our method with two other transcriptome assembly evaluation tools, RSEM-eval and TransRate.

The results showed that our evaluation score has better correlation with contig F-scores than RSEM-eval and TransRate. The whole framework can serve as an automatic evaluation tool for transcriptome assembly without reference genome. We will apply this tool to the long reads and short reads sequencing data from a parasitic species, O. cernua, to generate a high quality transcriptome for this species.

Research Symposium | 2017-11-08 | Abstracts

DATA SCIENCE FRAMEWORK AND INFRASTRUCTURE

Daniel Chen

Bench laboratory science labs have systematic protocols and notebooks that document the instructions, notes, and results of laboratory tasks and experiments. Typically there is a canonical laboratory notebook that documents the progress of a particular experiment. This provides provenance for research findings, which can be used for replicability and reproducibility.

Generally speaking, experimental science provides a luxury that many computational and data science domains do not have, highly coupled experimentation, data collection, and data analysis workflows. Even still, science is in a reproducibility crisis. In data science, we typically repurpose data from its original intention, convoluting the data collection and analysis phase. This cyclic process of preparing data is critical to the provenance of the data, and needs to be well documented to have a reproducible and replicable analysis. Unless there is a specific software product or library, data analysis workflows are scarcely documented. While there are many resources that talk about “good practices”, they seldom discuss the actual implementation of these practices.

Here we present the data science framework and infrastructure implemented at the Social and Decision Analytics Laboratory and how we reduce the barrier to reproducible pipelines and documentation.

Research Symposium | 2017-11-08 | Abstracts

THE ROLE OF AF2 DURING INVASIVE PULMONARY ASPERGILLOSIS AND CELL SPECIFIC CHALLENGES

Samuel Light, Tariq Ayubi, Nuria Tubau-Juni, Raquel Hontecillas, Josep Bassaganya-Riera, Shiv D. Kale

Aspergillus fumigatus is an opportunistic fungus that effects immunocompromised individuals. Infections by the fungus may cause Invasive Pulmonary Aspergillosis (IPA), a fungal infection of the respiratory system that often leads to rapid mortality. Using a combination of RNA-Seq analysis and functional genomics we have identified a subset of putative conserved secreted genes expressed during IPA. We genetically knocked out of one of these genes Af2 in A. fumigatus and saw a significant lost of fungal burden during chemotherapeutic and neutropenic models of IPA.

Immunological analysis of WT and AF2 deficient strains indicate altered immune response and differential production of intracellular cytokines. We became interested in determining if Af2 facilitates survival and virulence against human BEAS-2B airway epithelial cells, mouse bone marrow derived macrophages (BMDMs), and mouse bone marrow derived dendritic cells (BMDCs). Using a flow cytometry based assay we observed loss of survival in the Af2 knockout in a cell type specific manner. Complementing AF2 back into the knockout strain restored and in certain cases partially enhanced the wild type phenotype.

Our findings suggest A. fumigatus secretes proteins that are important for virulence during IPA, and one of these proteins, AF2, appears to function in a cell type specific manner.

Research Symposium | 2017-11-08 | Abstracts

APPLYING AN INFECTIOUS DISEASE PERSPECTIVE IN ADDRESSING THE OPIOID EPIDEMIC IN THE APPALACHIAN REGION

Van Truong, Bryan Lewis

Opioid overdose deaths are rising at an alarming rate in the United States. It is believed that the "Opioid Crisis" began in the 1990s due to the over-prescription of opioid painkillers. The Center for Disease Control reported in 2016 that almost half of all opioid deaths involved prescription opioids. Opioids are a class of painkiller drugs including oxycontin, hydrocodone, and fentanyl that synthetically resemble opiates.

As more and more families and communities fight back against this silent epidemic, our team is applying an infectious disease approach to understanding the factors contributing to a community's susceptibility to developing higher incidences of opioid addictions. With a particular focus on the Appalachian region, we performed a univariate analysis of multiple sociodemographic factors such as accessibility to physical therapy, median income level, High Intensity Drug Trafficking Areas (HIDTA) designation by the federal government, and many more against the annual rate of drug overdoses by county.

We are able to show that these sociodemographic factors have varying relationships to the rate of drug overdoses in our region of interest, with distinct differences when further divided between the type of drugs such as all drugs, all opioids, prescription opioids excluding fentanyl, or fentanyl alone. In order to address the opioid epidemic in our communities, we must first examine and understand the relationship that certain social "exposure" factors have with the increasing rate of opioid-related deaths.

Research Symposium | 2017-11-08 | Abstracts

SCHEV: SYNTHETIC DATA

Aaron Schroeder, Lata Kodali, Kyle Morgan, Sean Pili, Ronald Fesco

The Virginia Longitudinal Data System (VLDS) contains sensitive information about citizens that cannot be released. Data privacy regulations prohibit any researcher from having access to all of the data at once. If researchers wish to have access to certain parts of the data, they often must guess which variables they need for their research. An incorrect guess means that they must restart the data request process.

The goal of this project was to create a secure synthetic dataset that is sufficiently privatized, but also usable for machine learning purposes. Essentially, any relationships present in the original data should still be present in the synthetic data so that researchers can use the synthetic data to determine which variables to request from the original data.

For privacy purposes, any combination of variables that could be used to identify a specific individual should be changed or modified beyond recognition. Using the “synthpop” package in R along with other methods, we were able to make significant progress toward creating a fully privatized dataset that reasonably preserves the relationships present in the original data.

Research Symposium | 2017-11-08 | Abstracts

DESIGN OF A DISABLED-2-DERIVED PEPTIDE TO IMPAIR PLATELET-MEDIATED CANCER CELL EXTRAVASATION

Wei Song, Andrew Biscardi, Carla Finkielstein, Daniel Capelluto

Disabled-2 (Dab2) targets cellular membranes and triggers a wide range of biological events, including endocytosis and platelet aggregation. Dab2, through its phosphotyrosine-binding (PTB) domain, inhibits platelet aggregation by competing with fibrinogen for allbb3 integrin receptor binding.

We have shown that the N-terminal region, including the PTB domain (N-PTB), drives Dab2 to the platelet membrane surface by binding to sulfatide through two sulfatide-binding motifs (SBM), which modulate both the extent of platelet aggregation and platelet-leukocyte interaction by inhibiting the sulfatide-induced P-selectin surface expression. A Dab2 peptide, representing SBM, can reversibly bind to sulfatide with moderate affinity, and when added to a platelet mixture, reduces the number and size of sulfatide-induced aggregates.

In addition, tumor cells are reported to have the ability of aggregating platelets, which occurs following tumor cell intravasation into the vasculature, thereby facilitating tumor cell migration, invasion and arrest within the vasculature. Contributions of platelets aggregation to tumor cell survival and spread suggest platelets as a new avenue for therapy.

Overall, our findings identify and structurally characterize a minimal region in Dab2 that modulates platelet homotypic and heterotypic interactions, all of which provide the foundation for rational design of a new anti-aggregatory peptide that can impair extravasation of cancer cells.

Research Symposium | 2017-11-08 | Abstracts

THE DIVERSE EFFECTS OF GM-CSF ON GRANULOCYTE-MONOCYTE PROGENITOR CELL DIFFERENTIATION: MECHANISMS OF ACTION

Bronson Weston, Liwu Li, John Tyson

Differentiation of granulocyte-monocyte progenitor (GMP) cells results in a variety of white blood cells, such as neutrophils and macrophages, which execute diverse tasks of the innate immune system. Due to the specific demands of the body, this process is heavily regulated via cytokines, such as granulocyte/macrophage colony stimulating factor (GM-CSF).

While many cytokines strongly favor one lineage over another, GM-CSF has an intriguing, dose dependent, effect on differentiation, as it favors monopoiesis and granulopoiesis at low and high concentrations, respectively. Although this behavior is well documented, the mechanisms behind the diverse behavioral responses of GMP cells to GM-CSF are not well understood.

Here, we propose a network of interactions between the GM-CSF receptor and transcription factors that control GMP differentiation. We convert the interactions into a set of differential equations, and explore the properties of this mathematical model using dynamical systems theory.

Our model successfully reproduces the concentration-dependent behavior of GM-CSF induced differentiation, and we propose a three-component mechanism driving this behavior. Furthermore, our model predicts GM-CSF induced differentiation of a particularly interesting phenotype, the monocytic myeloid-derived suppressor cell (M-MDSC). These cells are well known for their ability to promote tumor angiogenesis and metastasis. We demonstrate that the same mechanisms that leads to the concentration-dependent response of GMP cells to GM-CSF, makes GM-CSF a capable inducer of the M-MDSC phenotype.

Research Symposium | 2017-11-08 | Abstracts

COMPUTATIONAL CONSTRUCTION OF TOXICANT SIGNALING NETWORKS

Jeffrey Law, Sophia Orbach, Bronson Weston, and T.M. Murali

Humans are constantly exposed to complex mixtures of environmental chemicals. Several ongoing efforts seek to increase our knowledge of chemical effects. For example, the EPA/NIH-funded ToxCast and Tox21 initiatives monitor the effect of chemicals on selected proteins using high-throughput screening assays. Toxicogenomic databases store gene expression profiles after chemical exposure. However, each dataset probes a different dimension of cell's response. Moreover, these experiments ignore complex networks through which proteins interact and computational methods to integrate these data are underdeveloped. These major barriers limit their usefulness of these data.

We propose computing toxicant signaling networks to address these challenges. For a chemical, such a network is composed of regulatory, signaling and physical interactions connecting proteins perturbed as a result of exposure to that chemical. We describe algorithms that connect the responding proteins in ToxCast/Tox21 assays in the context of the underlying network of regulatory and physical interactions. For well-studied chemicals, our networks are enriched in biological processes that are known to be perturbed by the chemicals. Toxicant signaling networks promise to reveal important intermediate proteins that have not have been tested and physiological processes that have not been previously implicated in connection with the chemical.

Research Symposium | 2017-11-08 | Abstracts

QUANTIFYING EFFECTS OF NEUTROPHIL MEMORY ON MIGRATION PATTERNS USING MICROFLUIDIC PLATFORMS AND ODE MODELING OF THE MECHANISTIC MOLECULAR PATHWAYS

Brittany P. Boribong, Mark J. Lenzi, Sarah Kadelka, Stanca Ciupe, Liwu Li, and Caroline N. Jones

During sepsis, neutrophils migrate and accumulate in healthy organs instead of migrating toward the infection. In this study, we present a microfluidic platform to measure neutrophil chemotaxis in an opposing chemoattractant gradient to quantify decision-making. We use two chemoattractants: a pro-resolution (fMLP) and pro-inflammatory (LTB4) chemoattractant. Despite advances in the understanding of signaling molecules and pathways within neutrophils, understanding of the directional decision-making process is limited, and consequently, our abilities to modulate the activity of neutrophils restricted.

To test the importance of leukocyte memory, we stimulate cells with a pro-inflammatory mediator (LPS) at both high and low-doses. We show unstimulated cells migrate toward fMLP over LTB4 in a 2:1 ratio. Cells stimulated with high-dose LPS show migration in a similar ratio. Surprisingly, cells stimulated with a low-dose of LPS migrate toward fMLP and LTB4 in a 1:1 ratio, showing an increase in migration toward LTB4. This study suggests that low-dose LPS stimulation can alter the decision-making properties of the neutrophil to migrate toward an inflammatory signal over a bacterial infection.

To understand the molecular mechanism of this cell memory, we used an ODE-based dynamical framework to model the interaction of the mutually inhibitory GRK2 and GRK5 proteins and its role in neutrophil decision-making. Our computational model results show a bimodal switch between high and low levels of GRK2, indicating that GRK2 may be a protein of importance in neutrophil decision-making. In the future, this platform can be used for early diagnosis of sepsis or to test the effect of pro-resolving mediators on neutrophil function.