REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
14-05-2018	Final Report	16-Aug-2012 - 14-Feb-2018

4. TITLE AND SUBTITLE
Final Report: Associating Growth Conditions with Cellular Composition in Gram-negative Bacteria

5a. CONTRACT NUMBER
W911NF-12-1-0390

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611103

6. AUTHORS

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Texas at Austin
101 East 27th Street
Suite 5.300
Austin, TX 78712 -1532

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
61789-MA-MUR.160

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Claus Wilke
UU	UU	UU	UU		19b. TELEPHONE NUMBER 512-471-6028

RPPR Final Report
as of 21-Jun-2018

Agency Code:

Proposal Number: 61789MAMUR **Agreement Number: W911NF-12-1-0390**
INVESTIGATOR(S):

Name: PhD Claus O. Wilke
Email: cwilke@mail.utexas.edu
Phone Number: 5124716028
Principal: Y

Organization: **University of Texas at Austin**
Address: 101 East 27th Street, Austin, TX 787121532
Country: USA
DUNS Number: 170230239 EIN: 746000203
Report Date: 14-May-2018 Date Received: 14-May-2018
Final Report for Period Beginning 16-Aug-2012 and Ending 14-Feb-2018
Title: Associating Growth Conditions with Cellular Composition in Gram-negative Bacteria
Begin Performance Period: 16-Aug-2012 **End Performance Period:** 14-Feb-2018
Report Term: 0-Other
Submitted By: Bart Smith Email: bart.smith@utexas.edu
Phone: (000) 000-0000
Distribution Statement: 1-Approved for public release; distribution is unlimited.

STEM Degrees: 1 **STEM Participants:** 8

Major Goals: The overarching goal of this project is to develop an understanding of how bacterial growth conditions relate to bacterial physiology, and more importantly, how we may predict growth conditions from physiology. The association between growth conditions and physiology (as measured by cellular composition) has important applications both in bacterial forensics (e.g., identifying the source of a pathogen used in a deliberate attack) and in engineering applications. We are carrying out theoretical and experimental work to address this question. First, we are developing general statistical theory for Multiple-Input-Multiple-Output data sets. Second, we are developing theoretical and computational models that link bacterial physiology back to growth conditions. Third, we are collecting a comprehensive experimental data set of E. coli grown under a variety of different conditions. In this data set, we obtain a variety of cellular composition measurements for our samples, include RNA expression data, protein expression data, lipid abundance data, and metabolic flux data. Fourth, we are applying the statistical and computational methods to the experimental data set we are compiling, with the ultimate goal to be able to predict the specific conditions under which a sample was grown from the measured cellular composition.

Accomplishments: Over the duration of the grant, we have made substantial progress on all four goals. We have developed statistical theory for high-dimensional heterogenous data, and we have developed theoretical and computational models that link bacterial physiology back to growth conditions. At the same time, we have collected a comprehensive experimental data set comprised of RNA expression data, protein expression data, lipid abundance data, and metabolic flux data, and we have performed analysis and modeling work on that data set.

A detailed progress report is provided as a separate attachment.

Training Opportunities: Undergraduates, graduate research assistants, and postdoctoral fellows were all provided various training opportunities throughout this project. Specific training opportunities have been reported in previous annual reports.

For the last reporting period, there are no additional training opportunities to report.

Results Dissemination: The results from this project have been disseminated to the general scientific community through poster presentations, oral presentations, and through publications. A detailed list of publications is provided separately under "Products".

Honors and Awards: No honors or awards during the last reporting period.

RPPR Final Report
as of 21-Jun-2018

Protocol Activity Status:

Technology Transfer: No tech transfer to report during the last reporting period.

PARTICIPANTS:

Participant Type: Postdoctoral (scholar, fellow or other postdoctoral position)
Participant: Mehmet Caglar
Person Months Worked: 6.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Postdoctoral (scholar, fellow or other postdoctoral position)
Participant: Adam Hockenberry
Person Months Worked: 4.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Undergraduate Student
Participant: Nelson Morrow
Person Months Worked: 1.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Postdoctoral (scholar, fellow or other postdoctoral position)
Participant: Bartram Smith
Person Months Worked: 1.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Daria Sydykova
Person Months Worked: 2.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Staff Scientist (doctoral level)
Participant: Daniel Boutz

Person Months Worked: 4.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Staff Scientist (doctoral level)
Participant: Ophelia Papoulas
Person Months Worked: 4.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Alexander Crofts
Person Months Worked: 3.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Staff Scientist (doctoral level)
Participant: Daniel Deatherage
Person Months Worked: 6.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Staff Scientist (doctoral level)
Participant: Dennis Mishler
Person Months Worked: 3.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Postdoctoral (scholar, fellow or other postdoctoral position)
Participant: Xue Zhang
Person Months Worked: 4.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)

RPPR Final Report
as of 21-Jun-2018

Participant: Arion Stettner
Person Months Worked: 2.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Taiyao Wang
Person Months Worked: 5.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Meghan Thommes
Person Months Worked: 4.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Postdoctoral (scholar, fellow or other postdoctoral position)
Participant: Jannell Bazurto
Person Months Worked: 4.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Undergraduate Student
Participant: Leah Lambert
Person Months Worked: 4.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Staff Scientist (doctoral level)
Participant: Sergey Stolyar
Person Months Worked: 1.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

RPPR Final Report
as of 21-Jun-2018

Participant Type: Postdoctoral (scholar, fellow or other postdoctoral position)
Participant: Tomislav Ticak
Person Months Worked: 3.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Co PD/PI
Participant: Pradeep Ravikumar
Person Months Worked: 2.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Postdoctoral (scholar, fellow or other postdoctoral position)
Participant: David Inouye
Person Months Worked: 6.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Arun Sai Suggala
Person Months Worked: 5.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

ARTICLES:

RPPR Final Report
as of 21-Jun-2018

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Infection and Immunity
Publication Identifier Type: DOI Publication Identifier: 10.1128/IAI.01046-12
Volume: 81 Issue: 2 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: EptC of Campylobacter jejuni Mediates Phenotypes Involved in Host Interactions and Virulence
Authors:
Keywords: EptC, Campylobacter jejuni
Abstract: Campylobacter jejuni is a natural commensal of the avian intestinal tract. However, this bacterium is also the leading cause of acute bacterial diarrhea worldwide and implicated in development of Guillain-Barré syndrome. Like many bacterial pathogens, C. jejuni assembles complex surface structures that interface with the surrounding environment and are involved in pathogenesis. Recent work in C. jejuni identified a gene encoding a novel phosphoethanolamine (pEtN) transferase, EptC (Cj0256), that serves a promiscuous role in modifying the flagellar rod protein, FlgG, the lipid A domain of lipooligosaccharide (LOS) and several N-linked glycans. In this work, we report that EptC catalyzes the addition of pEtN to the first heptose sugar of the inner core oligosaccharide of LOS, a fourth enzymatic target. We also examine the role pEtN modification plays in circumventing detection and/or killing by host defenses. Specifically, we show that modification of C. jejuni lipid A with pEtN results
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Journal of Biological Chemistry
Publication Identifier Type: DOI Publication Identifier: 10.1074/jbc.M113.453324
Volume: 288 Issue: 13 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: The Origin of 8-Amino-3,8-dideoxy-D-manno-octulosonic Acid (Kdo8N) in the Lipopolysaccharide of Shewanella oneidensis
Authors:
Keywords: Lipopolysaccharide, LipidA, sugar biosynthesis
Abstract: Lipopolysaccharide (LPS, endotoxin) is an essential component of the outer monolayer of nearly all Gramnegative bacteria. LPS is composed of a hydrophobic anchor, known as lipid A, an inner core oligosaccharide, and a repeating O-antigen polysaccharide. The first sugar bridging the hydrophopic lipid A and the polysaccharide domain is 3-deoxy-Dmanno-octulosonic acid (Kdo) in nearly all species, and thus is critically important for LPS biosynthesis. Modifications to lipid A have been shown to be important for resistance to antimicrobial peptides as well as modulating recognition by the mammalian innate immune system. Therefore, lipid A derivatives have been used for development of vaccine strains and vaccine adjuvants. One derivative that has yet to be studied is 8-amino-3,8-dideoxy-Dmanno- octulosonic acid (Kdo8N), which is found exclusively marine bacteria of the genus Shewanella. Using bioinformatics, a candidate gene cluster for Kdo8N biosynthesis was identified in Shewanella oneiden
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

RPPR Final Report
as of 21-Jun-2018

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: PLoS Computational Biology
Publication Identifier Type: DOI Publication Identifier: 10.1371/journal.pcbi.1003091
Volume: 9 Issue: 6 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: The Ability of Flux Balance Analysis to Predict Evolution of Central Metabolism Scales with the Initial Distance to the Optimum
Authors:
Keywords: Flux Balance Analysis, Evolution
Abstract: The most powerful genome-scale framework to model metabolism, flux balance analysis (FBA), is an evolutionary optimality model. It hypothesizes selection upon a proposed optimality criterion in order to predict the set of internal fluxes that would maximize fitness. Here we present a direct test of the optimality assumption underlying FBA by comparing the central metabolic fluxes predicted by multiple criteria to changes measurable by a 13C-labeling method for experimentallyevolved strains. We considered datasets for three Escherichia coli evolution experiments that varied in their length, consistency of environment, and initial optimality. For ten populations that were evolved for 50,000 generations in glucose minimal medium, we observed modest changes in relative fluxes that led to small, but significant decreases in optimality and increased the distance to the predicted optimal flux distribution. In contrast, seven populations evolved on the poor substrate lactate for 900 generation
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Infection and Immunity
Publication Identifier Type: DOI Publication Identifier: 10.1128/IAI.01008-12
Volume: 81 Issue: 3 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: The Polysaccharide Capsule of Campylobacter jejuni Modulates the Host Immune Response
Authors:
Keywords: Polysaccharide
Abstract: Campylobacter jejuni is a major cause of bacterial diarrheal disease worldwide. The organism is characterized by a diversity of polysaccharide structures, including a polysaccharide capsule. Most C. jejuni capsules are known to be decorated non-stoichiometrically with methyl phosphoramidate (MeOPN). The capsule of C. jejuni 81-176 has been shown to be required for serum resistance, but here we show that an encapsulated mutant lacking the MeOPN modification, mpnC, was equally sensitive to serum killing as the non29 encapsulated mutant. A non-encapsulated mutant, kpsM, exhibited significantly reduced colonization compared to wildtype 81-176 in a mouse intestinal colonization model, and the mpnC mutant showed an intermediate level of colonization. Both mutants were associated with higher levels of IL-17 expression from lamina propria CD4+ cells compared to cells from animals infected with 81-176. In addition, reduced levels of TLR-4 and -2 activation were observed following in vitro stimu
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

RPPR Final Report
as of 21-Jun-2018

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Nature Reviews Microbiology
Publication Identifier Type: DOI Publication Identifier: 10.1038/nrmicro3047
Volume: 11 Issue: 7 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Fortifying the barrier: the impact of lipid A remodelling on bacterial pathogenesis
Authors:
Keywords: lipid A, bacterial pathogenesis
Abstract: Gram-negative bacteria decorate their outermost surface structure, lipopolysaccharide, with elaborate chemical moieties, which effectively disguises them from immune surveillance and protects them from the onslaught of host defences. Many of these changes occur on the lipid A moiety of lipopolysaccharide, a component that is crucial for host recognition of Gram-negative infection. In this Review, we describe the regulatory mechanisms controlling lipid A modification and discuss the impact of modifications on pathogenesis, bacterial physiology and bacterial interactions with the host immune system.
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Proceedings of the National Academy of Sciences
Publication Identifier Type: DOI Publication Identifier: 10.1073/pnas.1218080110
Volume: 110 Issue: 4 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Modulating the innate immune response by combinatorial engineering of endotoxin
Authors:
Keywords: outer membrane, cell envelope, gram negative
Abstract: Despite its highly inflammatory nature, LPS is a molecule with remarkable therapeutic potential. Lipid A is a glycolipid that serves as the hydrophobic anchor of LPS and constitutes a potent ligand of the Toll-like receptor (TLR)4/myeloid differentiation factor 2 receptor of the innate immune system. A less toxic mixture of monophosphorylated lipid A species (MPL) recently became the first new Food and Drug Administration-approved adjuvant in over 70 y. Whereas wild-type Escherichia coli LPS provokes strong inflammatory MyD88 (myeloid differentiation primary response gene 88)-mediated TLR4 signaling, MPL preferentially induces less inflammatory TRIF (TIR-domain– containing adaptor-inducing IFN-?)-mediated responses. Here, we developed a systemfor combinatorial structural diversification of E. coli lipid A, yielding a spectrum of bioactive variants that display distinct TLR4 agonist activities and cytokine induction. Mice immunized with engineered lipid A/antigen emulsions exhibited rob
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

RPPR Final Report
as of 21-Jun-2018

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: mBio
Publication Identifier Type: DOI Publication Identifier: 10.1128/mBio.00305-13
Volume: 4 Issue: 3 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: The Outer Surface Lipoprotein VolA Mediates Utilization of Exogenous Lipids by Vibrio cholerae
Authors:
Keywords: Outer surface, lipids
Abstract: Previous work from our laboratory showed that the Gram-negative aquatic pathogen Vibrio cholerae can take up a much wider repertoire of fatty acids than other Gram-negative organisms. The current work elaborated on the ability of V. cholerae to exploit an even more diverse pool of lipid nutrients from its environment. We have demonstrated that the bacterium can use lysophosphatidylcholine as a metabolite for growth. Using a combination of thin-layer chromatography and mass spectrometry, we also showed that lysophosphatidylcholine-derived fatty acid moieties can be used for remodeling the V. cholerae membrane architecture. Furthermore, we have identified a lysophospholipase, VolA (Vibrio outer membrane lysophospholipase A), required for these activities. The enzyme is well conserved in Vibrio species, is coexpressed with the outer membrane fatty acid transporter FadL, is one of very few surface-exposed lipoprotein enzymes to be identified in Gram-negative bacteria and the first instance
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Molecular BioSystems
Publication Identifier Type: DOI Publication Identifier: 10.1039/c3mb25513a
Volume: 9 Issue: 4 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: The proteomic response to mutants of the Escherichia coli RNA degradosome
Authors:
Keywords: proteomic, E. coli, RNA
Abstract: The Escherichia coli RNA degradosome recognizes and degrades RNA through the coordination of four main protein components, the endonuclease RNase E, the exonuclease PNPase, the RhlB helicase and the metabolic enzyme enolase. To help our understanding of the functions of the RNA degradosome, we quantified expression changes of >2300 proteins using mass spectrometry based shotgun proteomics in E. coli strains deficient in rhlB, eno, pnp (which displays temperature sensitive growth), or rne(1-602) which encodes a C-terminal truncation mutant of RNase E and is deficient in degradosome assembly. Global protein expression changes are most similar between the pnp and rhlB mutants, confirming the functional relationship between the genes. We observe down-regulation of protein chaperones including GroEL and DnaK (which associate with the degradosome), a decrease in translation related proteins in Dpnp, DrhlB and rne(1-602) cells, and a significant increase in the abundance of aminoacyl-tRNA syn
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

RPPR Final Report
as of 21-Jun-2018

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: FEBS Letters
Publication Identifier Type: DOI Publication Identifier: 10.1016/j.febslet.2013.07.032
Volume: 587 Issue: 17 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: The average enzyme principle
Authors:
Keywords: Enzyme regulation, Michaelis-Menten, Metabolic network, Enzyme kinetics, Systems biology
Abstract: The Michaelis–Menten equation for an irreversible enzymatic reaction depends linearly on the enzyme concentration. Even if the enzyme concentration changes in time, this linearity implies that the amount of substrate depleted during a given time interval depends only on the average enzyme concentration. Here, we use a time re-scaling approach to generalize this result to a broad category of multi-reaction systems, whose constituent enzymes have the same dependence on time, e.g. they belong to the same regulon. This ‘‘average enzyme principle’’ provides a natural methodology for jointly studying metabolism and its regulation.
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Biology Direct
Publication Identifier Type: DOI Publication Identifier: 10.1186/1745-6150-8-7
Volume: 8 Issue: 1 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Invariance and optimality in the regulation of an enzyme
Authors:
Keywords: Regulation, Michaelis-Menten
Abstract: Background: The Michaelis-Menten equation, proposed a century ago, describes the kinetics of enzyme-catalyzed biochemical reactions. Since then, this equation has been used in countless, increasingly complex models of cellular metabolism, often including time-dependent enzyme levels. However, even for a single reaction, there remains a fundamental disconnect between our understanding of the reaction kinetics, and the regulation of that reaction through changes in the abundance of active enzyme. Results: We revisit the Michaelis-Menten equation under the assumption of a time-dependent enzyme concentration. We show that all temporal enzyme profiles with the same average enzyme level yield identical substrate degradation– a simple analytical conclusion that can be thought of as an invariance principle, and which we validate experimentally using a ?-galactosidase assay. The ensemble of all time-dependent enzyme trajectories with the same average concentration constitutes a space of functi
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

RPPR Final Report
as of 21-Jun-2018

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Proceedings of the National Academy of Sciences
Publication Identifier Type: DOI Publication Identifier: 10.1073/pnas.1307485110
Volume: 110 Issue: 24 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: The cost of efficiency in energy metabolism
Authors:
Keywords: metabolism
Abstract: In a universe being dragged into disorder by the second law of thermodynamics, living cells must expend energy to maintain their complex organization. In addition to providing a carbon source for biosynthesis, the classical Embden–Meyerhof–Parnas (EMP) and Entner-Doudoroff (ED) pathways help to satisfy this energetic demand by generating ATP during glucose metabolism (1). Based on simple stoichiometry of reactants and products, the EMP pathway appears, at first blush, greatly preferable to the ED pathway, yielding twice as much ATP per glucose. If glucose breakdown and energy conservation are tightly coupled, why is the less-efficient ED pathway so prevalent? What has kept prokaryotic life in its entirety from casting off the ED pathway in favor of the more profitable EMP pathway? In PNAS, Flamholz et al. (2) address these questions by drawing on thermodynamics, enzyme kinetics, mathematical optimization, and genomics.
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Operations Research
Publication Identifier Type: DOI Publication Identifier: 10.1287/opre.1120.1115
Volume: 60 Issue: 6 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Inverse Optimization: A New Perspective on the Black-Litterman Model
Authors:
Keywords: Finance: portfolio optimization. Programming: inverse optimization. Statistics: estimation
Abstract: The Black-Litterman (BL) model is a widely used asset allocation model in the ♀nancial industry. In this paper, we provide a new perspective. The key insight is to replace the statistical framework in the original approach with ideas from inverse optimization. This insight allows us to signi♀cantly expand the scope and applicability of the BL model. We provide a richer formulation that, unlike the original model, is exible enough to incorporate investor information on volatility and market dynamics. Equally importantly, our approach allows us to move beyond the traditional mean-variance paradigm of the original model and construct \BL"-type estimators for more general notions of risk such as coherent risk measures. Computationally, we introduce and study two new \BL"-type estimators and their corresponding portfolios: a Mean Variance Inverse Optimization (MV-IO) portfolio and a Robust Mean Variance Inverse Optimization (RMV-IO) portfolio. These two approaches are motivated by ideas fro
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

RPPR Final Report
as of 21-Jun-2018

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: IEEE Transactions on Signal Processing
Publication Identifier Type: DOI Publication Identifier: 10.1109/TSP.2013.2265679
Volume: 61 Issue: 17 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Model-Free Stochastic Localization of CBRN Releases
Authors:
Keywords: Source detection, source localization, compositehypothesis testing, large deviations, optimization, sensor placement.
Abstract: We present a novel two-stage methodology for locating a Chemical, Biological, Radiological, or Nuclear (CBRN) source in an urban area using a network of sensors. In contrast to earlier work, our approach does not solve an inverse dispersion problem but relies on data obtained from a simulation of the CBRN dispersion to obtain probabilistic descriptors of sensor measurements under a variety of CBRN release scenarios. At its first stage, subsequent sensor observations under nominal, event-free conditions are assumed to be independent and identically distributed and we rely on the method of types to detect a CBRN event. Conditional on such an event, subsequent sensor observations are assumed to follow a Markov process. Using composite hypothesis testing we map sensor measurements to a source location chosen out of a discrete set of possible locations. We leverage large deviation techniques to obtain a bound on the localization probability of error and propose several methodologies for fu
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Molecular BioSystems
Publication Identifier Type: DOI Publication Identifier: 10.1039/c3mb70606k
Volume: 0 Issue: 0 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Engineering reduced evolutionary potential for synthetic biology
Authors:
Keywords: synthetic biology, evolution, genome
Abstract: The field of synthetic biology seeks to engineer reliable and predictable behaviors in organisms from collections of standardized genetic parts. However, unlike other types of machines, genetically encoded biological systems are prone to changes in their designed sequences due to mutations in their DNA sequences after these devices are constructed and deployed. Thus, biological engineering efforts can be confounded by undesired evolution that rapidly breaks the functions of parts and systems, particularly when they are costly to the host cell to maintain. Here, we explain the fundamental properties that determine the evolvability of biological systems. Then, we use this framework to review current efforts to engineer the DNA sequences that encode synthetic biology devices and the genomes of their microbial hosts to reduce their ability to evolve and therefore increase their genetic reliability so that they maintain their intended functions over longer timescales.
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

RPPR Final Report
as of 21-Jun-2018

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Journal of Proteome Research
Publication Identifier Type: DOI Publication Identifier: 10.1021/pr4011684
Volume: 0 Issue: 0 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Protein-to-mRNA Ratios Are Conserved between
Authors:
Keywords: protein, mRNA, P. aeruginosa, -omics
Abstract: Recent studies have shown that the concentrations of proteins expressed from orthologous genes are often conserved across organisms and to a greater extent than the abundances of the corresponding mRNAs. However, such studies have not distinguished between evolutionary (e.g., sequence divergence) and environmental (e.g., growth condition) effects on the regulation of steady-state protein and mRNA abundances. Here, we systematically investigated the transcriptome and proteome of two closely related Pseudomonas aeruginosa strains, PAO1 and PA14, under identical experimental conditions, thus controlling for environmental effects. For 703 genes observed by both shotgun proteomics and microarray experiments, we found that the protein-to-mRNA ratios are highly correlated between orthologous genes in the two strains to an extent comparable to protein and mRNA abundances. In spite of this high molecular similarity between PAO1 and PA14, we found that several metabolic, virulence, and antibioti
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Nature Chemical Biology
Publication Identifier Type: DOI Publication Identifier: 10.1038/nchembio.1450
Volume: 10 Issue: 3 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Bacteriophages use an expanded genetic code on evolutionary paths to higher fitness
Authors:
Keywords: Bacteriophage, genetic code, evolution
Abstract: Bioengineering advances have made it possible to fundamentally alter the genetic codes of organisms. However, the evolutionary consequences of expanding an organism's genetic code with a noncanonical amino acid are poorly understood. Here we show that bacteriophages evolved on a host that incorporates 3-iodotyrosine at the amber stop codon acquire neutral and beneficial mutations to this new amino acid in their proteins, demonstrating that an expanded genetic code increases evolvability.
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

RPPR Final Report
as of 21-Jun-2018

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: IEEE Transactions on Biomedical Engineering
Publication Identifier Type: DOI Publication Identifier: 10.1109/TBME.2013.2280636
Volume: 61 Issue: 2 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Predicting and Evaluating the Effect of Bivalirudin in Cardiac Surgical Patients
Authors:
Keywords: Bivalirudin, regularized regression, optimization
Abstract: Bivalirudin, used in patients with heparin-induced thrombocytopenia, is a direct thrombin inhibitor. Since it is a rarely used drug, clinical experience with its dosing is sparse. We develop two approaches to predict the Partial Thromboplastin Time (PTT) based on bivalirudin infusion rates. The first approach is model-free and utilizes regularized regression. It is flexible enough to use as predictors bivalirudin infusion rates measured over several time instances before the time at which a PTT prediction is sought. The second approach is model-based and proposes a specific model for obtaining PTT which uses a shorter history of past measurements. We learn populationwide model parameters by solving a nonlinear optimization problem. We also devise an adaptive algorithm based on the extended Kalman filter that can adapt model parameters to individual patients. The latter adaptive model emerges as the most promising as it yields reduced mean error compared to the model-free approach. The
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Analytical Chemistry
Publication Identifier Type: DOI Publication Identifier: 10.1021/ac403796n
Volume: 86 Issue: 4 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: 193 nm Ultraviolet Photodissociation Mass Spectrometry for the Structural Elucidation of Lipid A Compounds in Complex Mixtures
Authors:
Keywords: lipidA, 193nm UVPD
Abstract: Here we implement ultraviolet photodissociation (UVPD) in an online liquid chromatographic tandem mass spectrometry (MS/MS) strategy to support analysis of complex mixtures of lipid A combinatorially modified during development of vaccine adjuvants. UVPD mass spectrometry at 193 nm was utilized to characterize the structures and fragment ion types of lipid A from Escherichia coli, Vibrio cholerae, and Pseudomonas aeruginosa using an Orbitrap mass spectrometer. The fragment ions generated by UVPD were compared to those from collision induced dissociation (CID) and higher energy collision dissociation (HCD) with respect to the precursor charge state. UVPD afforded the widest array of fragment ion types including acyl chain C-O, C-N, and C-C bond cleavages and glycosidic C-O and cross ring cleavages, thus providing the most comprehensive structural analysis of the lipid A. UVPD exhibited virtually no dependence on precursor ion charge state and was best at determining lipid A structure in
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

RPPR Final Report
as of 21-Jun-2018

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Molecular Microbiology
Publication Identifier Type: DOI Publication Identifier: 10.1111/mmi.12501
Volume: 91 Issue: 5 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Identification of a broad family of lipid A late acyltransferases with non-canonical substrate specificity
Authors:
Keywords: acyltransferase, lipidA, mass-spectrometry
Abstract: Most Gram-negative organisms produce lipopolysaccharide (LPS), a complex macromolecule anchored to the bacterial membrane by the lipid A moiety. Lipid A is synthesized via the Raetz pathway, a conserved nine-step enzymatic process first characterized in Escherichia coli. The Epsilonproteobacterium Helicobacter pylori uses the Raetz pathway to synthesize lipid A; however, only eight of nine enzymes in the pathway have been identified in this organism. Here, we identify the missing acyltransferase, Jhp0255, which transfers a secondary acyl chain to the 3'-linked primary acyl chain of lipid A, an activity similar to that of E.?coli LpxM. This enzyme, reannotated as LpxJ due to limited sequence similarity with LpxM, catalyses addition of a C12:0 or C14:0 acyl chain to the 3'-linked primary acyl chain of lipid A, complementing an E. coli?LpxM mutant. Enzymatic assays demonstrate that LpxJ and homologues in Campylobacter jejuni and Wolinella succinogenes can act before the 2' secondary acylt
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Infection and Immunity
Publication Identifier Type: DOI Publication Identifier: 10.1128/IAI.01057-13
Volume: 82 Issue: 2 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: The Vps/VacJ ABC Transporter Is Required for Intercellular Spread of Shigella flexneri
Authors:
Keywords: ABC Transporter, lipid asymmetry, S. flexneri
Abstract: The Vps/VacJ ABC transporter system is proposed to function in maintaining the lipid asymmetry of the outer membrane. Mutations in vps or vacJ in Shigella flexneri resulted in increased sensitivity to lysis by the detergent sodium dodecyl sulfate (SDS), and the vpsC mutant showed minor differences in its phospholipid profile compared to the wild type. vpsC mutants were unable to form plaques in cultured epithelial cells, but this was not due to a failure to invade, to replicate intracellularly, or to polymerize actin via IcsA for movement within epithelial cells. The addition of the outer membrane phospholipase gene pldA on a multicopy plasmid in a vpsC or vacJ mutant restored its resistance to SDS, suggesting a restoration of lipid asymmetry to the outer membrane. However, the pldA plasmid did not restore the mutant's ability to form plaques in tissue culture cells. Increased PldA levels also failed to restore the mutant's phospholipid profile to that of the wild type. We propose a du
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: PLoS Computational Biology
Publication Identifier Type: DOI Publication Identifier: 10.1371/journal.pcbi.1003195
Volume: 9 Issue: 8 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Flux Imbalance Analysis and the Sensitivity of Cellular Growth to Changes in Metabolite Pools
Authors:
Keywords: Flux balance analysis, metabolism, E. coli
Abstract: Stoichiometric models of metabolism, such as flux balance analysis (FBA), are classically applied to predicting steady state rates - or fluxes - of metabolic reactions in genome-scale metabolic networks. Here we revisit the central assumption of FBA, i.e. that intracellular metabolites are at steady state, and show that deviations from flux balance (i.e. flux imbalances) are informative of some features of in vivo metabolite concentrations. Mathematically, the sensitivity of FBA to these flux imbalances is captured by a native feature of linear optimization, the dual problem, and its corresponding variables, known as shadow prices. First, using recently published data on chemostat growth of Saccharomyces cerevisae under different nutrient limitations, we show that shadow prices anticorrelate with experimentally measured degrees of growth limitation of intracellular metabolites. We next hypothesize that metabolites which are limiting for growth (and thus have very negative shadow price)
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Annals of Operations Research
Publication Identifier Type: DOI Publication Identifier: 10.1007/s10479-013-1467-4
Volume: 0 Issue: 0 First Page #: 0
Date Submitted: Date Published:
Publication Location:
Article Title: Distribution-dependent robust linear optimization with applications to inventory control.
Authors:
Keywords: linear programming, distributional information
Abstract: This paper tackles linear programming problems with data uncertainty and applies it to an important inventory control problem. Each element of the constraint matrix is subject to uncertainty and is modeled as a random variable with a bounded support. The classical robust optimization approach to this problem yields a solution with guaranteed feasibility. As this approach tends to be too conservative when applications can tolerate a small chance of infeasibility, one would be interested in obtaining a less conservative solution with a certain probabilistic guarantee of feasibility. A robust formulation in the literature produces such a solution, but it does not use any distributional information on the uncertain data. In this work, we show that the use of distributional information leads to an equally robust solution (i.e., under the same probabilistic guarantee of feasibility) but with a better objective value. In particular, by exploiting distributional information, we establish stron
Distribution Statement: 1-Approved for public release; distribution is unlimited.
Acknowledged Federal Support:

CONFERENCE PAPERS:

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: IEEE International Conference on Robotics and Automation
Date Received: 13-Sep-2016 Conference Date: 14-May-2012 Date Published:
Conference Location: RiverCentre, Saint Paul, Minnesota, USA
Paper Title: Temporal logic motion control using actor-critic methods
Authors: Xu Chu Dingy, Jing Wangz, Morteza Lahijanianz, Ioannis Ch. Paschalidisz, Calin A. Belta
Acknowledged Federal Support: **Y**

RPPR Final Report
as of 21-Jun-2018

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: Advances in Neural Information Processing Systems (NIPS)
Date Received: 22-Aug-2016 Conference Date: 07-Dec-2015 Date Published: 12-Dec-2015
Conference Location: Montréal, Quebec, Canada
Paper Title: Fixed-Length Poisson MRF: Adding Dependencies to the Multinomial
Authors: David Inouye, Pradeep Ravikumar, Inderjit Dhillon
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: International Conference on Machine Learning (ICML) 33
Date Received: 22-Aug-2016 Conference Date: 19-Jun-2016 Date Published: 19-Jun-2016
Conference Location: New York, New York, USA
Paper Title: Square Root Graphical Models: Multivariate Generalizations of Univariate Exponential Families that Permit Positive Dependencies
Authors: David Inouye, Pradeep Ravikumar, Inderjit S. Dhillon
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: Advances in Neural Information Processing Systems (NIPS) 28
Date Received: 22-Aug-2016 Conference Date: 07-Dec-2015 Date Published: 07-Dec-2015
Conference Location: Montréal, Montréal CANADA
Paper Title: Consistent Multilabel Classification
Authors: Oluwasanmi Koyejo, Nagarajan Natarajan, Pradeep Ravikumar, Inderjit Dhillon
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: In International Conference on Machine Learning (ICML) 33
Date Received: 22-Aug-2016 Conference Date: 19-Jun-2016 Date Published: 19-Jun-2016
Conference Location: New York, New York, USA
Paper Title: PD-Sparse : A Primal and Dual Sparse Approach to Extreme Multiclass and Multilabel Classification
Authors: Ian Yen, Xiangru Huang, Kai Zhong, Pradeep Ravikumar, Inderjit S. Dhillon
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 2015 IEEE International Conference on Automation Science and Engineering (CASE)
Date Received: 22-Aug-2016 Conference Date: 24-Aug-2015 Date Published:
Conference Location: Gothenburg, Sweden
Paper Title: Sensing and classifying roadway obstacles: The street bump anomaly detection and decision support system
Authors: Theodora Brisimi, Setareh Ariafar, Yue Zhang, Christos Cassandras, Ioannis Paschalidis
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 2016 IEEE International Conference on Robotics and Automation (ICRA)
Date Received: 22-Aug-2016 Conference Date: 16-May-2016 Date Published:
Conference Location: Stockholm, Sweden
Paper Title: Cooperative multi-quadrotor pursuit of an evader in an environment with no-fly zones
Authors: Alyssa Pierson, Armin Ataei, Ioannis Paschalidis, Mac Schwager
Acknowledged Federal Support: **Y**

RPPR Final Report
as of 21-Jun-2018

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 2nd Workshop on Data Mining for Medical Informatics (DMMI 2015)
Date Received: 22-Aug-2016 Conference Date: 14-Nov-2015 Date Published: 14-Nov-2015
Conference Location: San Francisco, California USA
Paper Title: A Joint Clustering and Classification Approach for Healthcare Predictive Analytics
Authors: Wuyang Dai, Theodora Brisimi, Tingting Xu, Taiyao Wang, Venkatesh Saligrama, Ioannis Paschalidis
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 2015 54th IEEE Conference on Decision and Control (CDC)
Date Received: 22-Aug-2016 Conference Date: 15-Dec-2015 Date Published:
Conference Location: Osaka
Paper Title: An improved composite hypothesis test for Markov models with applications in network anomaly detection
Authors: Jing Zhang, Ioannis Paschalidis
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 2015 54th IEEE Conference on Decision and Control (CDC)
Date Received: 22-Aug-2016 Conference Date: 15-Dec-2015 Date Published:
Conference Location: Osaka
Paper Title: Learning cellular objectives from fluxes by inverse optimization
Authors: Qi Zhao, Arion Stettner, Ed Reznik, Daniel Segre, Ioannis Paschalidis
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 2015 54th IEEE Conference on Decision and Control (CDC)
Date Received: 22-Aug-2016 Conference Date: 15-Dec-2015 Date Published:
Conference Location: Osaka
Paper Title: Quadrotor deployment for emergency response in smart cities: A robust MPC approach
Authors: Armin Ataei, Ioannis Paschalidis
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: the 23rd ACM SIGKDD International Conference
Date Received: 23-Aug-2017 Conference Date: 13-Aug-2017 Date Published:
Conference Location: Halifax, NS, Canada
Paper Title: PPDsparse A Parallel Primal-Dual Sparse Method for Extreme Classification
Authors: Ian E.H. Yen, Xiangru Huang, Wei Dai†, Pradeep Ravikumar, Inderjit Dhillon, Eric Xing
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 30th Conference on Neural Information Processing Systems
Date Received: 23-Aug-2017 Conference Date: 05-Dec-2016 Date Published: 05-Dec-2016
Conference Location: Barcelona, Spain
Paper Title: Dual Decomposed Learning with Factorwise Oracles for Structural SVMs of Large Output Domain
Authors: Ian E.H. Yen, Xiangru Huang, Kai Zhong, Ruohan Zhang, Pradeep Ravikumar, Inderjit S. Dhillon
Acknowledged Federal Support: **Y**

RPPR Final Report
as of 21-Jun-2018

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: Proceedings of the 20th International Conference on Artificial Intelligence and Statistics
Date Received: 23-Aug-2017 Conference Date: 09-May-2017 Date Published: 09-May-2017
Conference Location: Fort Lauderdale, Florida, USA
Paper Title: Greedy Direction Method of Multipliers for MAP Inference with Large Output Domain.
Authors: Xiangru Huang, Ian En-Hsu Yen, Ruohan Zhang, Qixing Huang, Pradeep Ravikumar, Inderjit Dhillon
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 2017 IEEE 56th Annual Conference on Decision and Control (CDC)
Date Received: 14-May-2018 Conference Date: 12-Dec-2017 Date Published:
Conference Location: Melbourne, Australia
Paper Title: Strategic equilibrium bidding for electricity suppliers in a day-ahead market using inverse optimization
Authors: Ruidi Chen, Ioannis Paschalidis, Michael Caramanis
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 2017 IEEE 56th Annual Conference on Decision and Control (CDC)
Date Received: 14-May-2018 Conference Date: 12-Dec-2017 Date Published:
Conference Location: Melbourne, VIC
Paper Title: Data-driven estimation of travel latency cost functions via inverse optimization in multi-class transportation networks
Authors: Jing Zhang, Ioannis Paschalidis
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 2017 IEEE 56th Annual Conference on Decision and Control (CDC)
Date Received: 14-May-2018 Conference Date: 12-Dec-2017 Date Published:
Conference Location: Melbourne, Australia
Paper Title: Feature extraction in Q-learning using neural networks
Authors: Henghui Zhu, Ioannis Paschalidis, Michael E. Hasselmo
Acknowledged Federal Support: **Y**

DISSERTATIONS:

Publication Type: Thesis or Dissertation
Institution: Boston University
Date Received: 22-Aug-2016 Completion Date: 5/1/16 7:14PM
Title: Control and Optimization Methods in Biomedical Systems: From Cells to Humans
Authors: Qi Zhao
Acknowledged Federal Support: **Y**

Publication Type: Thesis or Dissertation
Institution: Boston University
Date Received: 22-Aug-2016 Completion Date: 4/28/15 12:14AM
Title: Anomaly detection and dynamic decision making for stochastic systems
Authors: JING WANG
Acknowledged Federal Support: **N**

RPPR Final Report
as of 21-Jun-2018

Publication Type: Thesis or Dissertation
Institution: Boston University
Date Received: 22-Aug-2016 Completion Date: 4/9/16 12:14AM
Title: Detection and prediction problems with applications in personalized health care
Authors: WUYANG DAI
Acknowledged Federal Support: **N**

Publication Type: Thesis or Dissertation
Institution: Boston University
Date Received: 22-Aug-2016 Completion Date: 4/27/16 12:14AM
Title: Optimization methods for side-chain positioning and macromolecular docking
Authors: Mohammad Moghadasi
Acknowledged Federal Support: **N**

Publication Type: Thesis or Dissertation
Institution: Boston University
Date Received: 22-Aug-2016 Completion Date: 2/18/16 10:57PM
Title: Functional and evolutionary implications of in silico gene deletions
Authors: CHRISTOPHER JACOBS
Acknowledged Federal Support: **N**

Publication Type: Thesis or Dissertation
Institution: Boston University
Date Received: 22-Aug-2016 Completion Date: 1/14/16 10:57PM
Title: Transcriptomic and computational approaches for interrogating metabolic interactions in the coral microbiome
Authors: BRIAN R. GRANGER
Acknowledged Federal Support: **N**

MURI Progress Report: Associating Growth Conditions with Cellular Composition in Gram-negative Bacteria

Final Report, May 2018

Statement: The overarching goal of this project has been to develop an understanding of how bacterial growth conditions relate to bacterial physiology, and more importantly, how we may predict growth conditions from physiology. The association between growth conditions and physiology (as measured by cellular composition) has important applications both in bacterial forensics (e.g., identifying the source of a pathogen used in a deliberate attack) and in engineering applications. We have carried out theoretical and experimental work to address this question. We have developed statistical theory for high-dimensional heterogenous data, and we have developed theoretical and computational models that link bacterial physiology back to growth conditions. At the same time, we have collected a comprehensive experimental data set comprised of RNA expression data, protein expression data, lipid abundance data, and metabolic flux data, and we have performed analysis and modeling work on that data set.

1. Developments in statistical theory and modeling

1.1 Statistical theory

We have developed a novel toolbox of parametric models for high-dimensional heterogenous data, for both unsupervised and supervised contexts. Parametric models use a finite set of parameters or weights to specify the statistical model and are useful in sample-scarce settings, which is the case for most biological datasets. Such parametric models however have the typical caveat of being too simple (for instance, linear models) to be able to sufficiently capture the nuances in the data. In our work, we finessed these bottlenecks, and developed flexible parametric models that have a small number of parameters, even for high-dimensional data with a large number of inputs and outputs, by imposing natural constraints such as sparsity of direct associations, which entail that each variable is not directly associated with too many other variables. Another typical caveat with parametric models is that they are typically designed for a single data type, such as numeric or categorical. Moreover, some data types, such as counts and ordinal data, due to their complexity, have comparatively less flexible models. Our toolbox of flexible models are applicable to heterogeneous data, where each variable could belong to a different data type, which opens up the use of statistical models for jointly modeling heterogeneous data such as genomics, transcriptomics, and imaging data, among others. Our modeling toolbox holds for both so-called supervised learning settings, where we predict a potentially multivariate response given covariates, and unsupervised learning settings, where we aim to jointly model the high-dimesional data.

1.2 Modeling of bacterial metabolism

Our work has focused on two challenges. The first was inferring bacterial cell metabolic objectives from metabolic reaction flux data. We have developed an efficient inverse Flux Balance Analysis (invFBA) approach, based on linear programming duality, to characterize the space of possible objective functions compatible with measured fluxes. After testing our

algorithm on simulated *E. coli* data and time-dependent *S. oneidensis* fluxes inferred from gene expression, we applied our inverse approach to flux measurements in long-term evolved *E. coli* strains, revealing objective functions that provide insight into metabolic adaptation trajectories.

A second major challenge has been understanding distributed metabolic activity among different members of a microbial community. As members of ecosystems, microbes often confront a tradeoff between metabolic independence, requiring a large number of costly functions, and reliance on adjacent organisms, through networks of metabolic exchange entailing a division of labor. This balance of conflicting strategies is likely a key determinant of microbial community structure and dynamics, with important implications for microbiome research and synthetic ecology. An ideal experiment to investigate this tradeoff would involve gradually limiting the number of metabolic reactions allowed in multiple initially identical organisms, until individual microbes are unable to grow in the absence of metabolically diverse cross-feeding partners. While experimentally challenging, this test can be easily implemented through in silico genome-scale models. We developed a set of algorithms to identify tradeoff solutions for this problem at varying numbers of allowed intracellular and transport reactions and apply them to study the division of labor in communities of *E. coli* variants. Our formulation finds division of labor strategies that are more complex than previously engineered cross-feeding of essential biomass components (e.g. amino acids). More broadly, we systematically mapped the landscape of possible 1-, 2-, and 3-species solutions at increasingly tight constraints on the number of allowed reactions. This landscape displays a nonlinear boundary, indicating that the loss of an intracellular reaction cannot be typically compensated by a single imported metabolite. Transitions between different regimes are accompanied by dramatic and consistent changes in the patterns of exchanged metabolites, suggesting the existence of an underlying hierarchy of division of labor strategies. We further predict the existence of specific areas in this landscape in which independent bacteria are feasible, but outcompeted by cross-feeding pairs, providing a theoretical ground for the rise and advantage of division of labor in evolutionary processes.

2. Experimental work

2.1 Changes in gene expression under different growth conditions

To determine cell composition changes in response to environmental and genetic perturbations, we have grown *E. coli* samples under multiple conditions (nearly 200 samples total over the duration of the grant) and have measured RNAseq transcriptomics data and mass-spec proteomics data. Cell composition measurements for the baseline *E. coli* B strain REL606 under different salt, nutrient, and temperature stresses were analyzed by the entire MURI team. In addition, the Barrick lab analyzed two further multiomics datasets with systematic genetic perturbations related to laboratory evolution on timescales of months to years, using strains from a long-term evolution experiment (LTEE) that started from the same REL606 strain. The first study searched for changes in cell composition that potentiated a rare evolutionary innovation: the appearance of aerobic citrate utilization in one *E. coli* population after 15 years. The second study examined a set of strains with different first-step mutations from early in the LTEE that lead to two distinct evolutionary pathways, one of which was more

beneficial in the short term but reliably went extinct in the longer term. This work has led to new hypotheses about how we can recognize changes in cell composition that constrain or enable future adaptive pathways and evolutionary innovations.

We have also taken advantage of a new discovery in the Marx lab – the first protein known to turn down or halt translation due to sensing an intracellular intermediate – to compare this form of translational inhibition to that via antibiotics. We have generated transcriptomic and proteomic data of these experiments, and our work has revealed several key genes involved in this process.

2.2 Changes in lipid composition

Lipidomic analysis of the bacterial membrane continues to reveal major alterations in lipid content in a manner dependent on growth parameters. We found that changes in lipid composition are reproducible and relatively facile to detect with our current protocols. Our lipid profiling in combination with RNA, protein, and metabolite analysis provided useful molecular indicators of specific growth parameters. These efforts ultimately contributed to the development of a more robust biomarker profile. The ability to deduce growth media composition, duration of growth, and other physiological parameters based on the molecular profile of any given bacterial sample will greatly improve our forensic analysis capabilities. For example, we saw that major alterations to the lipid content of the bacterial membrane for strains grown over extended periods to time including additional acylation of gram-negative lipopolysaccharide. Changes also occurred in the bacterial glycerophsopholipid fraction with increased cyclo-proponated species. Additionally, we investigated how key lipid species impacted overall bacterial fitness by performing saturating transposon (Tn) mutagenesis in multiple genetic backgrounds lacking specific phospholipids followed by identification of transposon insertion sites using NextGen sequencing (Tn-Seq). This work has allowed us to begin to determine how specific lipids impact fitness of bacterial under different growth conditions.

2.3 Changes in metabolic flux

We also examined changes in measured intracellular fluxes of *E. coli* under different growth conditions. Our major observation was that there was little change in flux ratios throughout growth, and for most of the experiment this initial labeling remained. There were changes at two weeks in one of the apparent flux ratios, however, given that there is not expected to be any net synthesis of amino acids after growth ceased, this likely represents either internal amino acid recycling or some de novo amino acid synthesis from recycling nutrients released by dead cells occurred after one week. We also determined metabolic fluxes through central metabolism during the high sodium and high magnesium experiments of that paper. We saw no significant changes in flux ratios with increasing Na+ or increasing/decreasing Mg2+ concentrations. These data collectively indicate that the relative flow of carbon through central metabolism was remarkably robust despite changes in growth phase or environmental stresses tested.