

Discovering drug targets via network analysis

Mario L. Arrieta-Ortiz
Institute for Systems Biology
Baliga Laboratory

Overview

- Introduction to network analysis
- Tolerance networks of MTB to bedaquiline
- R code overview

Networks are everywhere

Social interactions



Global communication



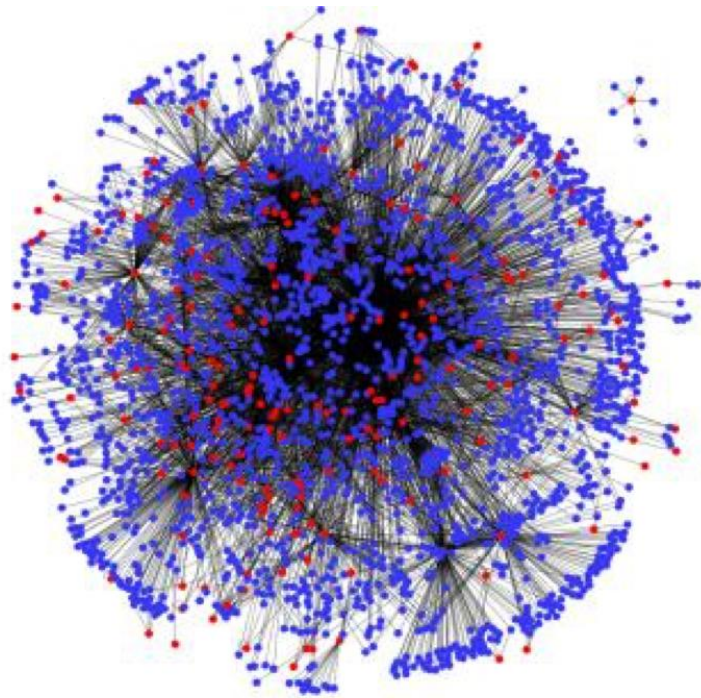
Transportation



Networks can be represented as graphs: nodes, edges

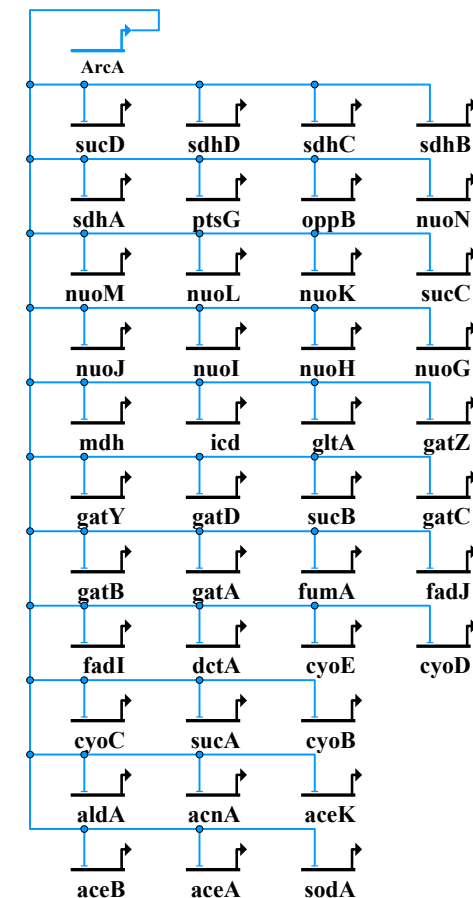
Interactions between transcriptional regulators (TFs) and their target genes are represented as networks

Transcriptional regulatory network (TRN)



TFs
Targets

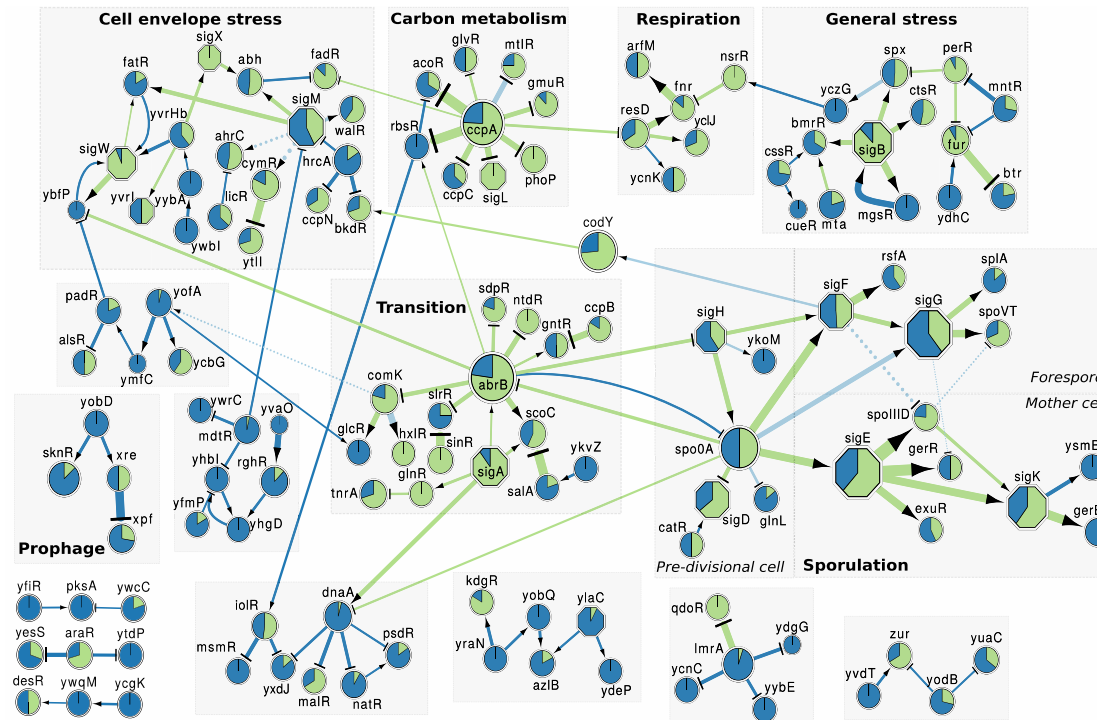
E. coli's ArcA regulon



Environmental and Gene Regulatory Influence Network (EGRIN)

5

- Global transcriptional networks are static representations and do not offer information about conditional regulation
- Biological networks are modular

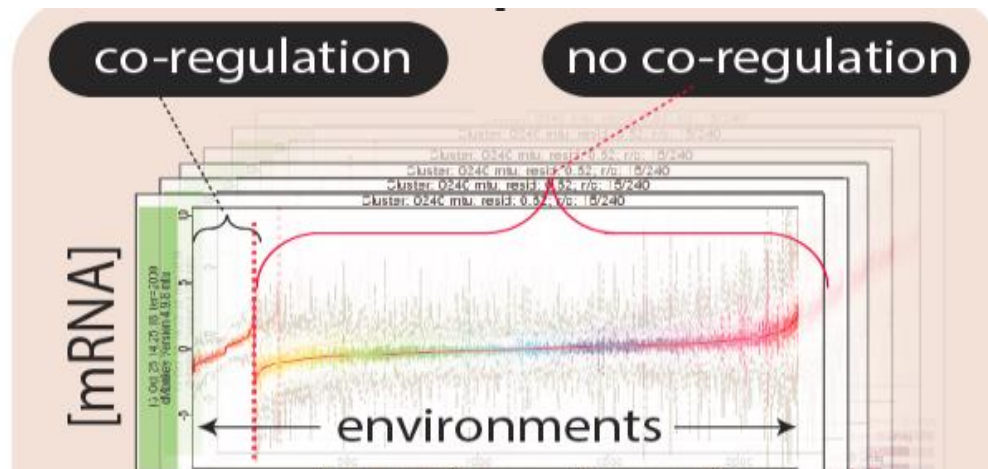


Arrieta-Ortiz, Hafemeister, et al., MSB (2015)

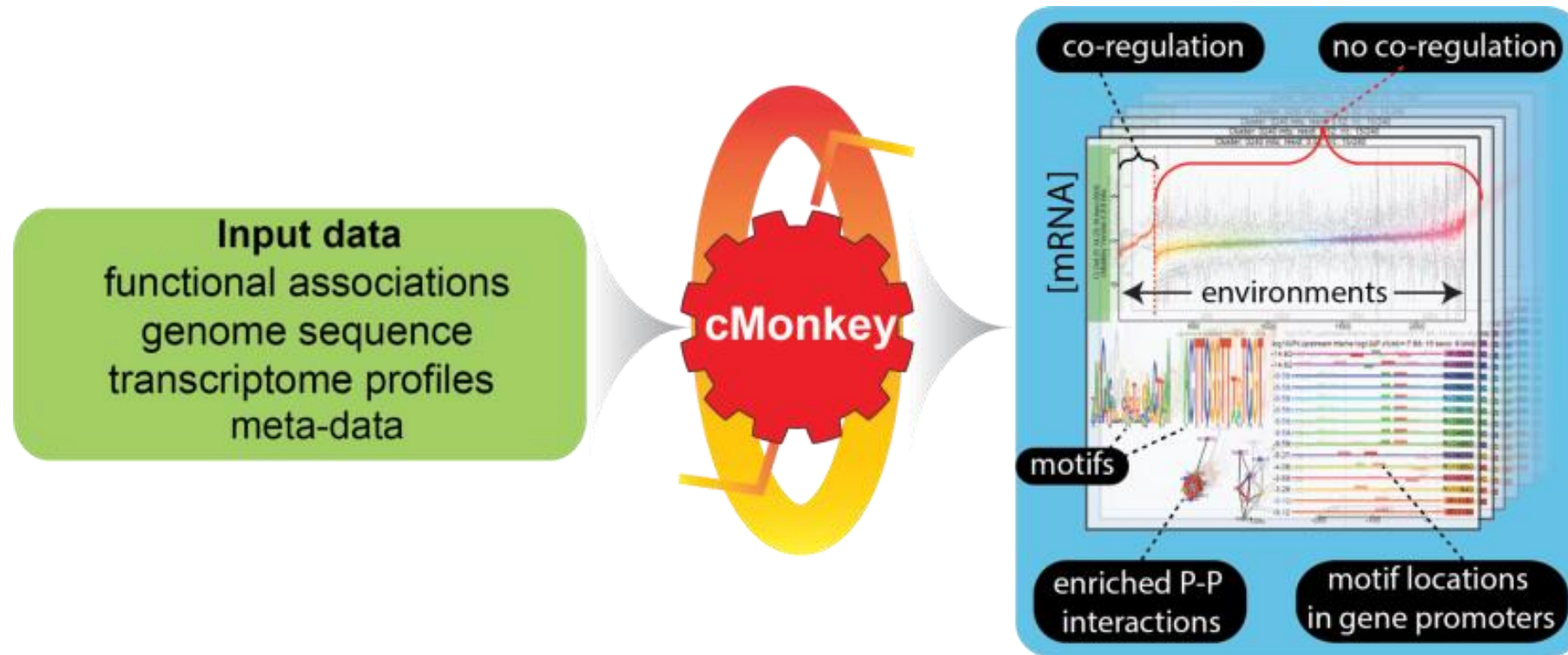
Environmental and Gene Regulatory Influence Network (EGRIN)

6

- Global transcriptional networks are static representations and do not offer information about conditional regulation
- Biological networks are modular
- EGRIN approach is based on bi-clustering (genes and conditions)
- Each bicluster represents a group of co-regulated genes in a specific set of conditions

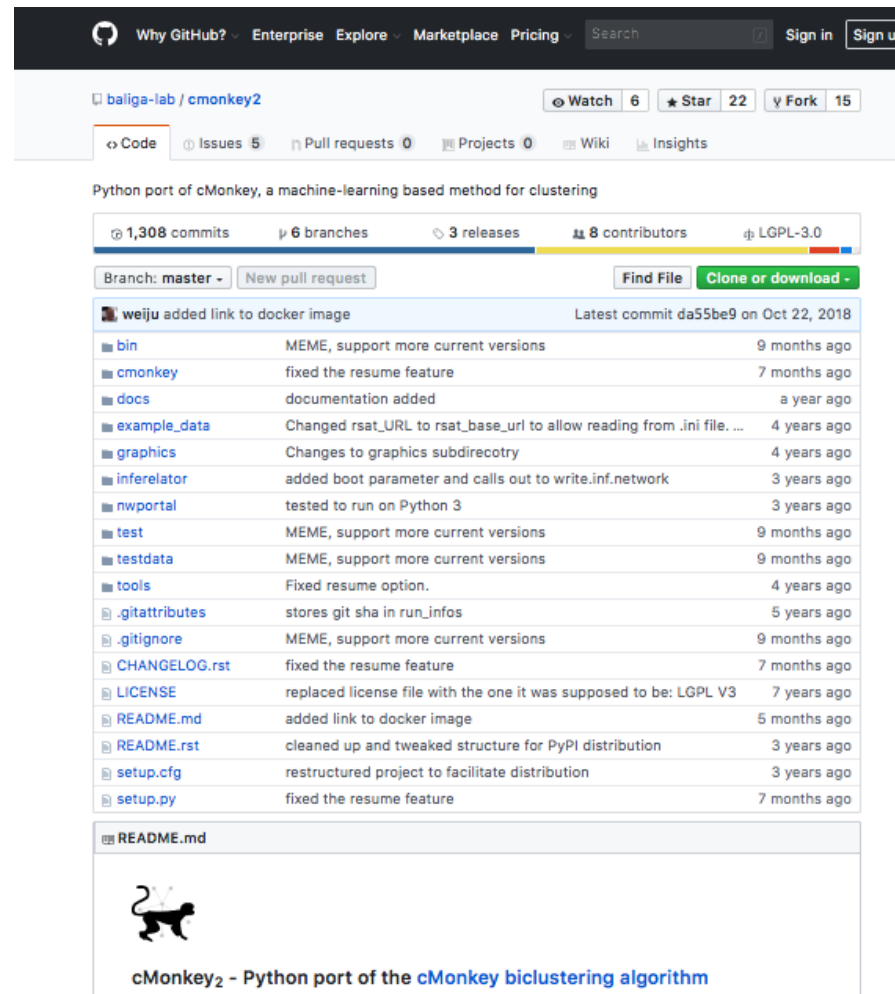


Generating an EGRIN model with cMonkey



cMonkey code availability

- <https://github.com/baliga-lab/cmonkey2> (Python version)




Python port of cMonkey, a machine-learning based method for clustering

1,308 commits · 6 branches · 3 releases · 8 contributors · LGPL-3.0

Branch: master · New pull request · Find File · Clone or download

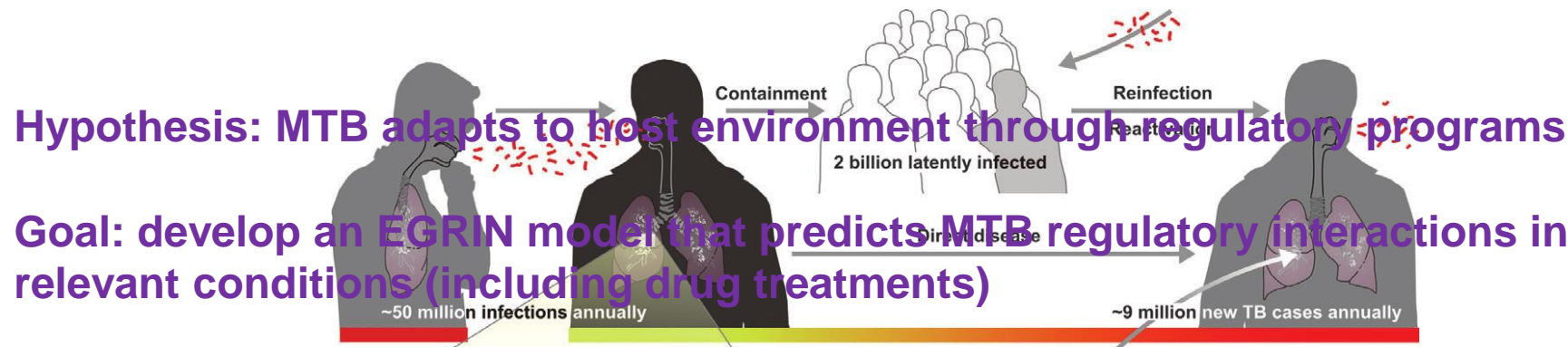
File/Folder	Description	Time Ago
bin	MEME, support more current versions	9 months ago
cmonkey	fixed the resume feature	7 months ago
docs	documentation added	a year ago
example_data	Changed rsat_URL to rsat_base_url to allow reading from .ini file. ...	4 years ago
graphics	Changes to graphics subdirecotry	4 years ago
inferelator	added boot parameter and calls out to write.inf.network	3 years ago
nwportal	tested to run on Python 3	3 years ago
test	MEME, support more current versions	9 months ago
testdata	MEME, support more current versions	9 months ago
tools	Fixed resume option.	4 years ago
.gitattributes	stores git sha in run_infos	5 years ago
.gitignore	MEME, support more current versions	9 months ago
CHANGELOG.rst	fixed the resume feature	7 months ago
LICENSE	replaced license file with the one it was supposed to be: LGPL V3	7 years ago
README.md	added link to docker image	5 months ago
README.rst	cleaned up and tweaked structure for PyPI distribution	3 years ago
setup.cfg	restructured project to facilitate distribution	3 years ago
setup.py	fixed the resume feature	7 months ago

README.md

 cMonkey₂ - Python port of the cMonkey biclustering algorithm

The burden of tuberculosis (TB)

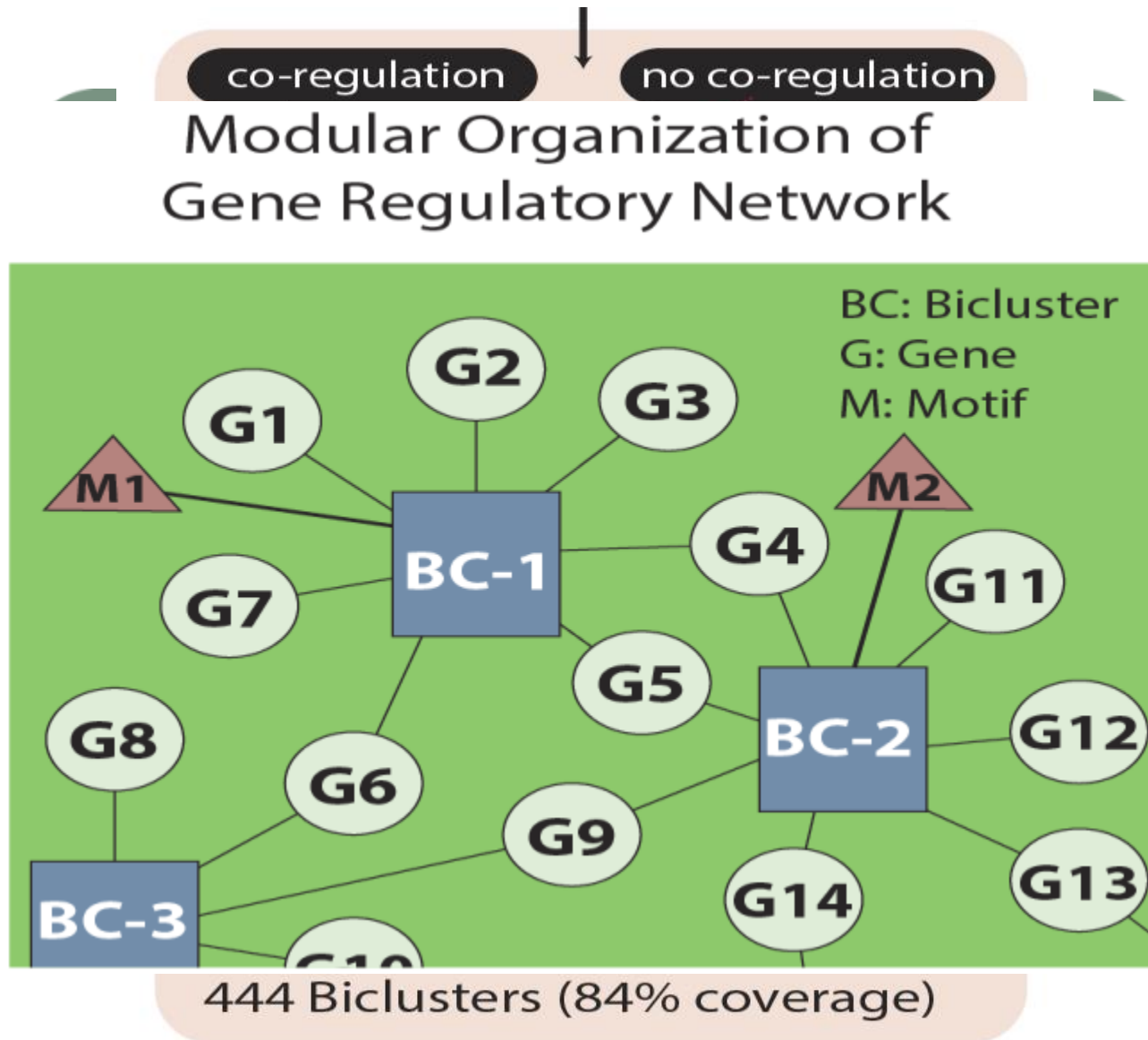
- #1 infection disease in terms of number of annual deaths (1.5 M)
- In need of new and shorter treatments (standard treatment includes four drugs for six months)
- Antimicrobial resistance on the rise → limited and longer therapy options
- *Mycobacterium tuberculosis* (MTB), the causal agent of TB, adapts to host's environment (immune response, low O₂, toxic compounds, nutrient depletion, low pH)



Adapted from Gengenbacher & Kaufmann, 2012

Construction of *Mycobacterium tuberculosis*'s EGRIN model

10



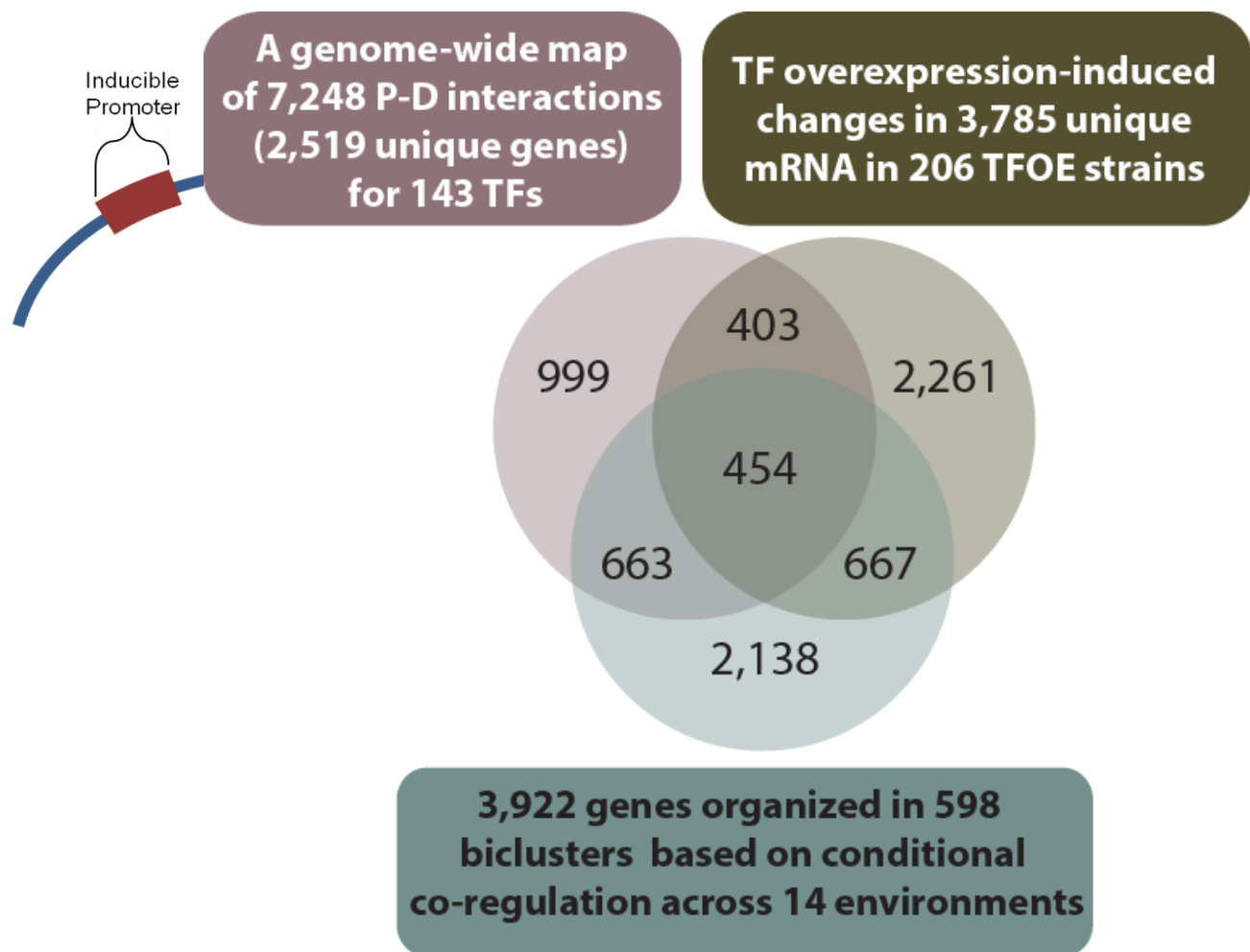
Biclusters enriched with a given set of genes are identified with hypergeometric tests

- Hypergeometric test evaluates whether the observed overlap between two groups is statistically significant

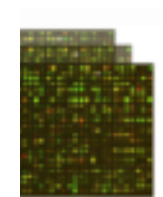
Number of successes, drawn WITHOUT replacement, from a source that contains a certain number of successes and a certain number of failures

Genes	Differentially expressed?	In functional category "Metabolism/Respiration"?
'Rv0001'	0	0
'Rv0002'	1	0
'Rv0003'	1	0
'Rv0004'	0	1
'Rv0005'	0	0

Validation of MTB's EGRIN model



torial
g

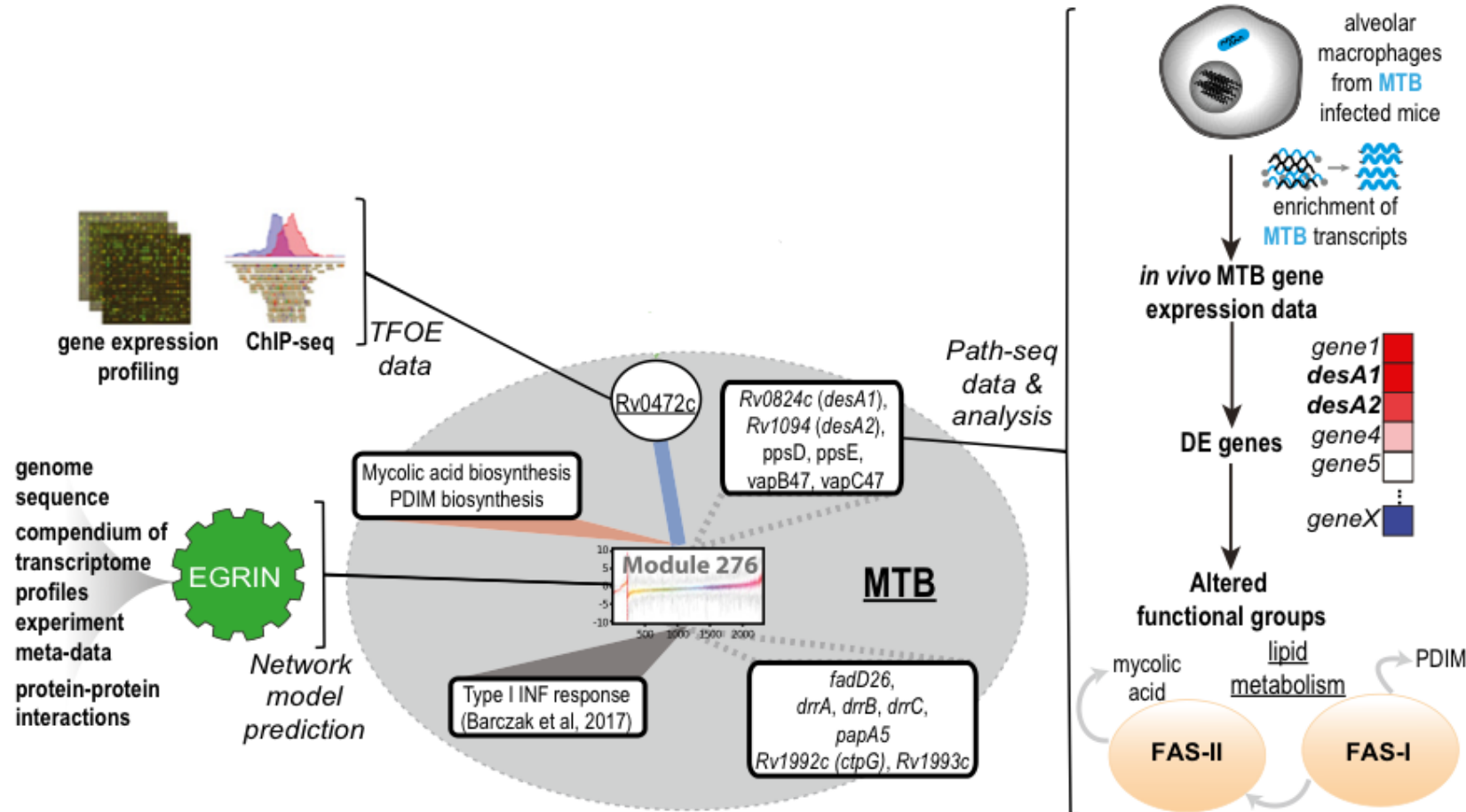


ion profiling by
icroarray

Peterson et al., NAR (2014)
Minch et al., Nat Comm (2014)
Rustad et al., Genome Biol (2014)
Turkarslan et al., Nature Sci Data (2015)

MTB EGRIN model offered insights about the regulation of mycolic acid biosynthesis genes in intracellular conditions

13



Peterson et al., MSB (2019)

Case study: network analysis to accelerate TB drug synergy and to elucidate mechanisms of action

14

nature
microbiology

ARTICLES

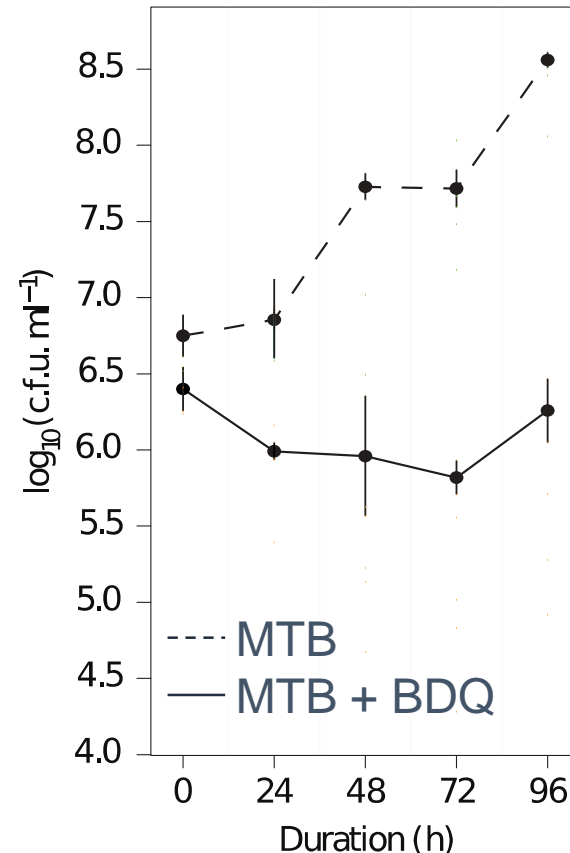
PUBLISHED: 6 JUNE 2016 | ARTICLE NUMBER: 16078 | DOI: 10.1038/NMICROBIOL.2016.78

Network analysis identifies Rv0324 and Rv0880 as regulators of bedaquiline tolerance in *Mycobacterium tuberculosis*

Eliza J. R. Peterson^{1†}, Shuyi Ma^{2†}, David R. Sherman^{2,3} and Nitin S. Baliga^{1,4,5★}

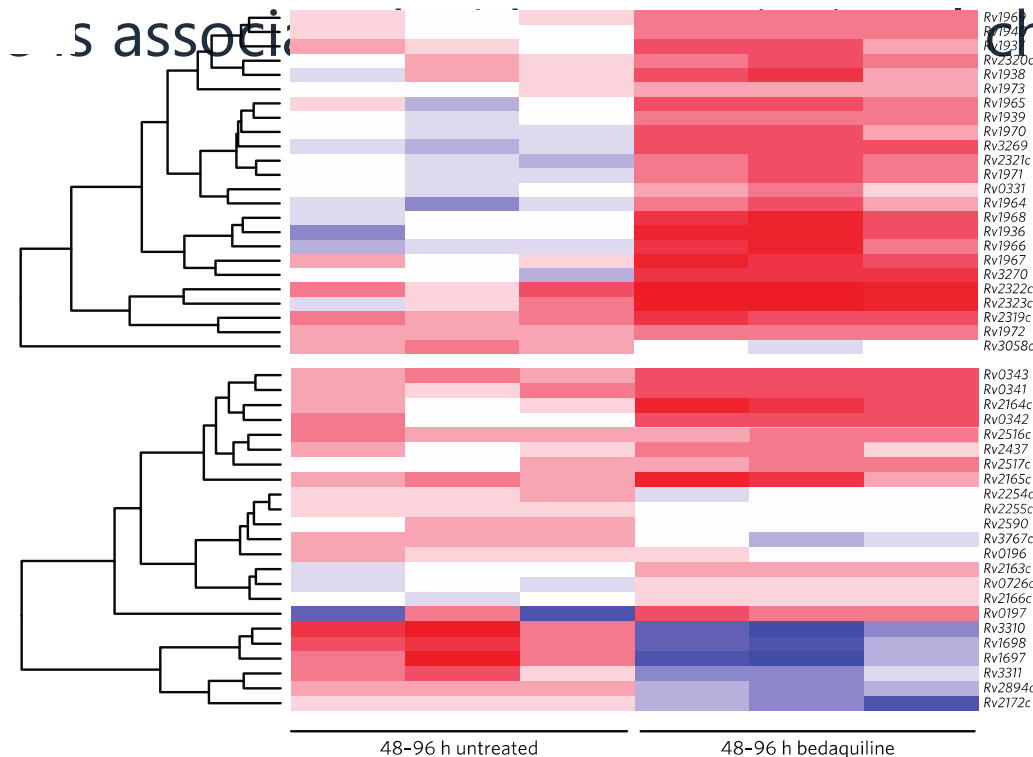
Modelling the transcriptional response of bedaquiline (BDQ)

- First anti-TB drug approved by FDA in nearly 40 years
- Kills relatively slow
- Regulatory mechanism pushes MTB into a tolerant state that resists BDQ killing

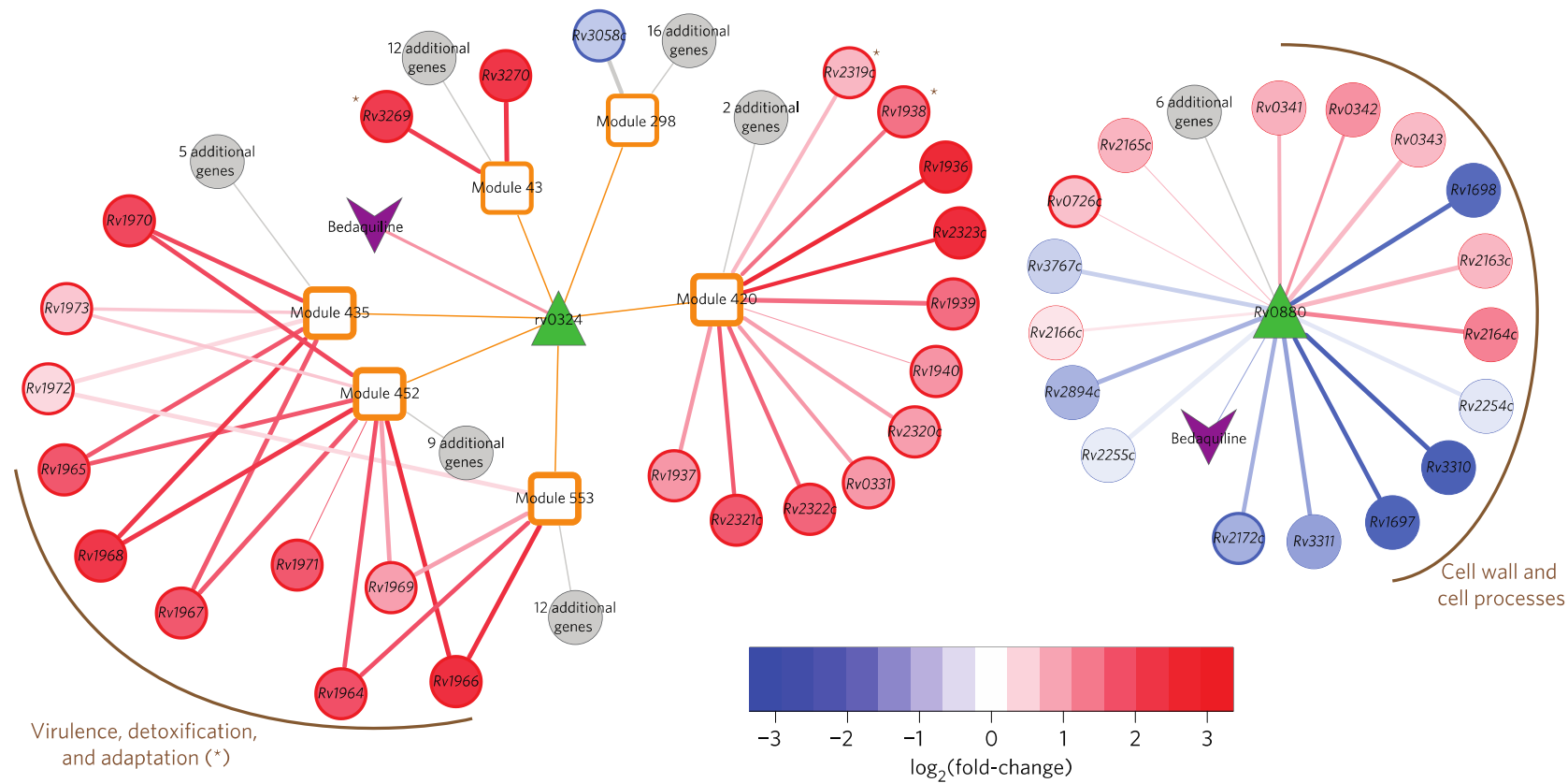


Modelling the transcriptional response of bedaquiline (BDQ)

- First anti-TB drug approved by FDA in nearly 40 years
- Kills relatively slow
- Regulatory mechanism pushes MTB into a tolerant state that resists BDQ killing
- BDQ tolerant state is associated with changes



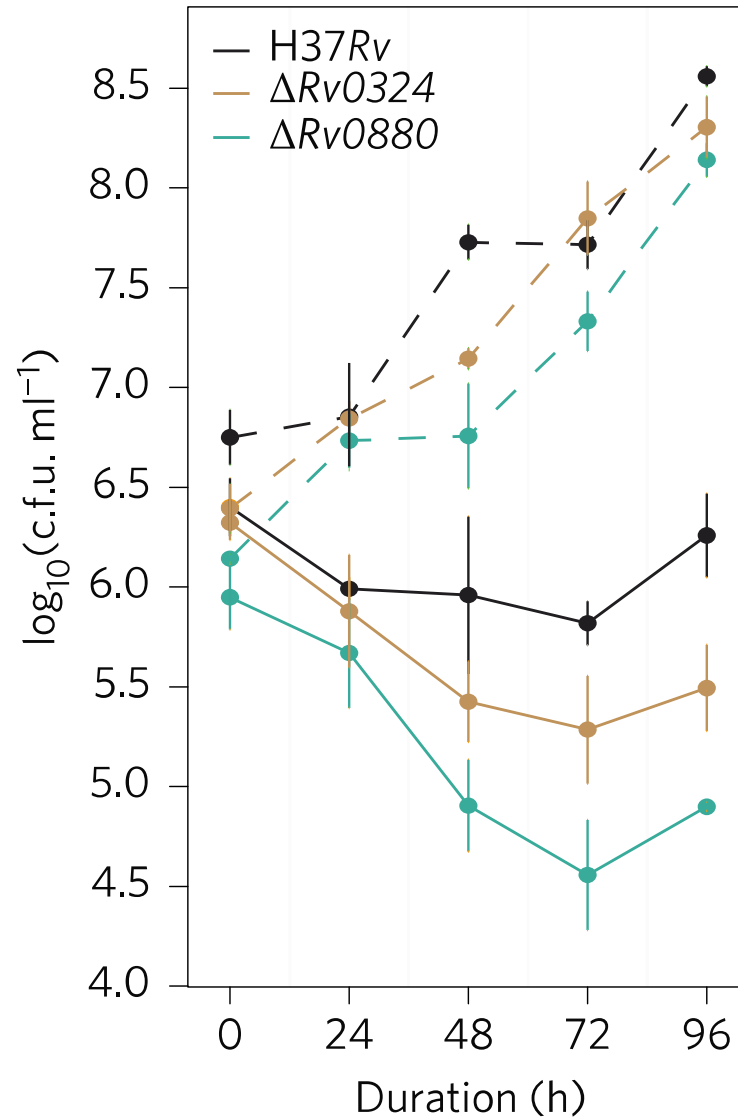
BDQ tolerance networks in MTB are driven by Rv0324 and Rv0880



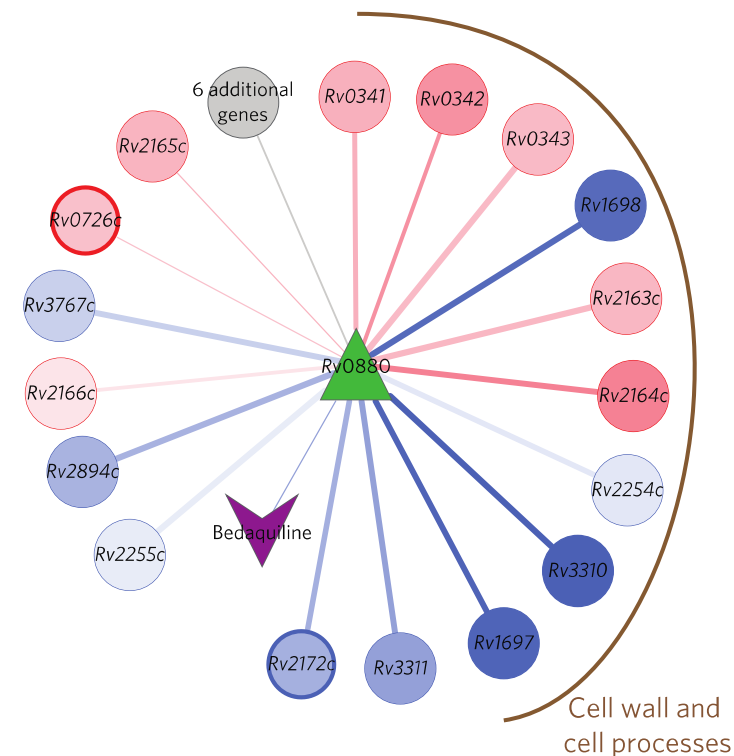
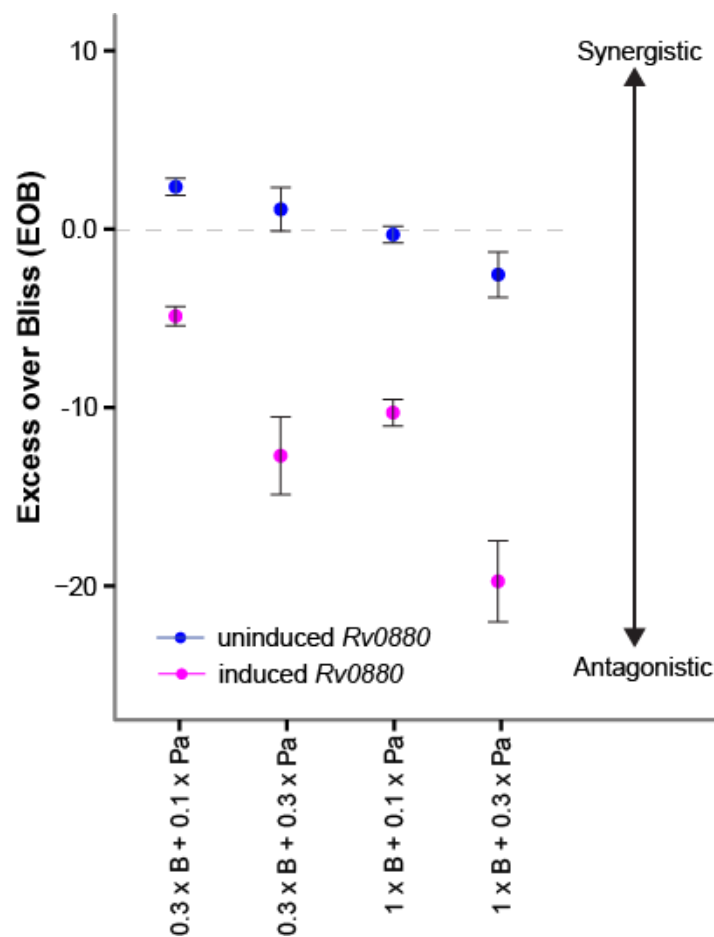
Peterson et al., Nature Micro (2016)

Disrupting tolerance networks increase killing by BDQ

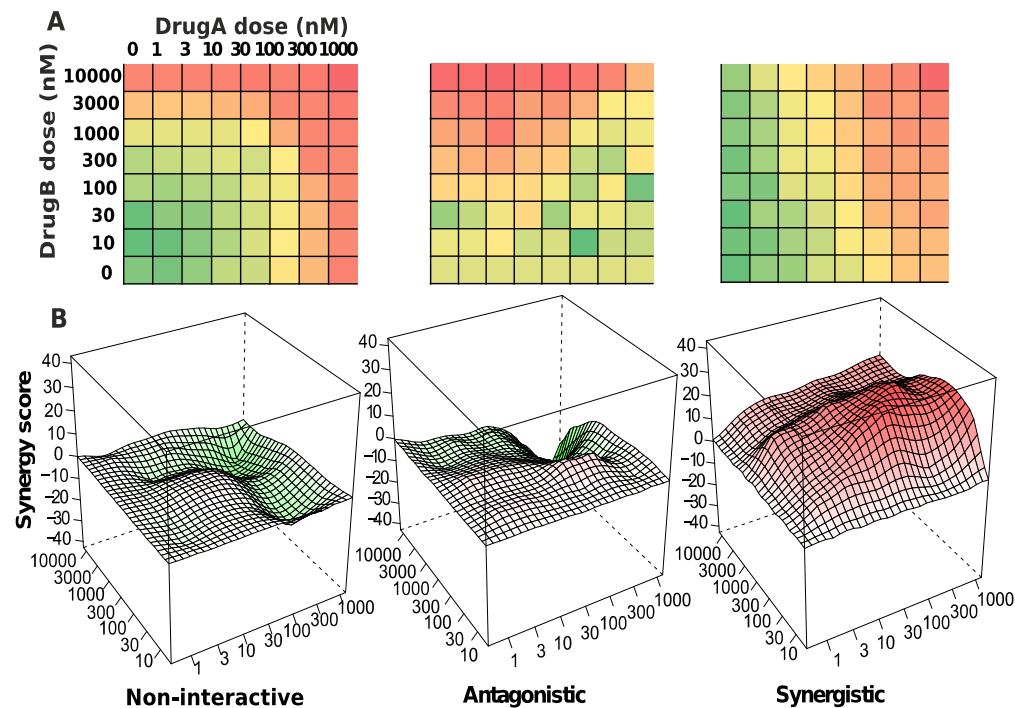
18



Pretomanid disrupts the tolerance network to potentiate killing by BDQ



Analyzing Drug Combinations- Drug synergy scoring



- HAS: expected effect is the highest monotherapy effect ($y_e = \max(y_1, y_2)$)
- Loewe: expected effect as if a drug is combined with itself ($y_e = y_1(x_1 + x_2) = y_2(x_1 + x_2)$)
- Bliss: expected effect as if the two drugs are acting independently ($y_e = y_1 + y_2 - y_1y_2$)
- ZIP: both the assumptions of the Loewe and Bliss models are met

R notebook overview

- Analyze and plot bedaquiline treated expression data
- Functional enrichment of differentially expressed genes
 - Hypergeometric test
- Network analysis of differentially expressed genes
- Identification of secondary drug to synergize with Bedaquiline
 - Analyze drug combination - synergyfinder

Datasets

Dataset	Description	Dataset	Description
wt_survival.csv	Experimental kill curve of wild-type MTB with and without Bedaquiline treatment (CFUs vs time)	Drug_treated_expression.csv	Differential expression of MTB treated with 35 different drugs, compared to untreated after 18 h.
BDQ.genes.ttest.p.value.ALL.csv	Differential expression of MTB treated with BDQ, compared to untreated after 48-96 h. Triplicate data used to calculate t-test p-value.	Drug_treated_pheno.csv	Description of replicates for Drug_treated_expression.csv
full_annotation.csv	Tuberculist categories for functional enrichment (http://tuberculist.epfl.ch/index.html)	Dose_response_matrix_1.csv	Inhibition of MTB growth at different concentrations and combinations of Bedaquiline and IMTB028
biclustergenes.csv	Sets of co-regulated genes, modeled into biclusters using cMonkey	Option 1: TFs_over_expression.csv	Expression data of the induced overexpression of 3 TFs. Triplicate data used to calculate t-test p-value.
Filtered.biclustergenes.444.fullname.csv	Biclusters filtered for mean residual < 0.55. The residual is a metric for gene co-expression among bicluster genes.	Option 1: rv324_ko_survival.csv	Experimental kill curve of wild-type MTB and Rv0324 knockout with and without Bedaquiline treatment (CFUs vs time)
Biclustergenes.TF.Overlap.csv	TFs predicted to regulated bicluster genes, based on TF overexpression (expression & ChIP-seq)	Option 2: Dose_response_matrix_2.csv	Inhibition of MTB growth at different concentrations and combinations of Bedaquiline and Isoniazid

Acknowledgements

Baliga Lab

- Nitin Baliga



- Eliza Peterson



- Chris Plaisier

Funding



National Institute of
Allergy and
Infectious Diseases

ISB Summer Course



<https://systemsbiology.org/course/#anchor2>

Collaborators



Seattle Children's
HOSPITAL • RESEARCH • FOUNDATION