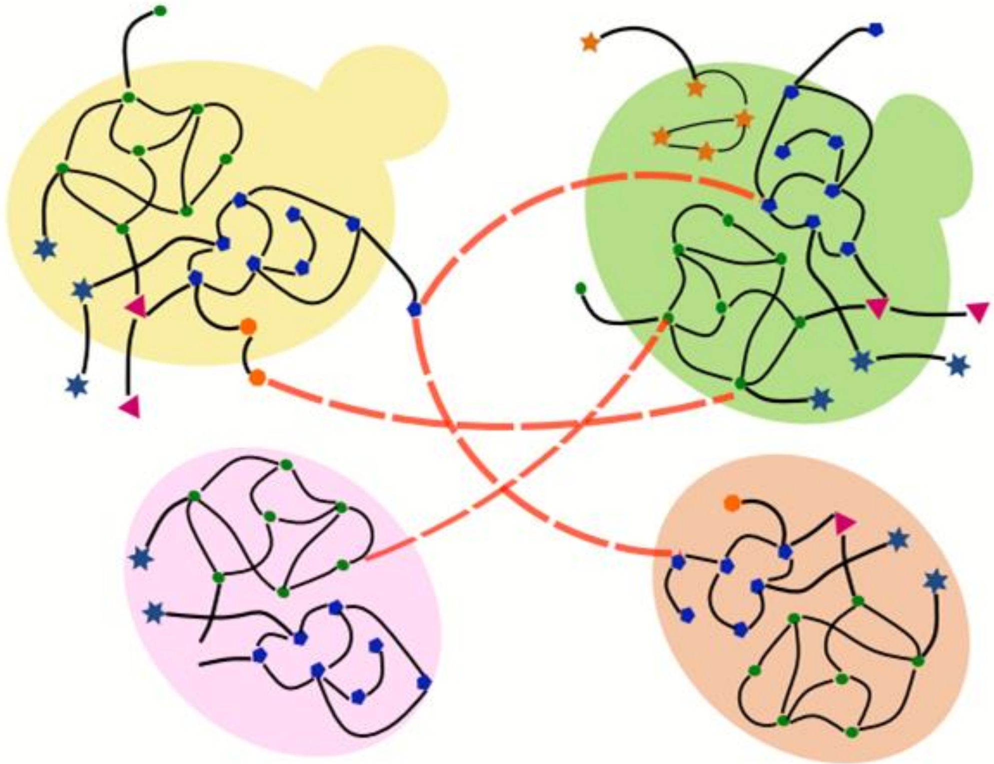
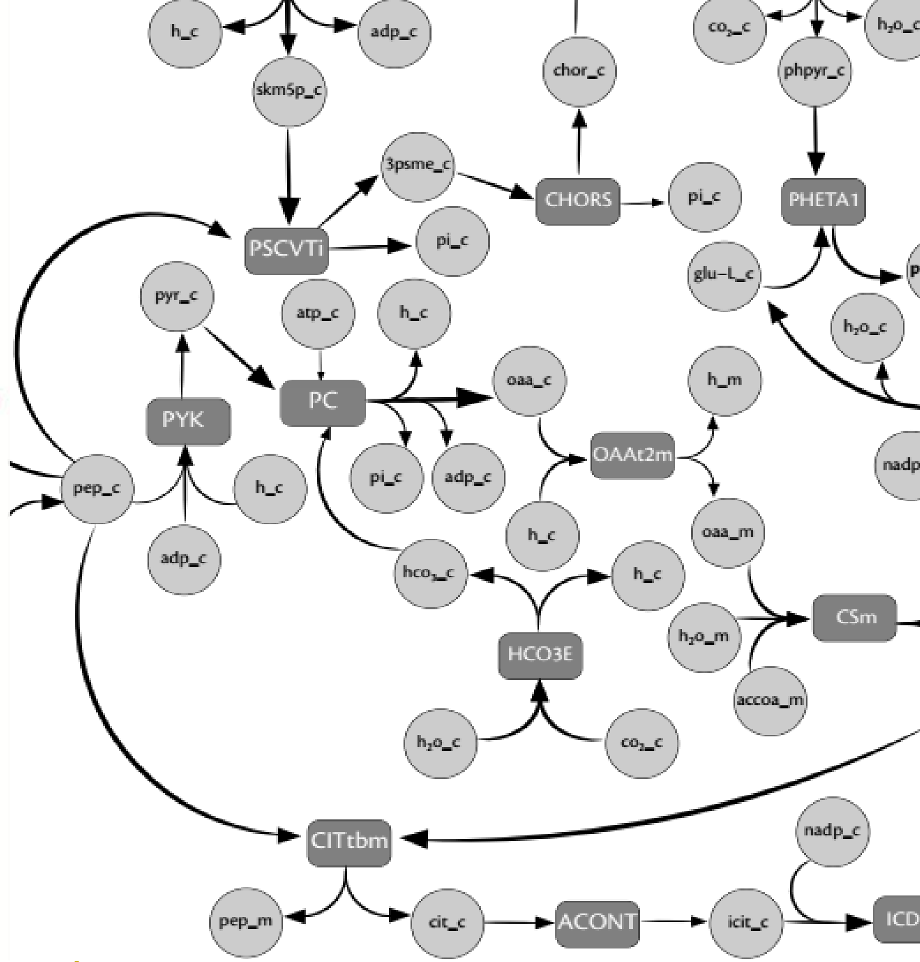




UNRAVELING MICROBIAL INTERACTIONS IN THE GUT MICROBIOME ASSOCIATED WITH ANTIBIOTIC RECOVERY

Karthik Raman

Bhupat and Jyoti Mehta School of Biosciences, IBSE & RBCDSAI

IIT MADRAS

INTRODUCTION

- Human gut is a complex ecosystem
- Many roles in health and disease
- Composition is highly variable
  - Stage of life
  - Diet
  - Environmental exposure
  - **Antibiotic usage**
- Antibiotic usage is a major cause of gut dysbiosis
  - Collateral damage to gut microflora
  - Induce changes in composition
  - Organisms develop resistance

INTRODUCTION

- Gut microbiome recovers post antibiotic treatment
- How long does the recovery take?
  - Varies from individual to individual
- Specific groups of organisms accelerate recovery
- Recovery Associated Bacteria (RABs)[1]
  - 20 species identified
  - Improved carbohydrate degrading capacity
  - Specific synergy between *Bacteroides thetaiotaomicron* and *Bifidobacterium adolescentis*
- Why do these organisms work well together?

[1]**Nandi et al** (2018) *bioRxiv doi:10.1101/350470*

UNRAVELLING THE COMPLEXITY OF MICROBIAL INTERACTIONS IN THE GUT

A METABOLIC PERSPECTIVE

GENOME-SCALE METABOLIC NETWORKS

- Have been reconstructed for many organisms
- Present a comprehensive picture of known metabolic reactions / transports happening in a cell
- ‘Draft’ reconstructions are readily obtained from genome sequence/databases like ModelSEED
- Many methods exist to analyse these networks

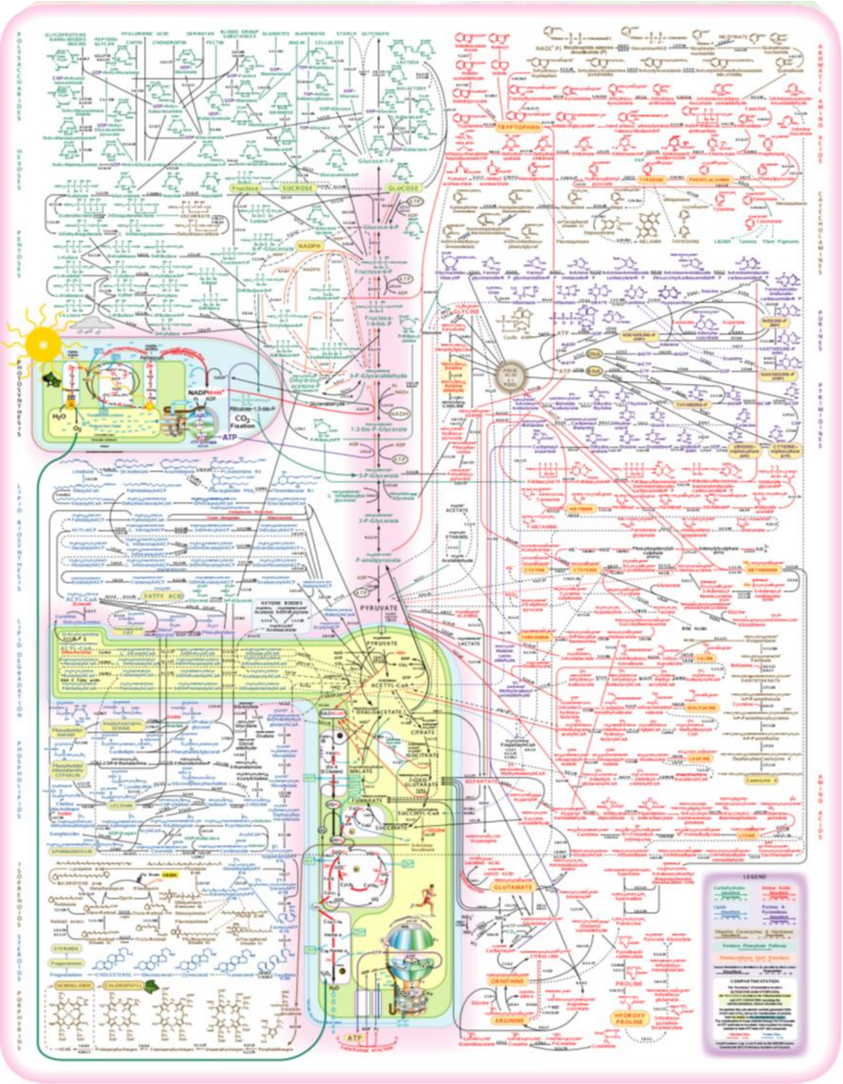
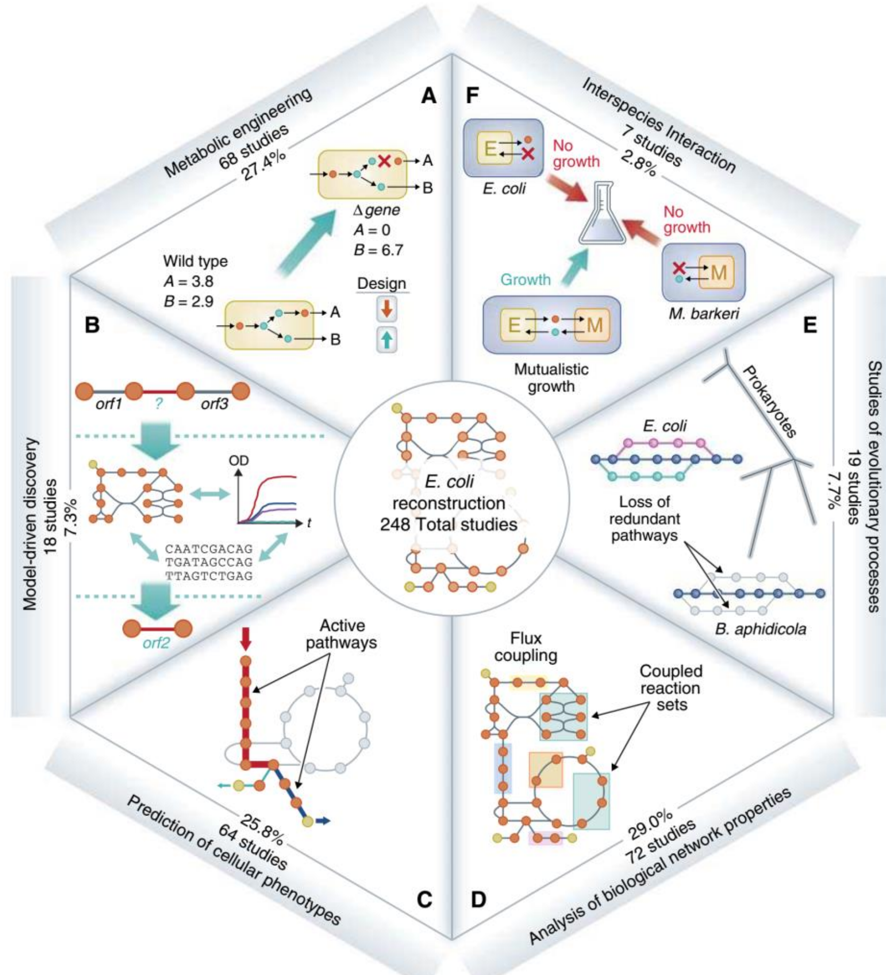


Image Courtesy of Sigma Aldrich, Order your copy from sigma-aldrich.com/metpath!

WHAT CAN GENOME-SCALE METABOLIC MODELS TELL US?

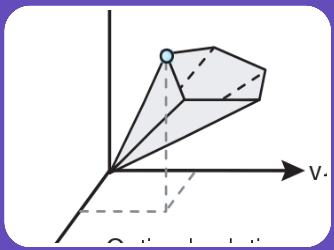
- Analysis of biological network properties
- Metabolic engineering[1]
- Prediction of cellular phenotypes
- Model-driven (biological knowledge) discovery
- Studies of evolutionary processes
- Interspecies interactions[2]



McCloskey D *et al* (2013) *Mol Syst Biol* **9**:661

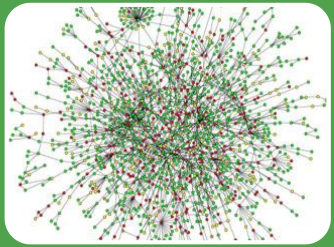
[1]**Badri A, Srinivasan A & Raman K** (2017)
In silico approaches to metabolic engineering
ISBN 978-0-444-63667-6 pp. 161-200

HOW TO ANALYSE GENOME-SCALE METABOLIC NETWORKS?



Constraint-based Modelling

- Popular for applications such as metabolic engineering
- Demands well-curated models



Network-based (Graph-based) Modelling

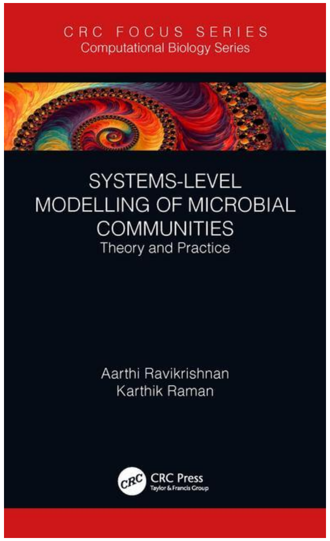
- Path-finding in metabolic networks
- Predicting ‘new’ pathways based on atom-atom mapping/reaction ‘rules’

Need Methods that

- are very scalable and accurate
- can figure all possible routes that exist

MODELLING MICROBIAL COMMUNITIES

- Again, many methods to model
  - Constraint-based
  - Graph-based
  - Population-based
  - Agent-based
- More methods being developed
- Many challenges[1]
  - Mostly draft reconstructions available
  - Difficult to make models talk to one another
- **Metabolic interactions/exchanges in communities are of particular interest**
- **Also has applications in understanding other communities, e.g. gut microbiome**



Ravikrishnan & Raman (2018)
ISBN: 978-113859671-9

[1]**Ravikrishnan A & Raman K** (2015) *Briefings in Bioinformatics* **16**:1057–1068

PATH-FINDING IN METABOLIC NETWORKS

CURRENT STATE-OF-THE-ART

- Rahnuma: Hypergraph-based method that performs DFS on hypergraph to find routes
- FMM: Constructs metabolic pathways between metabolites using substrate graph representation
- PathPred: Generates the pathways based on the structure transformation patterns and its comparison with reference pathway
- MetaPath: Calculates the *scope* of metabolic networks given a set of starting *seed*
- ATLAS: Finds possible transformations between two metabolites using reactions from KEGG and other (predicted) reactions specific to ATLAS
- Metabolic Route Explorer (MRE): Provides organism specific data from KEGG online tool for heterologous biosynthesis pathway design
- **These algorithms/methods are based on different heuristics, and aim to infer/predict the routes of conversions from source to the target molecules**
- Many of these methods are no longer available (broken link etc.) or do not scale well

GRAPH REPRESENTATIONS OF METABOLIC NETWORKS

- How to convert a metabolic network to a graph?
  - Substrate graph
    - Nodes: Metabolites
    - Edges connect metabolites participating in the same reaction / reactants to products
  - Reaction graph
    - Nodes: Reactions
    - Edges connect reactions sharing metabolites
  - Bi-partite graph / Hypergraphs
    - Nodes: two sets — metabolites and reactions
    - Edges: connect reactants to reaction nodes and reaction nodes to product nodes
    - No metabolite–metabolite or reaction–reaction links
- 'Currency' metabolites
  - Need to be eliminated from substrate graphs!
    - Else, we have a two-step glycolysis!

OUR ALGORITHM: METQUEST

Ravikrishnan, Nasre & Raman (2018) *Scientific Reports* **8**:9932

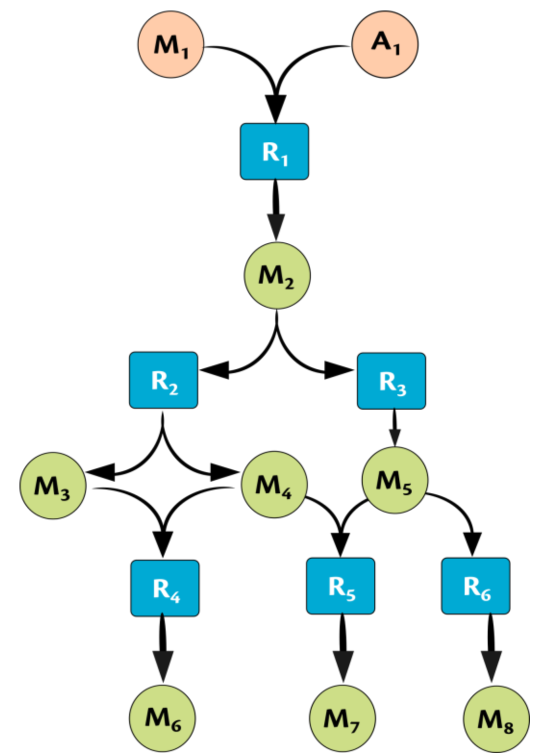
/RamanLab/MetQuest

MetQuest: Overview

- Novel dynamic-programming based enumeration, which assembles reactions into pathways of a specified size producing a given target from a given set of source molecules
- Employs two phases
  - Guided Breadth First Search (BFS)
  - Assembly of reactions into pathways
- Implemented on Python 3.6 & Python 2.7
- Key Features
  - Requires only the topology of reaction network (rather than stoichiometry / atom mapping)
  - Simple and scalable to large metabolic networks (especially those comprising >1 organism)
  - Efficiently handles cyclic and branched pathways
  - Examines multiple alternate routes of conversion

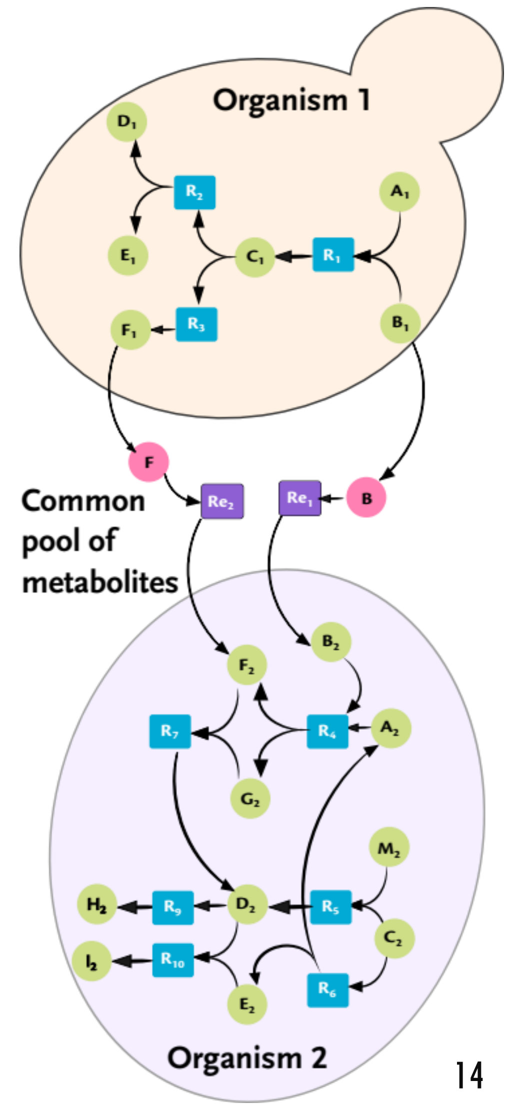
INPUT REPRESENTATION: BIPARTITE GRAPHS

- Any given metabolic network can be represented as a directed bipartite graph *G(M, R, E)*
  - *M* is the set of metabolites, *R* is the set of reactions, and *E* is the set of edges
- Directed edges connect metabolites ($m_i \in M$) to a reaction node ($r_i \in R$) or a reaction node to product metabolites
- Reversible reactions in the network are denoted by two separate reaction identifiers — representing the forward and reverse reactions, respectively
- Bipartite representations disallow invalid conversions as may be interpreted from substrate graphs and
- Help in generating valid paths with biologically meaningful conversions



HANDLING COMMUNITY METABOLIC NETWORKS

- Directed bipartite graph **G** of microbial communities (consisting of more than one metabolic network) are also easily constructed
- By connecting the graphs of individual organisms through a common extracellular medium, based on the overlapping set of exchange reactions
- The non-common exchange reactions are connected only to the extracellular environment



MetQuest: Inputs to the Algorithm

- Input to MetQuest
  - a directed bipartite graph G derived from a given metabolic network
  - a set of seed metabolites, S
  - a set of target metabolites, T
  - an integer β which bounds the size of any pathway generated

- Seed Metabolites
  - include the source metabolite(s)
  - as well as molecules such as co-factors and co-enzymes — commonly present in any cell
  - akin to a "medium" for growth

DEFINITIONS

- Reachable metabolite *m*
  - A metabolite *m* is reachable from a set *S* if either *m* is in the set *S* or there is a reaction *r* in the reaction network whose output is m and every input of *r* is producible
- Branched pathway producing *m*
  - An *S*-to-*m* pathway R' is a set of reactions such that *m* is the output of at least one reaction in R' and every input of every reaction in R' is producible from *S*
- Cyclic pathway producing *m*
  - A cyclic pathway R', from *S* to m is a set of reactions where *m*, which is the output of at least one reaction in R' is used in its own production by another reaction in R'
- Size of a pathway
  - It is the cardinality/number of reactions in the set R'

ALGORITHM WALKTHROUGH

PHASE 1: GUIDED BFS

"GUIDED" BFS

- BFS is a classic graph traversal technique that visits all the nodes of a given graph, starting at a source node, in a breadth-first fashion
- BFS employs a queue of vertices, where newly discovered vertices are enqueued, to be processed at a later stage
- We modify the standard BFS by *guiding* it, based on the availability of precursor metabolites
- Starting with the set of seed metabolites *S*, the algorithm first finds all the reactions from the set R, whose precursor metabolites are in *S*
- Such reactions are marked "visited" and added to the *visited reaction set* R_v
- The metabolites produced by these reactions, m_c, are then added to *S*
- The traversal continues in a breadth-first manner, incrementally adding *triggerable reactions* to the BFS queue

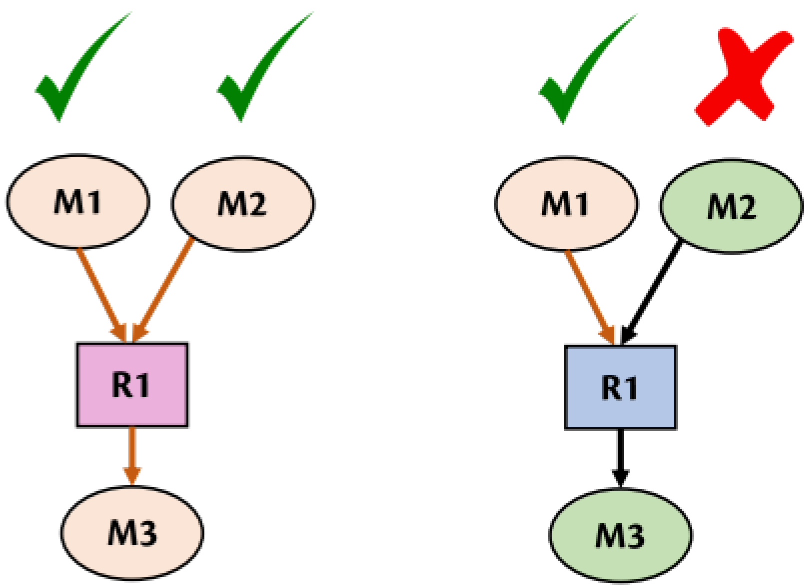
"GUIDED" BFS

- The expansion stops when there are no further reactions that can be visited; during the expansion, a reaction node is labelled as stuck, if it does not (yet) have the necessary precursors in S
- Such reactions are automatically triggered if the precursor metabolites are produced at any later stage
- The traversed graph consists of all reactions that can be visited
- At the end of the traversal, we obtain the scope $M_s \supseteq S$ and the set of visited reaction nodes R_v
- This process of graph traversal resembles the ideas of network expansion[1], and forward propagation[2] reported earlier
- However, we make a systematic note of the visited and stuck reaction nodes — later exploited for efficient and exhaustive enumeration

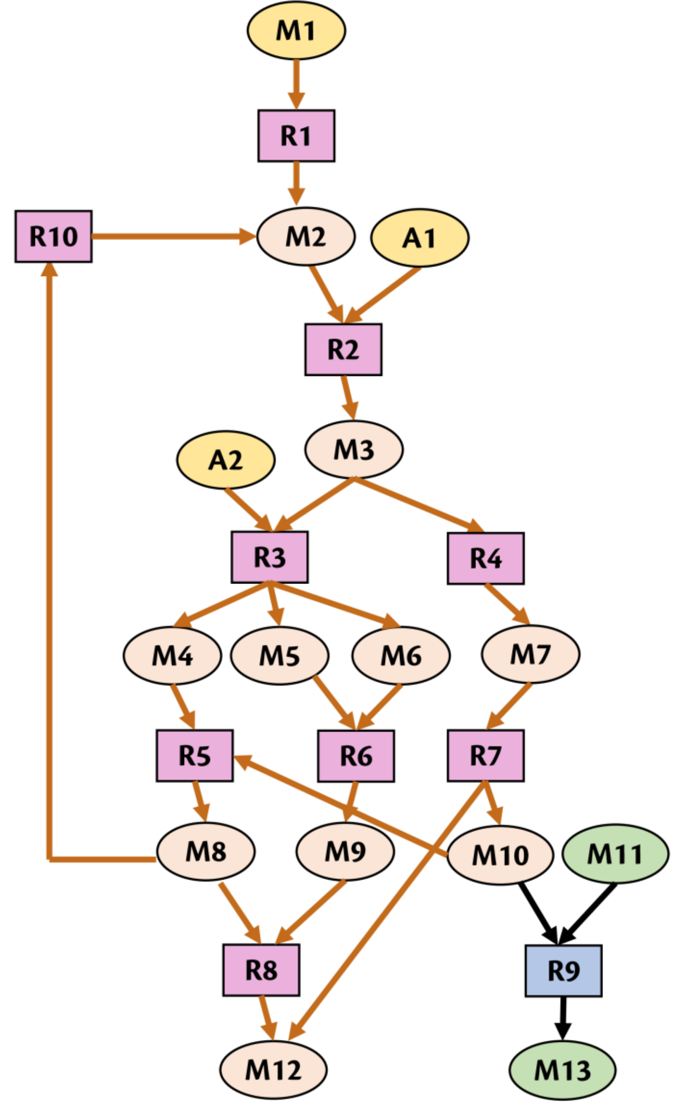
[1]**Handorf *et al*** (2005) *J Mol Evol* **61**:498
[2]**Acuña *et al*** (2012) *Bioinformatics* **28**:2474

GUIDED BFS: WALKTHROUGH

- Input – Directed bipartite graph ***G*** derived from metabolic network(s), *seed* metabolites
- Output – *Scope* of metabolites, Reaction set that can be *visited*



Visited   Stuck



Scope metabolite set – {M1, A1, A2, M2, M3, M4, M5, M6, M7, M9, M10, M12, M8}

Visited reaction set – {R1, R2, R3, R4, R6, R7, R5, R8, R10}

ALGORITHM WALKTHROUGH

PHASE 2: PATHWAY GENERATION

PATHWAY GENERATION

- Generates a large *Table*, of size $|M_s| \times \beta$
  - Enumerating all pathways of size $\leq \beta$
  - For every metabolite in the scope
- Goal of MetQuest: to populate all the entries of this *Table*
- We start filling the table entries by first considering the seed metabolite set S
- For every seed metabolite $m \in S$, the entry in corresponding cell $Table(m, 0) = \emptyset$, indicating that no reaction is required to produce it
- For every metabolite $m \in M_s \setminus S$, the entry $Table[m][0]$ remains as $\perp$
- At the end of the algorithm, for any metabolite $m \in M_s$ and an integer k ($0 \leq k \leq \beta$), the entry $Table[m][k]$ is a set of pathways or $\perp$
- If the entry is not $\perp$, each pathway in the set $Table[m][k]$ is of size k and produces the metabolite m starting from the seed metabolite set
- $Table[m][k] = \perp$ implies that m cannot be produced starting from the seed metabolite set S using exactly k reactions

RESULTS

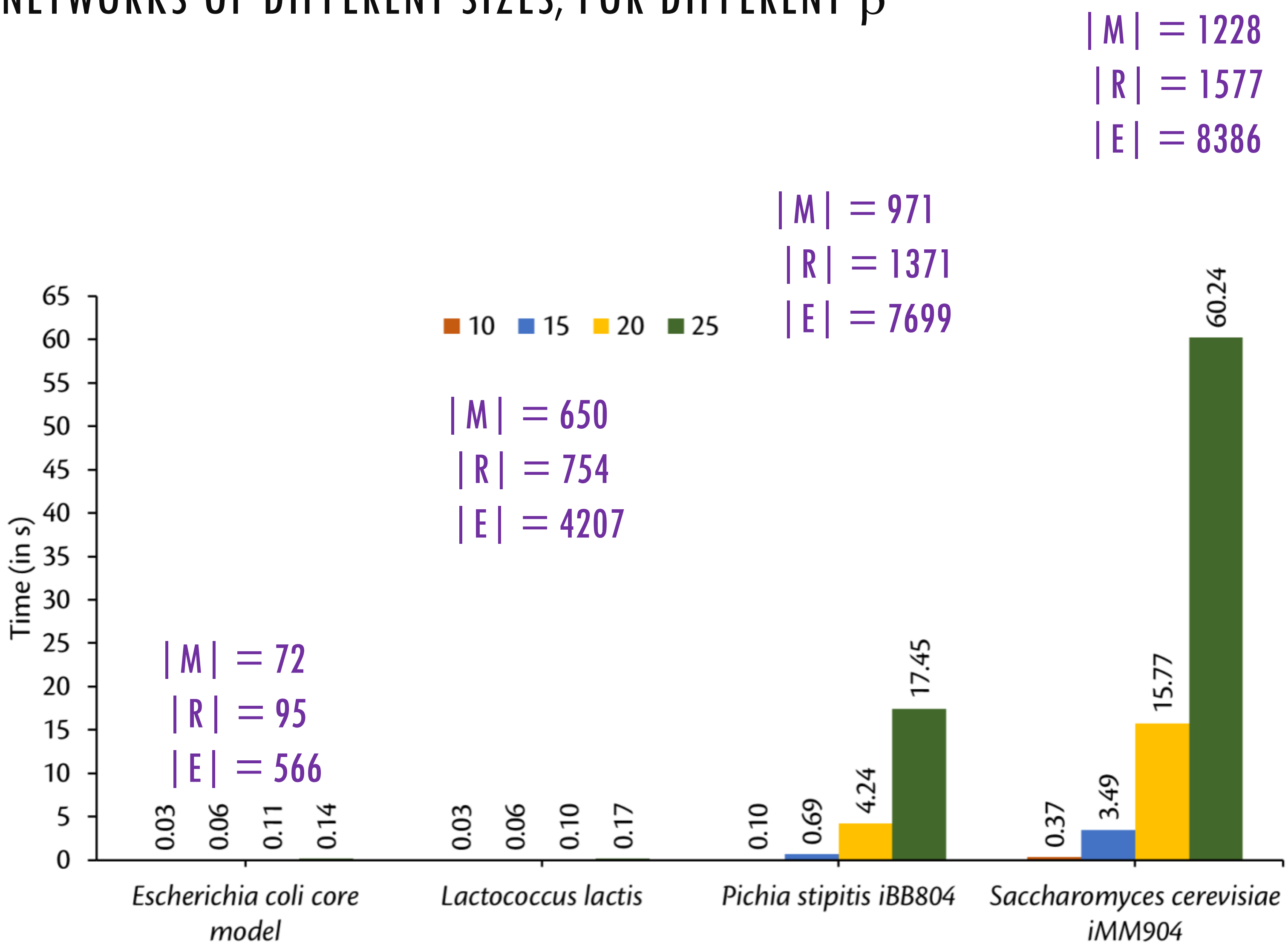
MetQuest excels in comparison with other algorithms

Source	Target	Size	Output sub-network	Comments
L-Arginine (C00062)	L-Citrulline (C00327)	2	R00551, R00665	Matches with ATLAS and FMM
Pyruvate (C00022)	Itaconate (C00490)	4	R02491, R00209, R00237, R02405	Matches with FMM, FMM does not report R00209 which produces C00024 – required by R02405†
Pyruvate (C00022)	Itaconate (C00490)	5	R00351, R02243, R00209, R00217, R01325	Matches with FMM, FMM does not report R00351 which produces C00036 – required by R00351†
L-Tyrosine (C00082)	Naringenin (C00509)	5	R02446, R00737, R01616, R01613, R06641	Matches with FMM, FMM does not report R06641†
L-Phenylalanine (C00079)	Resveratrol (C03582)	5	R01616, R00697, R02253, R06641, R01614	Matches with FMM, FMM does not report R06641 which produces malonyl-CoA required by R01614†
Mevalonic acid (C00418)	Amorpha-4,11-diene (C16028)	7	R01658, R03245, R02245, R01121, R01123, R07630, R02003	No paths found by FMM, ATLAS, however it is natively found in *S. cerevisiae*[51].
D-Erythrose 4-phosphate (C00279)	3-Amino-5-hydroxy-benzoate (C12107)	7	—	No paths reported by ATLAS, FMM and our algorithm

- Our output sub-networks are *complete* — they have all reactions necessary to produce every reactant in the pathway
- Smaller pathways of size 2 completely match with those generated by the other algorithms
- However, in many cases, we identify longer pathways, since these involve metabolites generated by branched pathways
- MetQuest correctly identified the already reported pathway between C00418 (Mevalonic acid) and C16028 (Amorpha-4,11-diene) — not identified by the other algorithms

MetQuest PERFORMANCE

NETWORKS OF DIFFERENT SIZES, FOR DIFFERENT β



MetQuest scales well to large genome-scale/community metabolic networks

- Consortium of *Clostridium cellulolyticum (cc)*, *Desulfovibrio vulgaris (dv)* & *Geobacter sulfurreducens (gs)*[1] ⇒ Directed Bipartite graph constructed
- Size of the network: 14265 nodes, 29073 edges
- Size of scope: 1135 metabolites
- Computed pathways of size 20, to all the metabolites within the scope of cellobiose and other seed metabolites
- Verified if the results contain pathways demonstrating experimentally proven metabolic exchanges
- In all the paths, acetate, pyruvate & ethanol were most frequently exchanged as previously shown[1]

[1]**Miller LD *et al* (2010)** *BMC Microbiol*. **10:**149

UNDERSTANDING METABOLIC INTERACTIONS IN THE GUT MICROBIOME

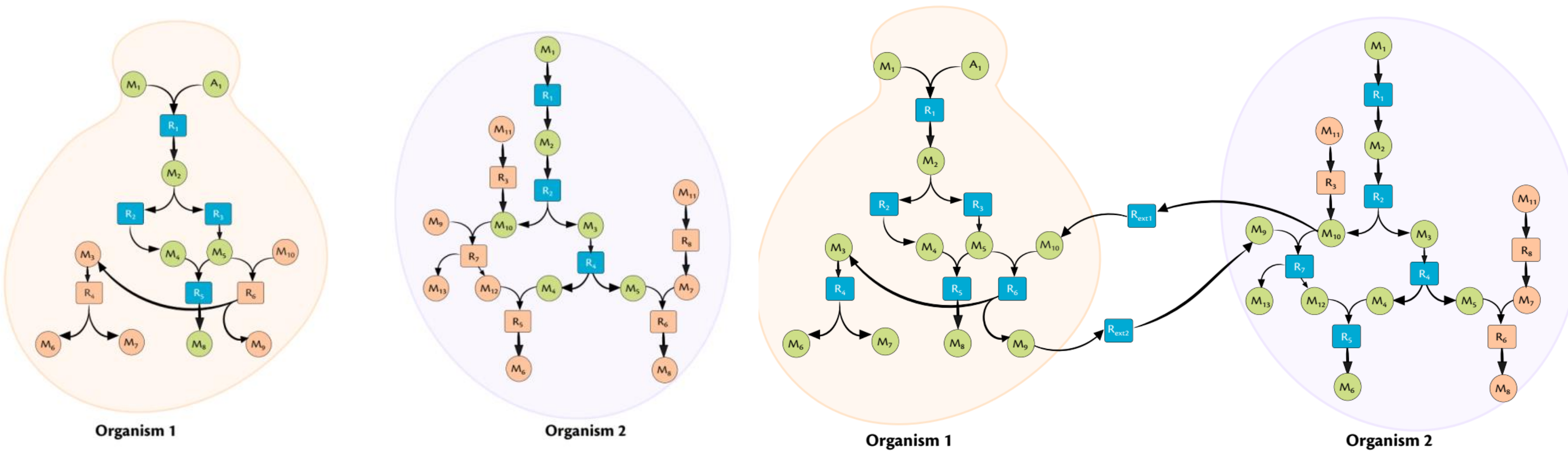
UNDERSTANDING POTENTIAL FOR SYNERGY BETWEEN ORGANISMS

- How well does one organism support another?
  - In terms of 'relieving' blocked reactions
  - In terms of improving metabolic capabilities

- Identifying exchanges that may contribute to better interactions
  - Possible targets for transporter overexpression

- Metabolic Support Index (MSI)
  - Fraction of blocked/stuck reactions relieved by the presence of another organism

METABOLIC SUPPORT INDEX



Individual metabolic networks **Community metabolic network**

- What fraction of stuck/blocked reactions in one is relieved by the presence of another organism?

$$MSI(A; A \cup B) = 1 - \frac{n_{stuck,\,A;A\cup B}}{n_{stuck,\,A;A}}$$

- If all stuck reactions remain stuck, there is no benefit, *i.e.* MSI = 0
- MSI seeks to quantify the *extent of benefit* (asymmetric)

METABOLIC SUPPORT: KNOWN CO-CULTURES

WHICH ORGANISM BENEFITS MORE?

Co-culture organisms	MSI of A	MSI of B	Experimental observations
Ketogulonicigenium vulgare (A) *and Bacillus megaterium*(B)	**0.224**	0.016	*B. megaterium* is a helper strain for *K. Vulgare*
Yarrowia lipolytica and Cellulomonas fimi	**0.120**	0.018	C. *fimi* provides additional metabolites to *Y. lipolytica* in a co-culture setup
Desulfovibrio vulgaris and Methanococcus maripaludis	**0.179**	0.036	*D. vulgaris* benefits the interaction with *M. maripaludis*
Clostridium cellulolyticum and *Clostridium acetobutylicum*	**0.052**	0.003	C. *acetobutylicum* helps C. *cellulolyticum* when grown together in a cellulose rich medium
Pichia stipitis and *Saccharomyces cerevisiae*	0.016	**0.032**	*P. stipitis* benefits the interactions with *S. cerevisiae* (*experimental validations were performed*)

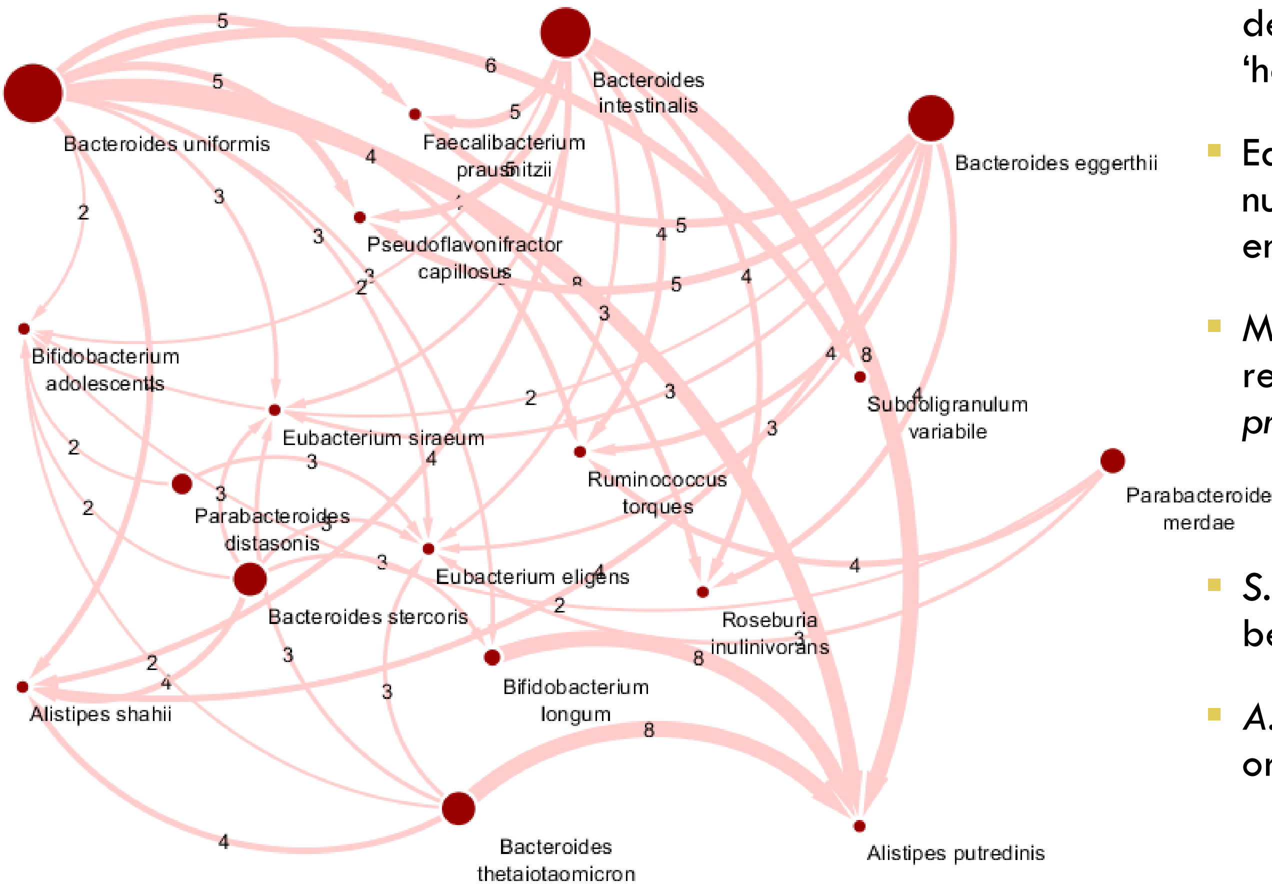
In *all cases*, organism with higher MSI exhibited higher biomass in co-culture!

SCOPE AND PATHWAY ANALYSES POINT TO KEY PLAYERS IN THE GUT

- 20 microbial species — antibiotic recovery associated bacteria (RABs) — chosen based on a previous study[1]
- Joint metabolic networks (of $20C_2 = 190$ combinations of RAB organisms) constructed
- Microbial association network constructed based on amino acid synthesising capabilities
- Analysed all the AA production pathways for exchange of metabolites
  - Two-way analyses performed—pathways originating from the source of one organism (glucose) and ending in amino acids of the other were analysed
  - Metabolites exchanged between the organisms were identified (from all the pathways)

[1]Nandi T *et al* (2018) bioRxiv doi:10.1101/350470

MICROBIAL ASSOCIATION NETWORKS

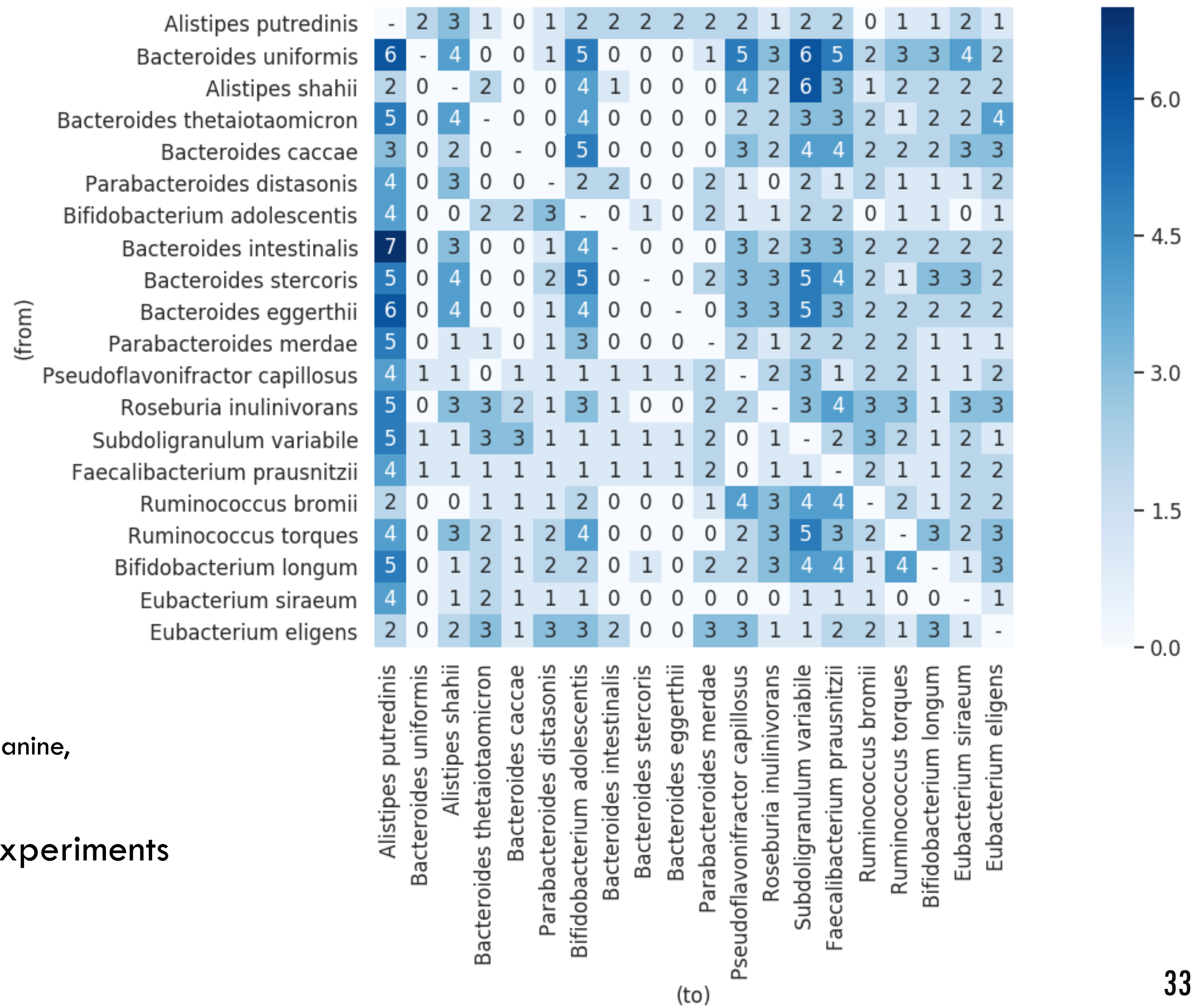


- Node size corresponds to out-degree, i.e. number of organisms 'helped'
- Edge weights correspond to number of new amino acids enabled to be produced
- Many organisms benefit the relationship with *B. uniformis, as previously demonstrated*[1]
- *S. variabile* exhibits maximum benefit only with *B. uniformis*
- *A. putredinis* does not help other organisms but only *receives*

[1]**Nandi et al** (2018) *bioRxiv doi:10.1101/350470*

AA PATHWAY ANALYSES REVEALS INTERESTING INTERACTIONS/EXCHANGES

- Figure shows number of metabolites exchanged towards AA production
- Some relationships are "two-way"
- Others very one-sided
- Very sensitive to the environment: most interactions lost in a "high fibre" diet
- Several metabolites exchanged
  - Fermentation products such as acetate, formate and L-lactate
  - Amino acids such as L-phenylalanine, L-glycine and L-threonine
- Need validation against experiments



(from) \ (to)	Alistipes putredinis	Bacteroides uniformis	Alistipes shahii	Bacteroides thetaiotaomicron	Bacteroides caccae	Parabacteroides distasonis	Bifidobacterium adolescentis	Bacteroides intestinalis	Bacteroides stercoris	Bacteroides eggerthii	Parabacteroides merdae	Pseudoflavonifractor capillosus	Roseburia inulinivorans	Subdoligranulum variabile	Faecalibacterium prausnitzii	Ruminococcus bromii	Ruminococcus torques	Bifidobacterium longum	Eubacterium siraeum	Eubacterium eligens
Alistipes putredinis	-	2	3	1	0	1	2	2	2	2	2	2	1	2	2	0	1	1	2	1
Bacteroides uniformis	6	-	4	0	0	1	5	0	0	0	1	5	3	6	5	2	3	3	4	2
Alistipes shahii	2	0	-	2	0	0	4	1	0	0	0	4	2	6	3	1	2	2	2	2
Bacteroides thetaiotaomicron	5	0	4	-	0	0	4	0	0	0	0	2	2	3	3	2	1	2	2	4
Bacteroides caccae	3	0	2	0	-	0	5	0	0	0	0	3	2	4	4	2	2	2	3	3
Parabacteroides distasonis	4	0	3	0	0	-	2	2	0	0	2	1	0	2	1	2	1	1	1	2
Bifidobacterium adolescentis	4	0	0	2	2	3	-	0	1	0	2	1	1	2	2	0	1	1	0	1
Bacteroides intestinalis	7	0	3	0	0	1	4	-	0	0	0	3	2	3	3	2	2	2	2	2
Bacteroides stercoris	5	0	4	0	0	2	5	0	-	0	2	3	3	5	4	2	1	3	3	2
Bacteroides eggerthii	6	0	4	0	0	1	4	0	0	-	0	3	3	5	3	2	2	2	2	2
Parabacteroides merdae	5	0	1	1	0	1	3	0	0	0	-	2	1	2	2	2	2	1	1	1
Pseudoflavonifractor capillosus	4	1	1	0	1	1	1	1	1	1	2	-	2	3	1	2	2	1	1	2
Roseburia inulinivorans	5	0	3	3	2	1	3	1	0	0	2	2	-	3	4	3	3	1	3	3
Subdoligranulum variabile	5	1	1	3	3	1	1	1	1	1	2	0	1	-	2	3	2	1	2	1
Faecalibacterium prausnitzii	4	1	1	1	1	1	1	1	1	1	2	0	1	1	-	2	1	1	2	2
Ruminococcus bromii	2	0	0	1	1	1	2	0	0	0	1	4	3	4	4	-	2	1	2	2
Ruminococcus torques	4	0	3	2	1	2	4	0	0	0	0	2	3	5	3	2	-	3	2	3
Bifidobacterium longum	5	0	1	2	1	2	2	0	1	0	2	2	3	4	4	1	4	-	1	3
Eubacterium siraeum	4	0	1	2	1	1	1	0	0	0	0	0	0	1	1	1	0	0	-	1
Eubacterium eligens	2	0	2	3	1	3	3	2	0	0	3	3	1	1	2	2	1	3	1	-

LIMITATIONS

- Only a static snapshot of interactions happening in the gut
- Nevertheless, graph-based approaches are very useful and complement constraint-based models
- Some predictions agree with previous experiments, but many remain to be validated
- MetQuest algorithm
  - Predictions are obviously heavily contingent on the quality of the input network
  - No weights or ranking attached to the metabolites/paths
  - Difficult to identify very long pathways — but they may not be very interesting!
- Also, we view cellular interactions only through a metabolic lens — lots more happening in reality!

SUMMARY

SUMMARY

- Gut microbiome suffers post-antibiotic treatment, yet recovers
  - Recovery facilitated much better by certain bacteria
- How to dissect the complexity of these interactions in the gut microbiome?
- We developed *MetQuest* – *a* novel dynamic-programming based enumeration
- Exhaustively identifies all possible pathways between a set of source and target molecules (within a size)
- Employs a two-phase approach: Guided BFS & Dynamic–programming based generation of pathways
- Overcomes the shortcomings of existing tools
- Scales well to large networks and identifies longer pathways

SUMMARY

- Particularly interesting to identify metabolic cross-talks happening between micro-organisms in a community
- Metabolic Support Index sheds light on nature of (pairwise) interactions between microbes
- MetQuest identifies several metabolic exchanges/dependencies in gut flora
- These exchanges are environment-dependent
- Several metabolites exchanged
  - Fermentation products & Amino acids
- Ongoing work: reaction rescues in gut microbiome / microbial communities

- **Generic algorithm/approach — can be applied to any microbial community to identify pathways and metabolic interactions**

AVAILABILITY/USAGE

- `$ pip3 install metquest`
- Ravikrishnan, Nasre & Raman (2018) *Scientific Reports* **8**:9932
- Ravikrishnan, Blank, Srivastava & Raman (2019) *bioRxiv*

10.1101/532184

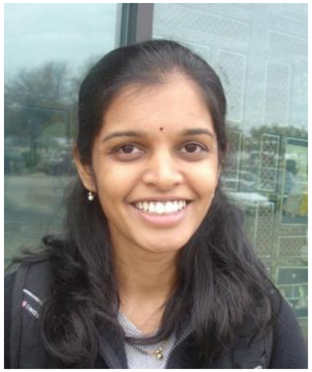
http://metquestdoc.readthedocs.io/

/RamanLab/MetQuest

ACKNOWLEDGEMENTS

- Aarthi Ravikrishnan
  - DST-INSPIRE
  - DAAD-UGC Exchange Programme

- Dr. Meghana Nasre (CSE, IITM)



THANK YOU!