Improved Predictive Models using Omics Data

Dr Vishwesh Kulkarni (warwick.ac.uk/kulkarnilab)

WARWICK
INTEGRATE ANTIMICROBIAL RESISTANCE

Objectives

1. Perform synthetic reaction and gene lethality analysis for genome-scale metabolic networks of *E. coli*, *S. enterica*, and *Mycobacterium tuberculosis*.
2. Characterize synthetic lethal gene deletions for these.
3. Improve on the *Flux Balance Analysis* (FBA) predictions using quadratic programming. Characterize the improvement via public domain data on the *transcription factor* (TF) knock-outs.

Materials

- SBML models of the organisms: *E.coli*: iAF1260 [2]; *S. enterica Typhimurium*: LT2 STM v1.0 [3]; *M. tuberculosis*: iNJ661 [4].
- MATLAB (R2015b) interfaced with COBRA Toolbox v2.X.
- PC: 2.4 GHz Intel Xeon E5645 processor with 16 GB DDR3 RAM running Windows 8.1 using the IBM CPLEX v12.5.1 solver.
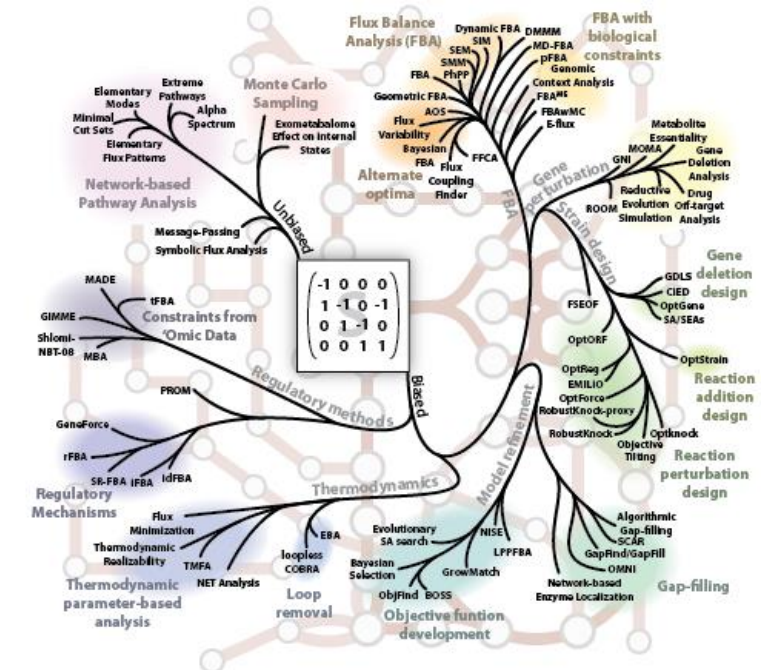
Methods

Fig. 1 Our results enrich the COBRA phylogeny [5].

- We solve the problem of minimizing the l_1 norm of the fluxes subject to the FBA constraints such that the biomass produced is the same as the wild type.
- For each reaction having a non-zero flux, it is deleted and identifies a synthetic lethal if the corresponding growth is less than 5% of the FBA prediction.
- Synthetic doubles/triplets/quadruplets found likewise.
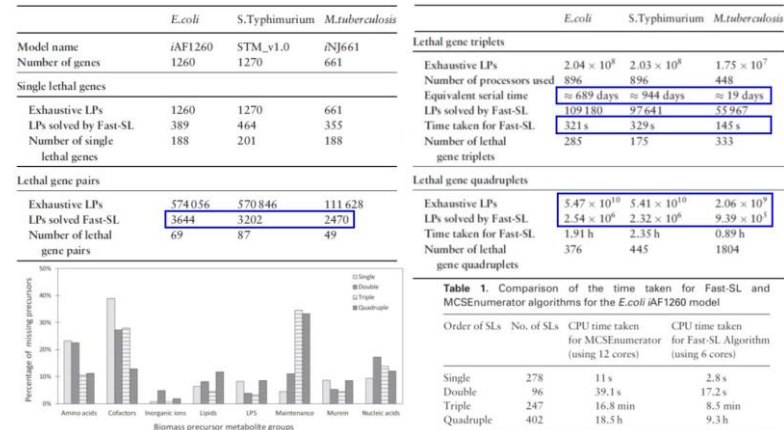
Results: Synthetic Lethal Discovery*

	E.coli	*S.Typhimurium*	*M.tuberculosis*
Model name	iAF1260	STM_v1.0	iNJ661
Number of genes	1260	1270	661
Single lethal genes			
Exhaustive LPs	1260	1270	661
LPs solved by Fast-SL	389	464	355
Number of single lethal genes	188	201	188
Lethal gene pairs			
Exhaustive LPs	574 056	570 846	111 628
LPs solved Fast-SL	3644	3202	2470
Number of lethal gene pairs	69	87	49

	E.coli	*S.Typhimurium*	*M.tuberculosis*
Lethal gene triplets			
Exhaustive LPs	2.04×10^8	2.03×10^8	1.75×10^7
Number of processors used	896	896	448
Equivalent serial time	≈ 689 days	≈ 944 days	≈ 19 days
LPs solved by Fast-SL	109 180	97 641	55 967
Time taken for Fast-SL	321 s	329 s	145 s
Number of lethal gene triplets	285	175	333
Lethal gene quadruplets			
Exhaustive LPs	5.47×10^{10}	5.41×10^{10}	2.06×10^9
LPs solved by Fast-SL	2.54×10^6	2.32×10^6	9.39×10^5
Time taken for Fast-SL	1.91 h	2.35 h	0.89 h
Number of lethal gene quadruplets	376	445	1804

Table 1. Comparison of the time taken for Fast-SL and MCSEnumerator algorithms for the *E.coli* iAF1260 model

Order of SLs	No. of SLs	CPU time taken for MCSEnumerator (using 12 cores)	CPU time taken for Fast-SL Algorithm (using 6 cores)
Single	278	11 s	2.8 s
Double	96	39.1 s	17.2 s
Triple	247	16.8 min	8.5 min
Quadruple	402	18.5 h	9.3 h

Fig. 2 Our algorithms speed up the computation of synthetic lethals (single/multiple) significantly.

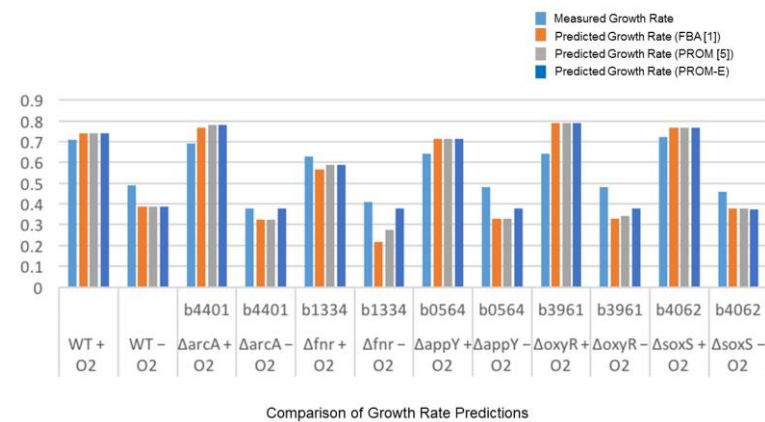## Results: Growth Prediction using QP

Fig 3 Prediction using our quadratic programming algorithm outperforms all other known results.

* Algorithms coded, in parts, by Sophie Clarke, an undergraduate student at the School of Engineering.

Future Directions

We aim to conduct more TF-knock out experiments and develop techniques to include miRNA and other datasets. **Contact us if you have datasets!**

References

[1] J. Orth *et al*. Nature Biotechnology 28(3):245-248, 2010. [2] A. Feist *et al*. Mol. Syst. Biol., 3:121, 2007. [3] I. Thiele *et al*. J. Bacteriol., 187: 5818–5830, 2005. [4] N. Jamshidi and B. Palsson. BMC Syst. Biol., 1:26, 2007. [5] N. Lewis *et al*. Nature Reviews Microbiology 10, 291-305, 2012. [6] S. Chandrasekaran and N. Price. PNAS 107(41):17845-50, 2012.