Efficient Computation of Synthetic Lethals

Vishwesh Kulkarni*
School of Engineering, University of Warwick
Email: V.Kulkarni@warwick.ac.uk

Karthik Raman†
Department of Biosciences, IIT Madras
Email: kraman@iitm.ac.in

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 networks.
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-out experiments.

Introduction

- Synthetic lethality arises when a combination of deficiencies in the expression of two or more genes leads to cell death, whereas a deficiency in only one of these genes does not. The deficiencies can arise through mutations, epigenetic alterations or inhibitors of one of the genes.
- The phenotype resulting from a gene/reaction deletion is termed a **lethal phenotype**, if the maximum growth-rate predicted by FBA is less than a cut-off, typically 1% of the maximum *in silico* wild-type growth rate.
- FBA [1] solves a *Linear Programming* (LP) problem in which, typically, flux through a set of biomass reactions is to be maximized subject to the constraints obtained from the stoichiometry of the metabolic network.

Materials

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

Methods

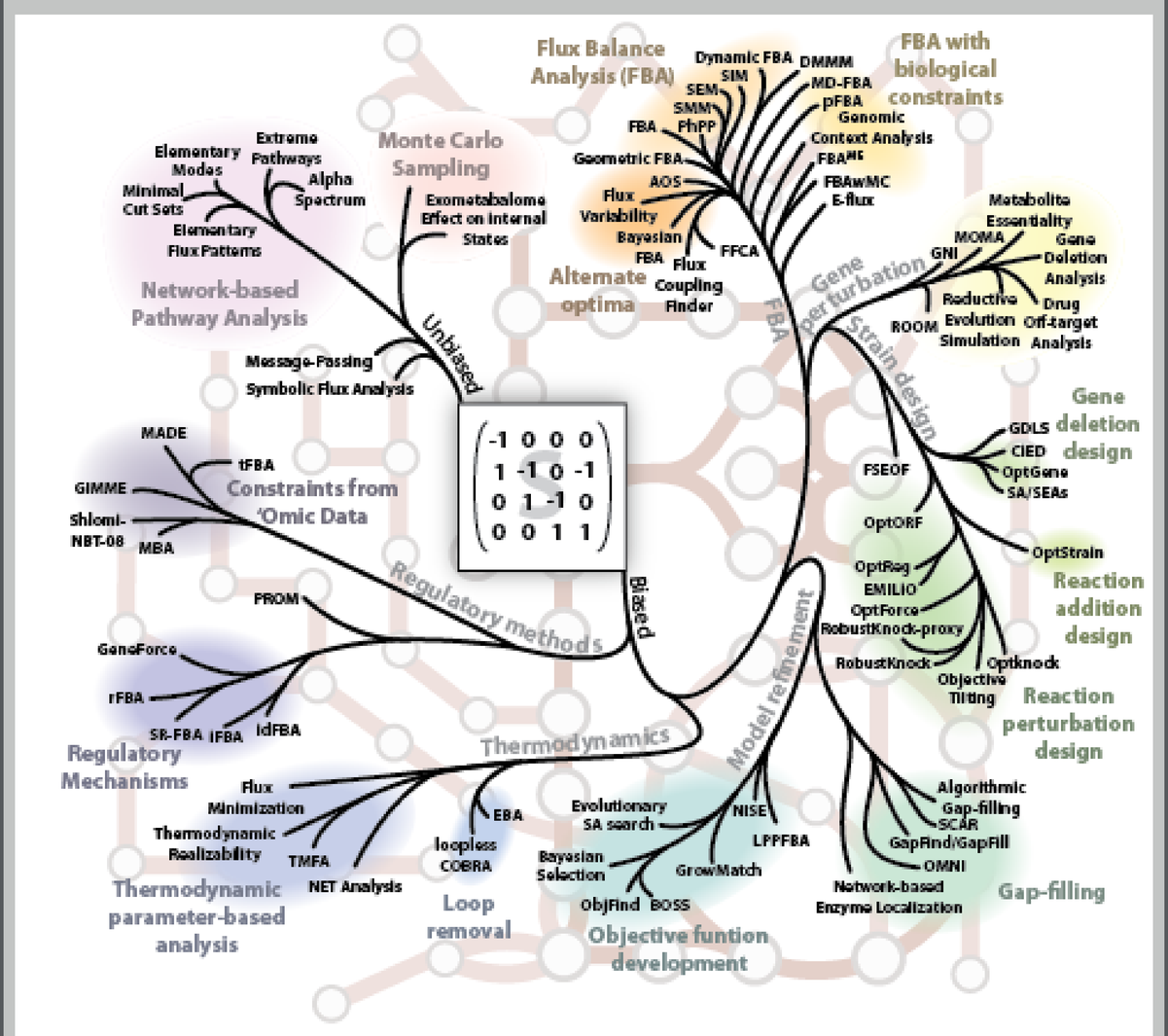

Figure 1: We enrich the constraint-based reconstruction & analysis (COBRA) phylogeny [5].

Mathematical Section

- 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. Let J_{nz} denote the set of reactions that have non-zero flux.

Single Lethals:

for each reaction $i \in J_{nz}$ **do**
  Set the upper and lower bounds of v_i to zero
  Do FBA to maximise growth rate, $v_{bio,i}$
  if $v_{bio,i} \leq v_{co}$ **then**
    Add i to the set J_{sl}
  end if
  Reset bounds on v_i
end for

Double Lethals:

for each reaction $i \in J_{nz} - J_{sl}$ **do**
  Set the upper and lower bounds of v_i to zero
  Do FBA to obtain minimum norm solution corresponding to $v_{bio,i}$
  Identify set of reactions $J_{nz,i}$, having non-zero fluxes
  for each reaction $j \in J_{nz,i} - J_{sl}$ **do**
    Set the upper and lower bounds of v_j to zero
    Do FBA to maximise growth rate $v_{bio,ij}$
    if $v_{bio,ij} \leq v_{co}$, **then**
      Add $\{i,j\}$ to the set J_{dl}
    end if
    Reset the bounds on v_j
  end for
  Reset bounds on v_i
end for

Results: Synthetic Lethal Discovery and Speed-Up

	E.coli	S.Typhimurium	*M.tuberculosis*
Model name	*i*AF1260	STM_v1.0	*i*NJ661
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
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

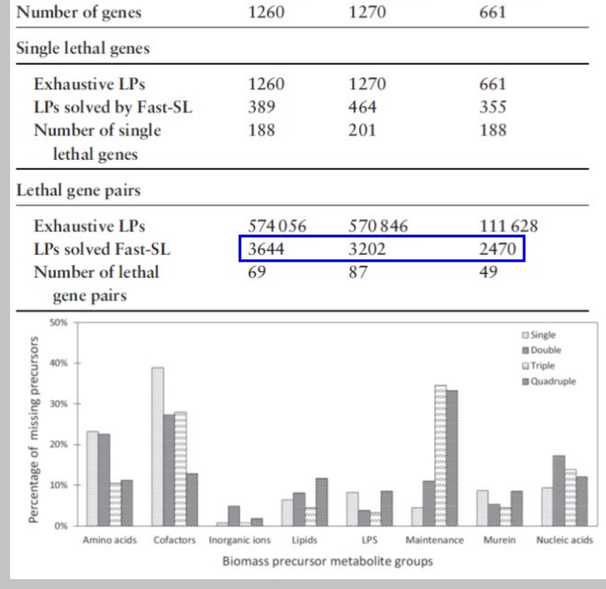

Table 1. Comparison of the time taken for Fast-SL and MCSEnumerator algorithms for the *E.coli* *i*AF1260 model

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

Figure 3: Our algorithms significantly speed up the synthetic lethal identification process.

Results: Quadratic Programming for Growth Predictions

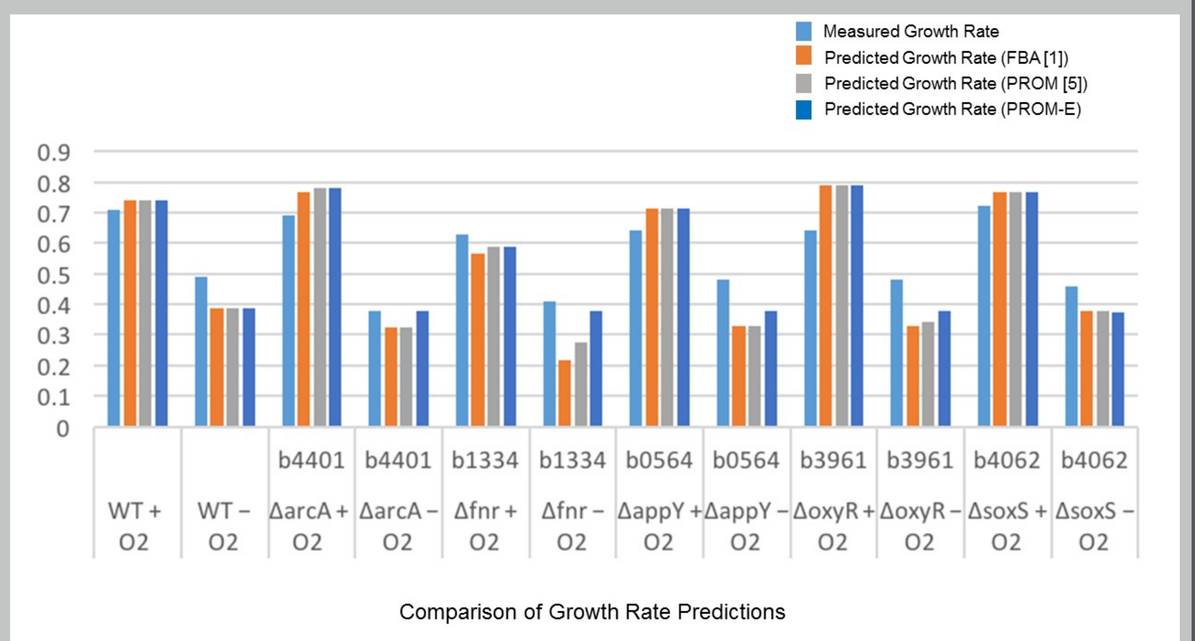

Figure 4: Our quadratic programming solution outperforms other algorithms.

Conclusion and Future Directions

- We have extended the scope of FBA for a more comprehensive and more efficient determination of synthetic lethals.
- Our quadratic programming algorithms predict growth better than other algorithms but require more data.
- We aim to conduct more TF-knock out experiments and develop techniques to include miRNA 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: 58185830, 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.

Acknowledgments

- VK acknowledges funding from Warwick Institute of Synthetic Biology (WISB), BBSRC, and EPSRC. KR acknowledges funding from IIT Madras and the grant BT/PR4949/BRB/10/1048/2012 from Govt. of India.

Contact Information

- Web: http://www2.warwick.ac.uk/fac/sci/eng/staff/vk/
- Email: V.Kulkarni@warwick.ac.uk
- Phone: +44 (024) 765 24806
- **We are looking to work with additional datasets. Do contact us if you have omics data to share!**