Subject Section

Integrating splice-isoform expression into genome-scale models characterizes breast cancer metabolism - Supplementary Material

Claudio Angione 1,*

1Department of Computer Science and Information Systems, Teesside University, Middlesbrough, UK

*To whom correspondence should be addressed.

Associate Editor: XXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Flux Balance Analysis (FBA)

To generate cancer-specific models from Cancer RNA-Seq Nexus, we propose an extended flux balance analysis (FBA) framework. FBA has been successfully applied to predict growth rate and other metabolic observables in cells [Palsson, 2015]. Starting from the stoichiometry matrix $S$ of all known metabolic reactions in a cell, and given an unknown vector $v$ of reaction flux rates, FBA models the steady-state of a cell through the condition $Sv = 0$. Further constraints are added as a lower- and upper-limit for the flux rate of each reaction. Through linear programming, a combination of flux rates is chosen as an objective to be maximized or minimized. The underdetermined linear system $Sv = 0$ is then solved and a distribution of flux rates is determined according to the objectives chosen and the constraints added on the flux rates. Standard FBA is extended through METRADE [Angione and Lió, 2015] to achieve multi-omic integration and multi-level linear programming, as detailed in the main manuscript.

GEMsplice pipeline

To build GEMsplice, we performed manual curation of the model using data associated with each gene in Recon1, retrieving transcript level information and all available identifiers published as supplementary material but not included in the model. We then mapped transcript identifiers to the Recon1 SBML file to transcript IDs used by databases, using BioMart [Smedley et al., 2015] and Ensembl [Yates et al., 2016]. We finally used these IDs within the breast cancer metabolic model by Jerby et al. [Jerby et al., 2012]. Using a RNA-Seq dataset and a method based on constraint-based programming, we were able to map transcript expression levels, including splice isoforms, onto the breast cancer model in order to generate 31 cancer-specific metabolic models. Cancer RNA-Seq Nexus was used as a RNA-Seq dataset [Li et al., 2016], including data from TCGA [Network et al., 2012], GEO [Barrett et al., 2013] and SRA [Kodama et al., 2012].

Analysis of invasive versus unlabeled cancer cells

Overall, unlabeled cancer cells, defined as the normal cells and the cells that have not been clearly detected as invasive, are predicted to produce more biomass than invasive cells. The trade-off frontiers in Figure S1 suggest that invasive cancer cells outperform unlabeled cells in carrying a positive flux of pyruvate kinase while ensuring a large flux for lactate dehydrogenase. More specifically, when comparable flux rates of pyruvate kinase are considered, the invasive breast cancer cells considered in our study are able to achieve a higher flux for the reverse reaction lactate dehydrogenase, which in turn represents an increased Warburg effect, i.e., an increased production of lactate from pyruvate. This result supports the hypothesis that aggressive phenotype and worse clinical outcome are positively correlated with Warburg effect [Gatenby and Gillies, 2004].

The pathways detected as outliers, with significantly different behavior between invasive and unlabeled breast cell, are reported in Figure S2. The plot also shows overall positive correlation between invasive and unlabeled perturbations (Pearson's $r = 0.22$, $p$-value = 1.17 $\times$ 10$^{-4}$, Spearman's $\rho = 0.63$, $p$-value = 3.87 $\times$ 10$^{-7}$). This suggests that our method behaves consistently across different cancer subsets, but also highlights pathways whose behavior is significantly changed in the two cases.

References


Invasive - normalized relative pathway flux change
-40
-30
-20
-10
0
10
20
30
Unlabeled - normalized relative pathway flux change
0
10
20
30

been detected as invasive.

unlabeled
invasive
invasive
unlabeled
invasive
unlabeled
invasive
unlabeled
invasive
unlabeled
invasive
invasive
unlabeled

Pyruvate kinase [mmolh⁻¹ gDW⁻¹]
Biomass [h⁻¹]

Lactate dehydrogenase [mmolh⁻¹ gDW⁻¹]
Biomass [h⁻¹]

Pyruvate kinase [mmolh⁻¹ gDW⁻¹]
Lactate dehydrogenase [mmolh⁻¹ gDW⁻¹]
Biomass [h⁻¹]

Pyruvate kinases [mmolh⁻¹ gDW⁻¹]
Biomass [h⁻¹]

Lactate dehydrogenase [mmolh⁻¹ gDW⁻¹]
Biomass [h⁻¹]

Fig. S1. Starch and Sucrose Metabolism
Vitamin A Metabolism
Fatty Acid Elongation
Sphingolipid Metabolism
IMP Biosynthesis
Pyrimidine Biosynthesis

Fig. S2. Transcript-based pathway analysis. Indicator deployment for invasive breast cancer versus unlabeled breast cells. The size of each point represents the number of reactions in the pathway. Colors are given according to a normalized error of the computed mean flux in each pathway, namely \( NE = \sigma / \mu \), where \( \sigma \) is the standard deviation of the flux rate in the pathway, computed across all breast cancer subsets, and \( \mu \) represents the size of the pathway. The \( y = x \) line is shown on the plots to highlight outliers. By definition of \( x \), if a pathway lies outside the line, its perturbation in cancer samples is different from its perturbation in healthy samples (both perturbations are computed with respect to the default breast cell model run without omic-derived constraints).