Mechanistic effects of influenza in bronchial cells through poly-omic genome-scale modelling

Elisabeth Yaneske
Department of Computer Science and Information Systems, Teesside University
Middlesbrough, United Kingdom
e.yaneske@tees.ac.uk

Claudio Angione
Department of Computer Science and Information Systems, Teesside University
Middlesbrough, United Kingdom
c.angione@tees.ac.uk

ABSTRACT

In this work we propose regularised bi-level constraint-based modelling to determine the fluxomic profiles for four different influenza viruses, H7N9, H7M7, H3N2 and H5N1. We report here the first step of the analysis of the flux data using AutoSOME clustering, where we identify novel biomarkers of infection. This is a work in progress that can directly lead to novel therapeutic targets.

CCS CONCEPTS

• **Applied computing** → **Computational biology**; **Systems biology**;

KEYWORDS

genome-scale models; regularisation; bi-level optimisation.

ACM Reference Format:
Elisabeth Yaneske and Claudio Angione. 2018. Mechanistic effects of influenza in bronchial cells through poly-omic genome-scale modelling. In *Proceedings of 10th International Workshop on Bio-Design Automation (IWBDA).* ACM, New York, NY, USA, 3 pages. https://doi.org/10.475/123_4

1 INTRODUCTION

Previous work [6] analysed transcriptomic data to identify FDA-approved antiviral drugs that would be effective against the H7N9 Anhui01 influenza virus. This was done by infecting human bronchial epithelial cells cells with H7N9 and comparing the transcriptomic profile of these with cells infected with H3N2, H5N1 and H7N7. Four replicate samples were taken at 3, 7, 12, and 24 hours. A control batch of uninfected cells was also sampled at the same time. Here we extend on this work by applying genome-scale modelling to the transcriptomic profiles of the four strains of influenza virus H3N2, H5N1, H7N7, H7N9 in order to determine their metabolic phenotypes.

Standard metabolic models created using FBA (Flux Balance Analysis) and constraint-based modelling have no unique solution for the optimal flux vector. The Cobra 3.0 toolbox [4] introduced a regularisation function so that the optimisation problem has a single unique solution. We here adapt the regularisation function to create a novel bi-level linear program with FBA and regularisation. To our knowledge, this is the first time this has been reported in the literature. This modelling procedure enables us to predict how the distribution of flux rates within the cell responds to infection with different influenza viruses. The transcriptomic data from each individual virus is used to constrain the model to generate a virus-specific metabolic model for each of the four influenza strains at each of the four time points sampled.

2 METHODS

Data processing and metabolic modelling

After retrieving the transcripomic data was from GEO (GSE49840), the probe data was matched to HGNC IDs. Where multiple probes were associated with a single HGNC ID, the gene expression values were averaged. The replicate samples were averaged to give a single transcriptomic profile for each time point. The transcriptomic data was normalised by taking the ratio of the influenza data to the control data to obtain the fold change. The normalised transcriptomic profiles of the influenza viruses were then used to create virus specific bronchial epithilel cell metabolic models. The metabolic models were created using constraint based modelling and flux balance analysis (FBA) of the human epithelial cell augmented with transcriptomics [6] through GEMsplice [1].

Constraint-based modelling with regularisation

In FBA the cell is assumed to be in steady state, $Sv = 0$, where S is a stoichiometric matrix of all known metabolic reactions (metabolites by reactions) and v is the vector of

IWBDA, July 31 - August 3 2018, Berkley, CA USA

ACM ISBN 123-4567-24-567/08/06…$15.00
https://doi.org/10.475/123_4

reaction by flux rates. Additionally, every reaction flux is constrained by lower- and upper- bounds ($v^{\min}$ and $v^{\max}$). Here we constrain the strain-specific metabolic models generated from the transcriptomics data with upper- and lower-bounds on reactions set as a function of the expression level of the genes involved in the reactions using GEMsplice [1]. We set the primary objective as maximisation of hexokinase [7] and the secondary objective as maximisation of UDP-N-acetylglucosamine diphosphorylase [10]. We additionally apply regularisation to the secondary objective function such that it is maximised subject to the primary objective being maximised with a penalty term defined as a multiple of $v^T v$. This is achieved by adding a function that drives minimisation of the squared flux rates. This state reflects the most efficient metabolic network. We use the following bi-level program with regularisation:

$$\begin{aligned} &\max\ g^\intercal v - \frac{\sigma}{2} v^T v \\ &\text{such that} \qquad \max\ f^\intercal v, \quad Sv = 0, \qquad (1) \\ &\qquad\qquad v^{\min}\varphi(\Theta) \le v \le v^{\max}\varphi(\Theta). \end{aligned}$$

The Boolean vectors f and g are weights to select the first and second objectives respectively to be maximised from the vector v i.e. hexokinase and UDP-N-acetylglucosamine diphosphorylase. The vectors $v^{\min}$ and $v^{\max}$ represent the lower- and upper-bounds for flux rates. The regularisation function ($\frac{\sigma}{2} v^T v$) requires that the sum of the square of the fluxes is minimised for the maximisation of the second objective to be obtained. To maintain the optimal value of the original linear objective whilst minimising the square of the fluxes, the coefficient, σ, is set to 10^{-6}.

The vector Θ represents the set of gene expression values for the enzymes catalysing the biochemical reactions associated with the vector of fluxes v. The upper- and lower-bounds are constrained depending on the expression levels of the enzymes and a rule based on the type of enzyme (single enzyme, isozyme, or enzymatic complex) using the function φ [2]. Simulations were performed in Matlab R2016b.

Clustering

To cross-compare the fluxomics of the four viruses, flux distributions were clustered using AutoSOME [8], an unsupervised SOM-based method for high-dimensional data that uses a combination of density equalisation, minimum spanning tree clustering and ensemble averaging strategies. AutoSOME has the advantage that it does not require prior knowledge of the number of clusters and is not skewed by outliers in the data.

3 RESULTS AND CONCLUSIONS

Clustering the influenza sample subsystems according to their flux profile using AutoSOME resulted in four clusters.

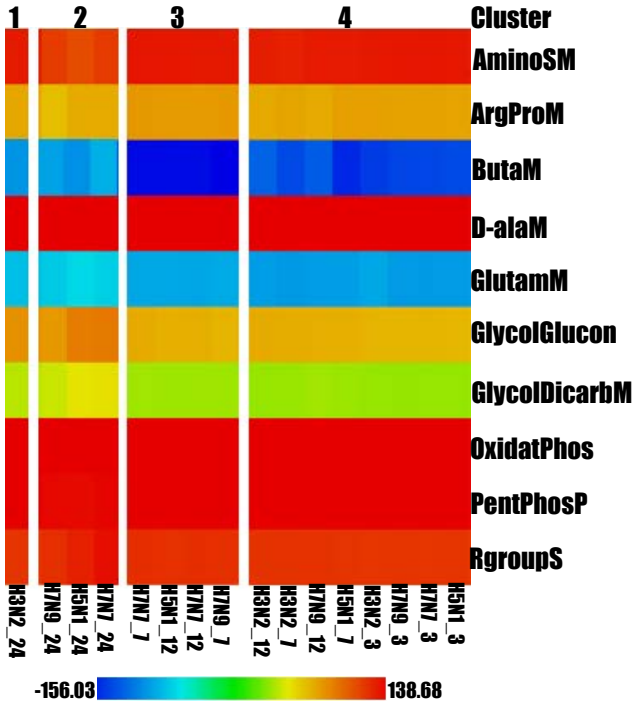

Figure 1: Heatmap of AutoSOME clustering. A subset of the subsystems is shown illustrating the variability between the four clusters.

In agreement with [6], H7N9 shows its own unique metabolic profile. Among the avian viruses at 24 hours of influenza infection, the metabolic profile of H7N9 is closest to H5N1 though it shares similarities with the H3N2 virus at both 12 and 24 hours [6]. Of the pathways showing strongest perturbations, the pentose phosphate pathway [9], oxidative phosphorylation [3] and r-group (de novo viral protein) synthesis [11] have previously been identified as important in viral replication. The importance of amino sugar metabolism may be due to its links with glycolysis [9] and glycoprotein production [10]. D-alanine metabolism has not previously been reported but may be important in the production of pyruvate [9] for viral replication. Butanoate metabolism shows a different profile across the four clusters. Butanoate metabolism has also not previously been reported but may highlight differences in viral cAMP signalling [5]. These results identify novel biomarkers of infection, suggesting that further analysis of the data using machine learning techniques focussed on these metabolic features could contribute to the identification of novel therapeutic targets.

REFERENCES

[1] C Angione. 2018. Integrating splice-isoform expression into genome-scale models characterizes breast cancer metabolism. *Bioinformatics* 34, 3 (2018), 494–501.

[2] C Angione and P Lió. 2015. Predictive analytics of environmental adaptability in multi-omic network models. *Scientific reports* 5 (2015), 15147.

[3] Adi Bercovich-Kinori, Julie Tai, Idit Anna Gelbart, Alina Shitrit, Shani Ben-Moshe, Yaron Drori, Shalev Itzkovitz, Michal Mandelboim, and Noam Stern-Ginossar. 2016. A systematic view on influenza induced host shutoff. *Elife* 5 (2016), e18311.

[4] L Heirendt, S Arreckx, T Pfau, S N Mendoza, A Richelle, A Heinken, H S Haraldsdottir, S M Keating, V Vlasov, J Wachowiak, et al. 2017. Creation and analysis of biochemical constraint-based models: the COBRA Toolbox v3. 0. *arXiv preprint arXiv:1710.04038* (2017).

[5] Noriyuki Hirata, Futoshi Suizu, Mami Matsuda-Lennikov, Tatsuma Edamura, Jyoti Bala, and Masayuki Noguchi. 2014. Inhibition of Akt kinase activity suppresses entry and replication of influenza virus. *Biochemical and biophysical research communications* 450, 1 (2014), 891–898.

[6] L Josset, H Zeng, S M Kelly, T M Tumpey, and M G Katze. 2014. Transcriptomic characterization of the novel avian-origin influenza A (H7N9) virus: specific host response and responses intermediate between avian (H5N1 and H7N7) and human (H3N2) viruses and implications for treatment options. *MBio* 5, 1 (2014), e01102–13.

[7] Hinissan P Kohio and Amy L Adamson. 2013. Glycolytic control of vacuolar-type ATPase activity: a mechanism to regulate influenza viral infection. *Virology* 444, 1-2 (2013), 301–309.

[8] A M Newman and J B Cooper. 2010. AutoSOME: a clustering method for identifying gene expression modules without prior knowledge of cluster number. *BMC bioinformatics* 11, 1 (2010), 117.

[9] Joachim B Ritter, Aljoscha S Wahl, Susann Freund, Yvonne Genzel, and Udo Reichl. 2010. Metabolic effects of influenza virus infection in cultured animal cells: Intra-and extracellular metabolite profiling. *BMC systems biology* 4, 1 (2010), 61.

[10] David J Vigerust and Virginia L Shepherd. 2007. Virus glycosylation: role in virulence and immune interactions. *Trends in microbiology* 15, 5 (2007), 211–218.

[11] Tokiko Watanabe, Shinji Watanabe, and Yoshihiro Kawaoka. 2010. Cellular networks involved in the influenza virus life cycle. *Cell host & microbe* 7, 6 (2010), 427–439.