Journal of Molecular and Genetic Medicine

ISSN: 1747-0862

Marín de Mas, J Mol Genet Med 2019, 13:3

Commentary Open Access

Enhancing Transcriptomics-Based Model-Driven Performance *via* Stoichiometric Gene-to-Reaction Associations: Metabolic Reprogramming in Prostate Cancer

Igor Marín de Mas*

The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark

Abstract

Genome-scale metabolic models (GSMMs) have been widely used to study the molecular mechanisms underlying a variety of diseases with a strong metabolic component such as diabetes or cancer. GSMMs incorporate logical rules associating genes, proteins, and reactions (GPR rules), enabling the integration of either proteomics or transcriptomics. However, current GPR formulation do not account for the necessary stoichiometry to describe the number of transcript copies that are necessary to generate a catalytically active enzyme, which limits our understanding of how gene expression modulates metabolism. Thus, in this short commentary article, we introduce the stoichimetric-GPR (S-GPR) concept, presented in Marin de Mas I et al. The novel S-GPRs associations were implemented to study the metabolic reprogramming in DU145 prostate cancer cells associated to the chronic exposure to the endocrine disruptor Aldrin. The results showed that S-GPRs outperformed previous approaches by significantly improving the GSMM predictions. Thus, the novel S-GPR concept that we have developed enables a more precise integration of transcriptomics data into GSMM-based methods and can be extended to proteomics data, with an important impact in the environmental and the clinical fields.

Keywords: Metabolism; Proteomic; Transcriptomic; Prostate cancer

Commentary

Genetic background and environmental factors are some of the main triggers underlying tumor progression and metastasis. This multi-factorial disease is one of the first leading cause of death worldwide representing an important economic and social impact on health care systems [1,2].

In this sense, genome-scale metabolic models (GSMM) have been widely used in cancer research to study the aberrant tumoral metabolism [3]. GSMMs gather the current knowledge concerning the metabolic reactions taking place in a cell from a given organism/tissue [4]. This systems biology tool accounts for logical rules describing relationship between the genes encoding the enzymes that catalyze the metabolic reactions, the so-called GPRs, which makes GSMMs an excellent platform to integrate transcriptomic and proteomic data. Briefly, GPRs use logical "or" and "and" operators to represent isoenzymes encoded by different genes or genes encoding sub-units of a complex, respectively. The gene expression is incorporated into the GPRs that can be represented either in absolute or relative values, depending on the approach, the logical "or" is replaced by the "mean" or "max" operators and the logical "and" by the "minimum" operator. Next, by solving the mathematical expressions, the value of the metabolic reactions is calculated. Finally, the reactions values are implemented in the different transcriptomic-based model-driven methods to infer a case-specific metabolic flux profile. Thus, GPRs allow the integration of large omics data-sets generated by different high-throughput platforms which ultimately enhances the predictive capabilities of the GSMM-based methods [5]. In the last decade, a variety of GPR-based algorithms have been developed to integrate both proteomic and transcriptomic data into GSMMs [6]. However, despite the recent advance in the field, current GPR formulation does not consider the number of transcript copies that are required to produce a catalytically active enzyme. In other words, if the classical GPR formulation is applied on a complex with, for instance, three sub-units, one encoded by the gene "a" and two by the gene "b", both genes have the same weight in the calculation of the activity state of the associated metabolic reaction. Thus, classical GPRs assume that the same expression of gene "a" and "b" is needed to produce a catalytically active enzyme, regardless the complex has double sub-units encoded by gene "b" than by gene "a". Thus, this lack of stoichiometry in the GPR formulation limits our understanding on how gene expression modulates metabolism.

In order to overcome the existing lack of stoichiometry in the gene-to-reaction associations, we have enriched the current GPR formulation by accounting the stoichiometry of the transcripts: Stoichimetric-GPR (S-GPR) [7].

In brief, S-GPR formulation follows the same principle than GPRs rules but incorporating the transcripts stoichiometry. Following the same example described above, the GPR corresponding to the complex is "a and b", while the S-GPR formulation is "1*a and 2*b". Replacing the logical operators by min/max operators the GPR will be formulated as follow: "min (a, b)". However, S-GPR approach in addition divides the gene expression by its stoichiometric coefficient: "min (a/1, b/2)". Thus, if for instance, the expression of gene "a" is 3 and the expression of gene "b" is 4, based on GPR formulation, gene "a" is the limiting factor (exp.gene_a < exp.gene_b), while if S-GPR approach is implemented, the limiting factor is gene "b" (exp.gene_a /1 > exp.gene_b/2), which is more realistic. The accuracy in this step of the integration process is especially relevant as the way the gene expression data is incorporated to the analysis has an important impact on the reliability of the model

***Corresponding author:** Dr. Igor Marín de Mas, The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark, Tel: +4593511997; E-mail: igmar@biosustain.dtu.dk

Received September 06, 2019; **Accepted** September 20, 2019; **Published** September 27, 2019

Citation: Marin de Mas I (2019) Enhancing Transcriptomics-Based Model-Driven Performance via Stoichiometric Gene-to-Reaction Associations: Metabolic Reprogramming in Prostate Cancer. J Mol Genet Med 13: 432

predictions.

The S-GPR formulation was applied to study the metabolic reprogramming in the prostate cancer (PC) cell line DU145, associated to a prolonged exposure to sub-lethal concentrations of the endocrine disruptor (ED) Aldrin. This organic pollutant is found at low concentrations in different ecosystems [8]. Numerous studies have already shown a clear relation between increasing ED concentration and an enhanced tumor malignancy. However, how a long exposure to low and non-lethal EDs concentrations affects cancer metabolism remains poorly understood. In this sense, Bedia et al. reported important alterations in the metabolic and lipidomic profile of the DU145 prostate cancer cells that are associated with an enhanced tumor malignancy, after a chronic exposure to low concentrations of Aldrin [9].

In this study, DU145 cells were exposed for a long period (50 days) to a pollutant concentration that didn't compromise their growth rate. As consequence, it exists a high isogeneity between Aldrin-exposed and non-exposed cells. Indeed, only 0.35% of genes are deferentially expressed between conditions and among these genes only 1.62% are metabolic genes. However, it has been reported that Aldrin-exposed cells present an enhanced malignant phenotype compared with non-exposed DU145 cells [9].

To evaluate how the incorporation of stoichiometry into the gene-to-reaction formulation improved models' predictive capabilities, we applied some of the most widely used transcriptomic-based model-driven methods to integrate the transcriptomic data of Aldrin-exposed and non-exposed DU145 cells *via* either S-GPR or GPR [10-14]. The predicted and the experimentally measured metabolic consumption/production rates were compared (up to 244 uptake/secretion rates from different species), showing a significant improvement of model's predictions when S-GPR were implemented (up to 6% depending on the integration method). In addition, our computational analyses unveiled a marked metabolic reprogramming of key metabolic pathways in Aldrin-exposed cells. More specifically, carnitine shuttle, that transports long-chain fatty acids from cytosol to mitochondria and prostaglandine biosynthesis metabolisms were predicted to be significantly over-activated in Aldrin-exposed DU145 cells compared with control DU145 cells. Our computational analysis's predictions are consistent with reported metabolic changes associated to an enhanced malignant phenotype in prostate cancer and are supported by experimental observations [15].

Furthermore, this case of study is especially sensitive to the incorporation of stoichiometry into the gene-to-reaction formulation as the high isogeneity between Aldrin-exposed and non-exposed DU145 prostate cancer cells makes harder to infer metabolic changes from transcriptomic data. By contrary, the impact of adding stoichiometry to the gene-to-reaction associations may be disguise if more important differences exist in the transcriptomic profile between conditions. For instance, if important differences exist between conditions, considering the stoichiometry of a complex with several sub-units encoded by the same gene may not have any further effects due to the expression of the gene would unlikely become limiting at producing a catalytically active enzyme.

The novel S-GPR formulation that we have developed, incorporates stoichiometry to the gene-to-reaction associations enables a more precise integration of the transcriptomic data that ultimately enhances the predictive capabilities of the GSMM-based analyses. This new grasp outperformed previous approaches by significantly improving predictions of metabolism of multi-factorial diseases in a complex scenario. In addition, the novel S-GPR concept that we have developed, has the potential to be extended to proteomic data integration, which increases the impact that this new slant can have in key research areas such as environmental or clinical fields.

Acknowledgments

This work is supported by the Novo Nordisk Foundation (NNF10CC1016517 and NNF14OC0009473).

References

1. Roy PS, Saikia BJ (2016) Cancer and cure: A critical analysis. Indian J Cancer 53: 441-442.
2. Yabroff KR, Lund J, Kepka D, Mariotto A (2011) Economic burden of cancer in the US: Estimates, projections, and future research. Cancer Epidemiol Biomarkers Preval 20: 1-18.
3. Ghaffari P, Mardinoglu A, Nielsen J (2015) Cancer metabolism: A modelling perspective. Front Physiol 6: 1-9.
4. Thiele I, Palsson BØ (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat Protoc 5: 93-121.
5. Marin de Mas I (2018) Data Analysis for Omic Sciences: Methods and Applications. Multi-omic data integration and analysis via model-driven approaches, 1st (edn), Elsevier Publishers, United States.
6. Marin de Mas I, Esther A, Anusha J (2014) Cancer cell metabolism as new targets for novel designed therapies. Future Med Chem 6: 1791-1810.
7. Marin de Mas I, Torrents L, Bedia C, Nielsen LK, Cascante M, et al. (2019) Stoichiometric gene-to-reaction associations enhances model-driven analysis performance: Metabolic response to chronic exposure to Aldrin in prostate cancer. BMC Genomics 20: 652.
8. Official State Gazette (BOE) (2011) BO of E. Royal Decree 1799/2010, of December 30, which regulates the process of preparation and marketing of prepared waters packaged for human consumption. BOE 6292-6304.
9. Bedia C, Dalmau N, Jaumot J, Tauler R (2015) Phenotypic malignant changes and untargeted lipidomic analysis of long-term exposed prostate cancer cells to endocrine disruptors. Environ Res 140: 18-31.
10. Mardinoglu A, Agren R, Kampf C, Asplund A, Uhlen M, et al. (2014) Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. Nat Commun 5: 3083.
11. Becker SA, Palsson BO (2008) Context-specific metabolic networks are consistent with experiments. PLoS Comput Biol 4: e1000082.
12. Zur H, Ruppin E, Shlomi T (2010) iMAT: An integrative metabolic analysis tool. Bioinformat 26: 3140-3142.
13. Gonçalves E, Pereira R, Rocha I, Rocha M (2012) Optimization Approaches for the In-Silico discovery of optimal targets for gene over/underexpression. J Comput Biol 19: 102-114.
14. Jensen PA, Papin JA (2011) Functional integration of a metabolic network model and expression data without arbitrary thresholding. Bioinformat 27: 541-547.
15. Marin de Mas I, Aguilar E, Zodda E, Dallmann G, Thomson TM, et al. (2018) Model-driven discovery of metabolic vulnerabilities associated with epithelial-mesenchymal transition in prostate cancer. PLoS Comput Biol 14: e1005914.