Note on cancer genome-scale metabolic models

Francesco Gatto (gatto@chalmers.se), 2013-12-04
(Updated: 2014-02-19)

In this note, we give a brief overview on the published models and some tips on how to use these models for analysis and simulations. We refer the reader to the SI Appendix, SI Materials and Methods in the paper "*Chromosome 3p Loss of Heterozygosity Uniquely Shapes Clear Cell Renal Carcinoma Metabolic Network*" by F. Gatto, I. Nookaew, and J. Nielsen (doi: 10.1073/pnas.1319196111) for details on the reconstruction.

Overview on the models

There are 5 genome-scale metabolic models (GEMs) for cancer cells in the Human Metabolic Atlas extracted from the above-mentioned paper. These are: iBreastCancer1771, iLiverCancer1715, iLungCancer1472, iRenalCancer1410, and iUrothelialCancer1647. The models are provided in SBML format.

When using these models, the reader should bear in mind that they have been reconstructed using the following criteria:

1. The template GEM is the Human Metabolic Reaction (HMR) database from (Mardinoglu et al, 2013). At that time, HMR comprised 7,943 reactions, 3'158 unique metabolites and 3'674 genes.
2. Each GEM represents a primary cancer tissue**.** The resulting network therefore includes reactions from any cell type part of the solid tumor.
3. Each GEM is based on proteome evidence provided by the Human Protein Atlas (http://www.proteinatlas.org/), v. 11. Each GEM may contain reactions with no or little evidence if during the gap filling process such reactions were needed for connectivity or functionality reason.
4. Each GEM represents an average metabolic network for that cancer type. The median staining level for a reaction-encoding protein defines the evidence score for that reaction to be included in the network. As such, these GEMs are not generic cancer GEMs, but they are fair representation of that cancer type on average. They are well suited for comparative studies, but may also limit their applicability.
5. Each GEM collects data from a unique histological site, yet it typically encompasses more than one morphological type. For example, iLungCancer1472 ensembles both squamous cell and adenocarcinoma. Refer to the Human Protein Atlas for more details on the morphological sites of the samples.

Tips on simulations

As all GEMs, these cancer GEMs are simulation-ready models. We recommend the RAVEN Toolbox for MATLAB to simulate these GEMs (Agren et al, 2013). Each GEM can be simply imported using the function `importModel.m`. The network is balanced, connected and functional, that is it can carry flux in all elementary metabolic tasks (refer to SI Appendix in the reference paper). To facilitate the simulation of these models, the readers may benefit from the following instructions:

Biomass equation:

- The biomass equation is described in the SI Appendix, SI Materials and Methods. This is denoted as "CancerBiomass_OF" and produces one unit of "cancer_biomass[s]". The coefficients are derived assuming that metabolic fluxes are measured in terms of mmol g_{DW}^{-1} h^{-1}. Units of cancer_biomass are secreted by the exchange rxn "CancerBiomass_Ex". Make sure, when maximizing for biomass production, that both reactions are correctly set in terms of bounds.
- There is no growth-associated maintenance (GAM) in the biomass equation. This may however be estimated as 13.9 ATP [mmol g_{DW}^{-1}] and 11.1 GTP [mmol g_{DW}^{-1}] (Stephanopoulos et al, 1998).

- There is no non-growth associated maintenance (NGAM) in any GEM. This may however be estimated as 2.36 ATP [mmol g_{DW}^{-1}].
- In the uploaded version, a flux in the biomass equation is guaranteed when the list of metabolites in the appendix of this note are available for uptake. Note that other lists of metabolites were not tested, but the models can presumably grow in way less rich media.

Exchange reactions:

- Each GEM is closed. This is denoted by the presence of the field "unconstrained" in the model structure. It means that each exchange rxn is in the form "*metA*[s] ⇔ *metA*[x]". "x" is a fictitious outside compartment. Many constraint-based analysis require the network to be closed. Other analyses, such as FBA, work on the assumption that fluxes around a metabolites should be balanced. In this case, the "x" metabolites, that can only be produced or consumed, must be removed. A quick way to accomplish this is to use `simplifyModel.m` in the RAVEN Toolbox to do so.
- With respect to HMR, only exchange reactions that are needed to accomplish the different metabolic tasks are retained (these include biomass growth). Note that there are a variable number of new exchange reactions that are absent in HMR but added to fulfill the tasks. Evidence for all these reactions is taken from (Jain et al, 2012). If uptake, these reactions are denoted as "NewExRxn#", if secretion "NewDiffusion#".

Metabolites:

- Biomass lipids are represented by so-called pool metabolites. There are both catabolic and anabolic reactions to degrade and biosynthesize such pool metabolites. These reactions are inherently unbalanced (like the biomass equation) but are overall balanced when synthesis and degradation are coupled (i.e. the same amount of carbon is used to build a unit of lipid pool and is released when a unit of lipid pool is degraded). Lipid pool metabolites are identified by "pool" in their names.
- Glycogen is produced from the elementary polymerization reaction of UDP-glucose (1:1 stoichiometry).

Bounds: There are no artificial bounds in the network. If a reaction is irreversible, the lower bound is 0. If a reaction in unbounded, the corresponding bound is Inf.

References

1. Agren R, Liu L, Shoaie S, Vongsangnak W, Nookaew I, Nielsen J (2013) The RAVEN toolbox and its use for generating a genome-scale metabolic model for Penicillium chrysogenum. *PLoS computational biology* **9:** e1002980

2. Gatto F, Nookaew I, Nielsen J (2014) Chromosome 3p loss of heterozygosity is associated with a unique metabolic network in clear cell renal carcinoma. *PNAS* doi: 10.1073/pnas.1319196111

3. Jain M, Nilsson R, Sharma S, Madhusudhan N, Kitami T, Souza AL, Kafri R, Kirschner MW, Clish CB, Mootha VK (2012) Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. *Science* **336:** 1040-1044

4. Mardinoglu A, Agren R, Kampf C, Asplund A, Nookaew I, Jacobson P, Walley AJ, Froguel P, Carlsson LM, Uhlen M, Nielsen J (2013) Integration of clinical data with a genome-scale metabolic model of the human adipocyte. *Molecular systems biology* **9**

5. Stephanopoulos G, Aristidou AA, Nielsen JH (1998) *Metabolic engineering : principles and methodologies*, San Diego: Academic Press.

Appendix

1. List of metabolites that, if available for uptake, guarantee a flux in the biomass equation

Metabolite ID	Metabolite name [compartment]
m01965s	glucose[s]
m02387s	linoleate[s]
m02125s	histidine[s]
m02184s	isoleucine[s]
m02360s	leucine[s]
m02426s	lysine[s]
m02471s	methionine[s]
m02724s	phenylalanine[s]
m02993s	threonine[s]
m03089s	tryptophan[s]
m03135s	valine[s]
m02040s	H2O[s]
m02630s	O2[s]
m01596s	CO2[s]
m01307s	alanine[s]
m01369s	asparagine[s]
m01975s	glutamine[s]
m03101s	tyrosine[s]
m01628s	cysteine[s]
m01365s	arginine[s]
m01986s	glycine[s]
m02770s	proline[s]
m02896s	serine[s]
m01370s	aspartate[s]
m01974s	glutamate[s]
m02751s	Pi[s]
m02578s	NH3[s]
m02946s	sulfate[s]
m03120s	urate[s]
m01821s	Fe2+[s]
m02519s	Na+[s]
m02046s	HCO3-[s]
m02039s	H+[s]
m03157s	zinc[s]
m02200s	K+[s]
m01413s	Ca2+[s]
m01513s	choline[s]
m01797s	ethanolamine[s]
m01983s	glycerol[s]
m02658s	ornithine[s]
m02949s	sulfite[s]
m00970s	4-aminobutyrate[s]
m01822s	Fe3+[s]
m02332s	lactose[s]
m01401s	biotin[s]
m01253s	acetoacetate[s]
m02819s	pyruvate[s]
m02403s	L-lactate[s]
m02586s	nicotinate[s]
m02680s	pantothenate[s]
m01830s	folate[s]
m02174s	iodide[s]
m02588s	nitrite[s]
m01442s	chloride[s]
m01327s	alpha-tocopherol[s]
m01330s	alpha-tocotrienol[s]
m01935s	gamma-tocopherol[s]
m01938s	gamma-tocotrienol[s]
m02050s	hemoglobin[s]
m01368s	ascorbate[s]
m02982s	thiamin[s]
m02136s	homoserine[s]
m02440s	malonate[s]
m02661s	oxalate[s]
m01588s	citrulline[s]
m01633s	D-3-amino-isobutanoate[s]
m01279s	adenine[s]
m01280s	adenosine[s]
m01306s	AKG[s]
m01334s	AMP[s]
m01397s	bilirubin-bisglucuronoside[s]
m01398s	bilirubin-monoglucuronoside[s]
m01396s	bilirubin[s]
m01988s	glycocholate[s]
m02963s	taurocholate[s]
m01987s	glycochenodeoxycholate[s]
m02962s	taurochenodeoxycholate[s]
m01450s	cholesterol[s]
m01587s	citrate[s]
m01590s	CMP[s]
m01619s	creatine[s]
m02348s	L-carnitine[s]
m01615s	cortisol[s]
m01630s	cytidine[s]
m01668s	deoxycytidine[s]
m01666s	deoxyadenosine[s]
m01673s	deoxyuridine[s]
m01393s	betaine[s]
m02016s	GMP[s]
m02159s	hypoxanthine[s]
m02167s	IMP[s]

m02170s	inosine[s]
m02583s	nicotinamide[s]
m02769s	progesterone[s]
m02783s	prostaglandin D2[s]
m02785s	prostaglandin E1[s]
m02786s	prostaglandin E2[s]
m02897s	serotonin[s]
m02943s	succinate[s]
m02945s	sucrose[s]
m02961s	taurine[s]
m02997s	thymine[s]
m02996s	thymidine[s]
m02998s	thyroxine[s]
m03052s	triiodothyronine[s]
m02969s	testosterone[s]
m03114s	UMP[s]
m03118s	uracil[s]
m03123s	uridine[s]
m01621s	creatinine[s]
m01682s	D-glucitol[s]
m02659s	orotate[s]
m02926s	spermine[s]
m00674c	2-phospho-D-glycerate[c]
m01005c	4-hydroxyphenylpyruvate[c]
m00775c	3-hydroxyanthranilate[c]
m00913c	3-phospho-D-glycerate[c]
m00995c	4-hydroxybenzoate[c]
m03037c	trans-4-hydroxy-L-proline[c]
m01103c	5-hydroxyindoleacetate[c]
m02871c	SAH[c]
m01342c	anthranilate[c]
m00923c	3-ureidopropionate[c]
m01423c	carnosine[c]
m02738c	phosphocholine[c]
m02349c	L-cystathionine[c]
m01644c	dCMP[c]
m01690s	DHAP[s]
m01708c	dimethylglycine[c]
m01798c	ethanolamine-phosphate[c]
m01862c	fumarate[c]
m02912c	sn-glycerol-3-PC[c]
m01973c	glucuronate[c]
m01981c	glyceraldehyde[c]
m02914c	sn-glycerol-3-phosphate[c]
m01978c	glutathione episulfonium ion[c]
m02036c	guanidinoacetate[c]
m02133c	homocysteine[c]
m00788c	3-hydroxy-L-kynurenine[c]
m02183c	isocitrate[c]
m00990c	4-hydroxy-2-quinolinecarboxylic acid[c]
m02322c	L-2-aminoadipate[c]
m02319c	kynurenine[c]
m02439c	malate[c]
m02585c	nicotinate ribonucleotide[c]
m02660c	orotidine-5-phosphate[c]
m02696c	PEP[c]
m02772s	propanoate[s]
m02822c	quinolinate[c]
m02923s	spermidine[s]
m03109c	UDP-glucuronate[c]
m03148c	xanthine[c]
m03150c	xanthosine-5-phosphate[c]
m03149c	xanthosine[c]
m02577c	NG,NG-dimethyl-L-arginine[c]
m01313c	allantoin[c]
m01580c	cis-aconitate[c]
m02123c	hippurate[c]
m02479c	methylmalonate[c]
m02402s	lithocholate[s]
m02965s	taurolithocholate[s]
m02646s	oleate[s]