Rapid reconstruction of metabolic network models for human tissue metabolism: Application to liver metabolism

Livnat Jerby [1], Tomer Shlomi [2], Eytan Ruppin [1,3]

(1) School of Computer Science, Tel-Aviv University; (2) Computer Science Department, Technion; (3) School of Medicine, Tel-Aviv University.

A FEW QUESTIONS…

- What is the difference between eukaryote metabolism to prokaryote metabolism?
- What is the difference between the metabolism of a multi-cell organism and the metabolism of a uni-cell organism.
- What about cancer metabolism?
- Why should we nonetheless study human metabolism?

MBA – Model Building Algorithm ☺

Input: A. *A general (human) model*.

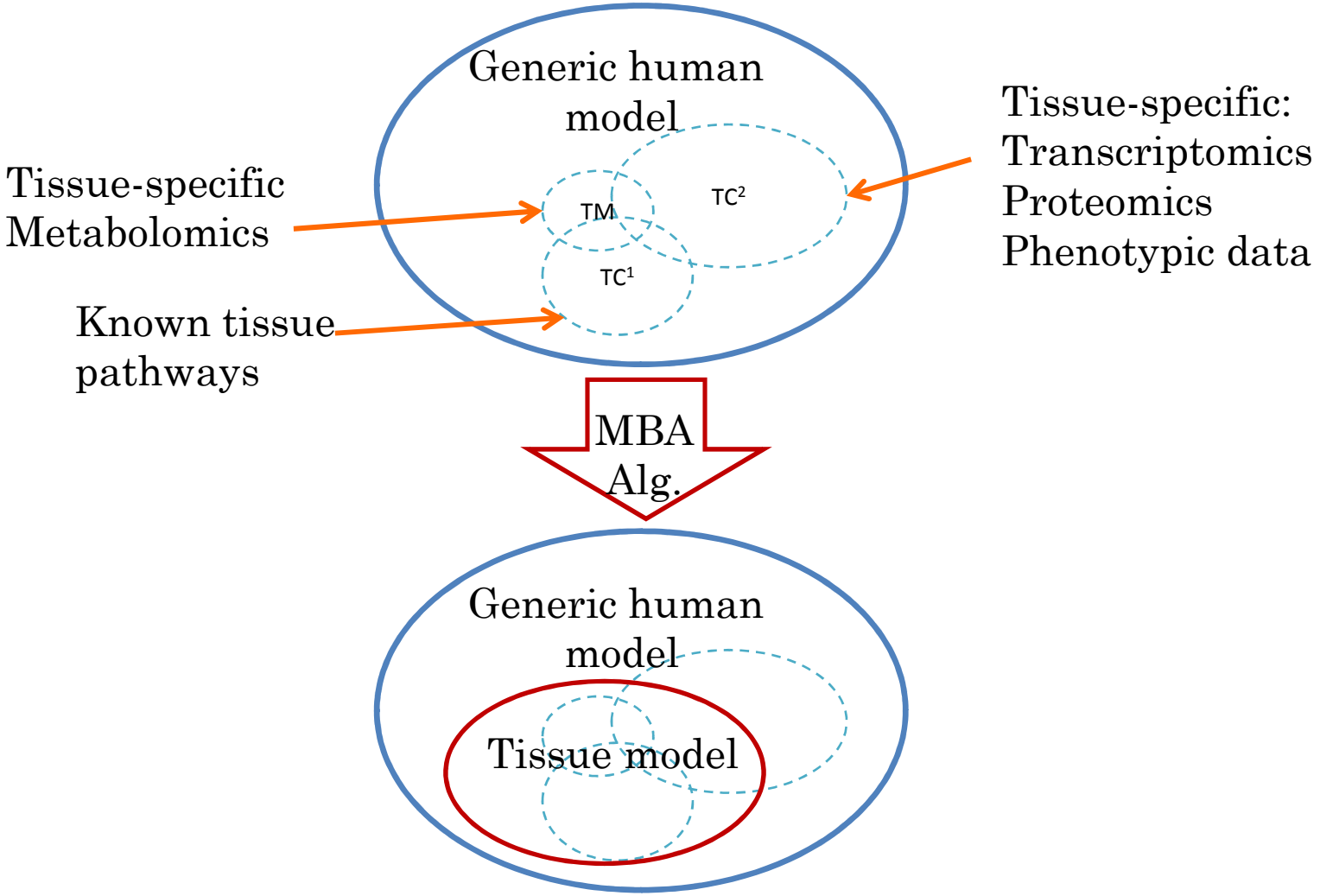
B. *Tissue Core (TC)* – an initial set of tissue-specific reactions composed from the different data sources. The *TC* is composed of two conceptually different sets of reactions: TC^1 (known biochemical pathway data) and the TC^2 (omics data) with high and moderate reliability levels, respectively.

Observation: the TC is not a viable/consistent model.

Output: additional tissue-specific reactions that enable the activation of the *TC,* and *together with the TC compose the tissue-specific model*.

- **The algorithm** starts from a global set of reactions (e.g., in our case, from the human model) and attempts to eliminate each of the reactions that *TC*, without disrupting the activation of the *TC (i.e.* a full activation of the TC^1 and a partial activation of the TC^2).
- The algorithm halts after all the potentially removable reactions have been scanned.

ONE PICTURE IS WORTH..



MBA Algorithm (II)

- The scanning order is random. It effects the resulting set, such that the output varies accordingly. Therefore the algorithm is executed with (1000) different, random elimination orders (each run results in a viable candidate model).
- **Aggregative model** – a model built from considering the essentiality results across all runs, by incrementally adding the most essential reactions until a viable model is obtained (the TC^1 and the TC^2 can be fully and a partially activated, respectively).

Speedup Techniques

- The time complexity of the naïve algorithm is computationally prohibitive.
- Acceleration techniques were implemented to circumvent this hurdle:
  - **On the *TC* side:** A recursive procedure aiming at activating multiple *TC* reactions together in the same LP problem, instead of each one at a time.
  - **On the elimination side:** In each attempt to eliminate a reaction set we examine whether the TC is still active. During that process we account for other potentially removable reactions that were inactive during the activation of the TC, and remove them as a set.

Applying MBA to build a liver model

- The liver is the main metabolic organ in the human body
- Involved in many important clinical conditions (e.g., obesity, diabetes, fatty liver, alcoholism etc.).
- The development of bioartificial liver (BAL) devices.
- Drug discovery and development.
- Because its there…

Liver Tissue-specific Data Sources

Tissue specific reactions are inferred from

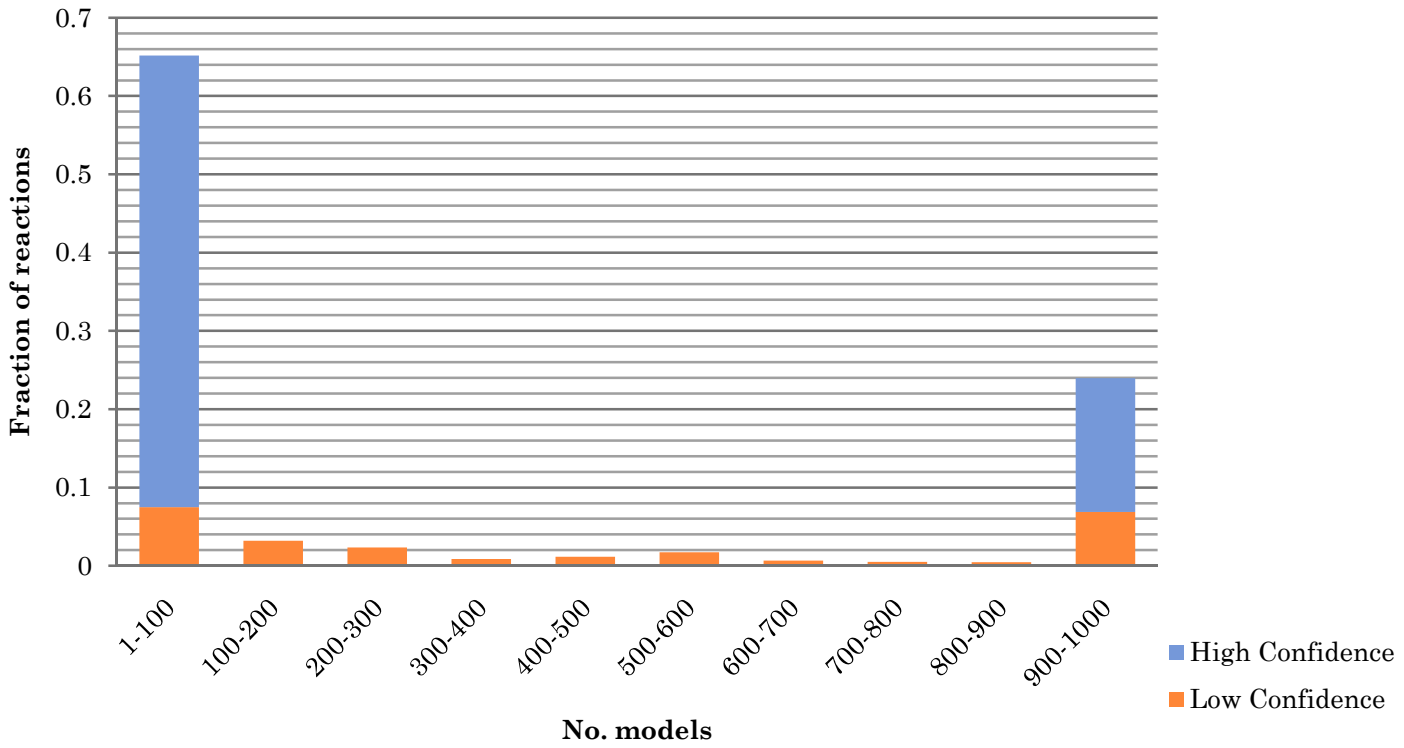
- Transcriptomic data [1]
- Proteomic data [2,4,5]
- Metabolic data [3]
- Phenotypic data[6]
- Literature-based knowledge[7].

A reaction is considered tissue specific only if it is inferred from at least two datasets.

1. Shmueli et al, 2003.
2. He, F., Human Liver Proteome Project: Plan, Progress, and Perspectives. *Molecular & Cellular Proteomics* 4 (12), 1841-1848 (2005).
3. Wishart, D.S. et al. HMDB: the Human Metabolome Database. *Nucleic Acids Res.* **35**, D521–D526 (2007).
4. Yan, Q. & Sadee, W. Human membrane transporter database: a Web-accessible relational database for drug transport studies and pharmacogenomics. *AAPS PharmSci* **2**, E20 (2000).
5. Saier, M.H., Jr., Tran, C.V. & Barabote, R.D. TCDB: the Transporter Classification Database for membrane transport protein analyses and information. *Nucleic Acids Res.* **34**, D181–D186 (2006).
6. McKusick, V.A. Mendelian Inheritance in Man and its online version, OMIM. *Am. J. Hum. Genet.* **80**, 588–604 (2007).
7. Karl Walter Bock, W. Gerok, S. Matern. (1991) *Hepatic metabolism and disposition of endo- and xenobiotics*. Germany: Kluwer Academic Publishers.

Validation I: Consistency of the Resulting Models

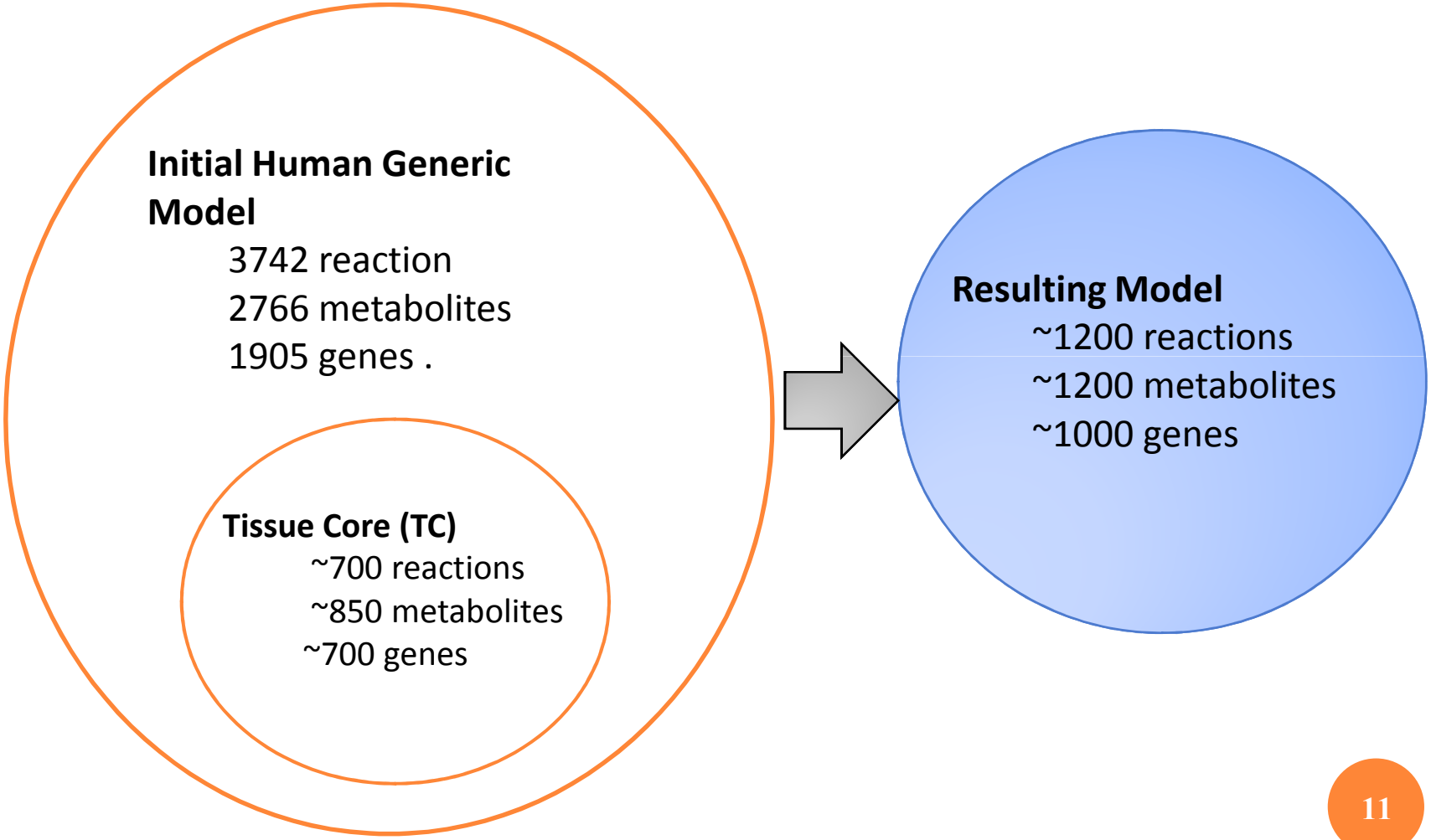
- For 76% of the reactions the algorithm provides a constant prediction regardless of the scanning order.



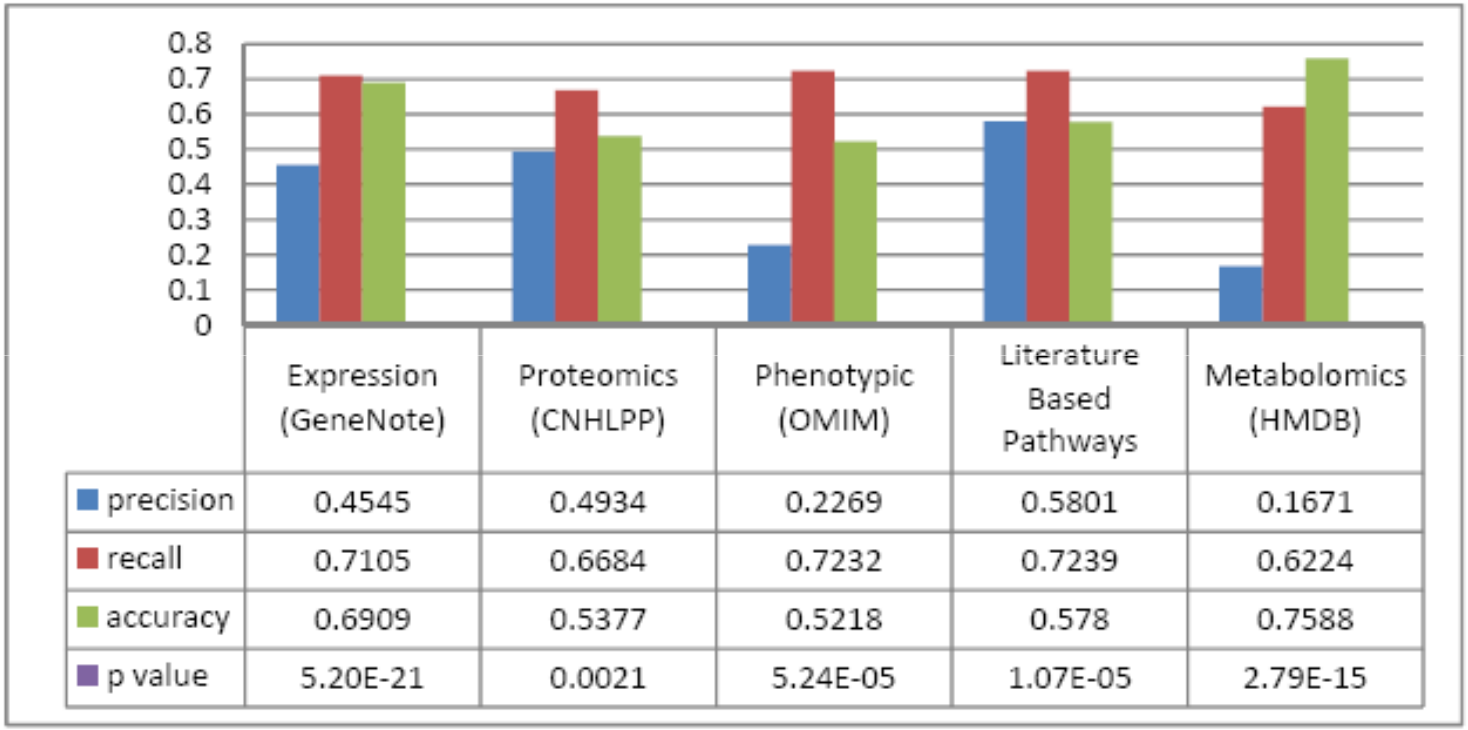
Validation II: Prediction of Missing Core Data

- The modeling process was repeated in a standard cross validation process 5 times, such that each time one of the datasets was omitted from the construction stage, and the model's prediction ability vs. the missing set was evaluated.
- The TC (TC=TC2) was formed from reactions testified by at least two of the remaining input sources.
- The evaluation was performed using the aggregative model described beforehand.

Typical Model vs. Initial Model

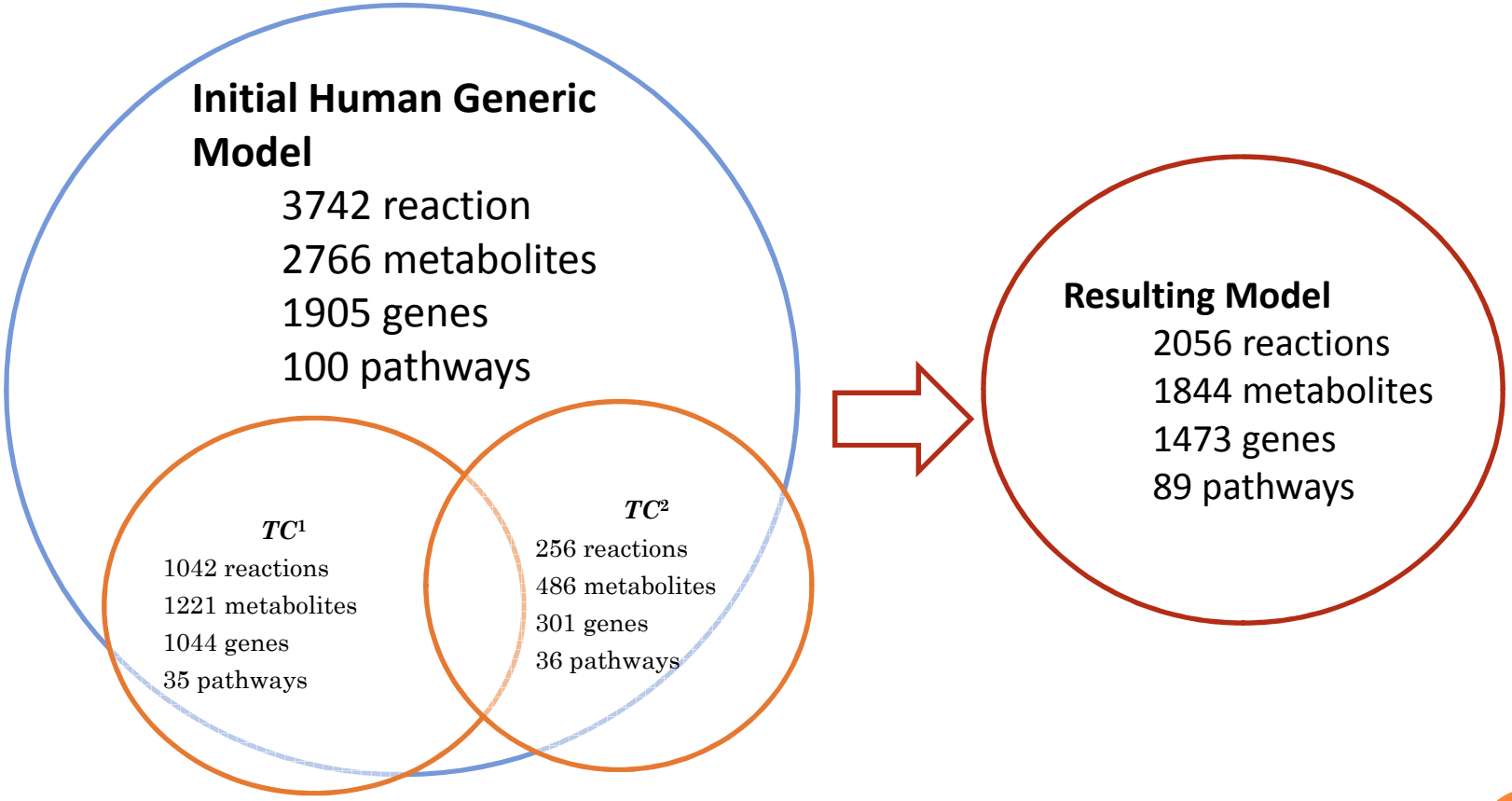


Validation II: Prediction of Missing Core Data



	Expression (GeneNote)	Proteomics (CNHLPP)	Phenotypic (OMIM)	Literature Based Pathways	Metabolomics (HMDB)
precision	0.4545	0.4934	0.2269	0.5801	0.1671
recall	0.7105	0.6684	0.7232	0.7239	0.6224
accuracy	0.6909	0.5377	0.5218	0.578	0.7588
p value	5.20E-21	0.0021	5.24E-05	1.07E-05	2.79E-15

Genome-Scale Metabolic Liver Model



VALIDATION III: MODEL FUNCTIONALITY

- The capacity of the model to carry out hepatic metabolic functions was assessed by simulating different metabolic scenarios:
  - Gluconeogenesis (i.e. the synthesis of glucose from glucogenic amino acids, lactate or glycerol)
  - Glycogenolysis (i.e. the breakdown of glycogen into glucose)
  - Glycogenesis (i.e. the formation of glycogen from glucose)
  - Cholesterol synthesis
  - Urea production.
- It is important to emphasize that even though the major hepatic metabolic pathways were already included in the TC it is not trivial that the model would have the capacity to perform more complex functions that involve the integration of several metabolic pathways.

VALIDATION IV
PREDICTING *IN-VITRO* FLUX MEASUREMENTS

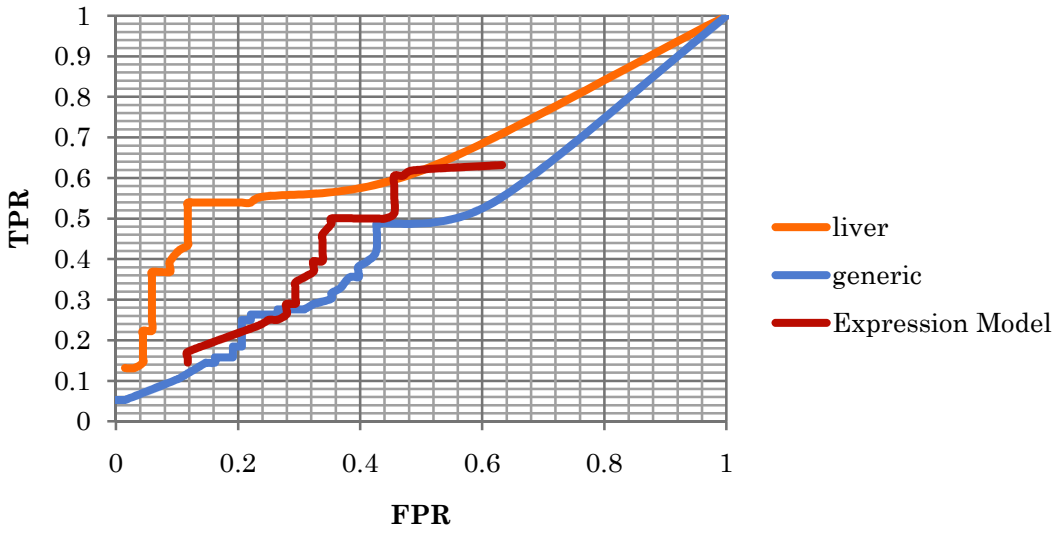
- Metabolic fluxes were measured in *rat cells in four conditions [1]:
  - Low-insulin preconditioning (L):
    - plasma (LP)
    - plasma+ amino acids (LPA)
  - High-insulin preconditioning (H):
    - plasma (HP)
    - plasma+ amino acids (HPA)
- The measured fluxes include:
  - 22 exchange fluxes (uptake of metabolites)
  - 20 inner fluxes.
- Both the human generic and the liver models were constrained using QP to depict the four metabolic states.

*Rat cells are the major component of bioartificial liver (BAL) devices.

1. Chan C, Berthiaume F, Lee K, Yarmush ML. Metabolic flux analysis of cultured hepatocytes exposed to plasma. *Biotechnol Bioeng* **81**(1), 33-49 (2003).

Flux Predictions

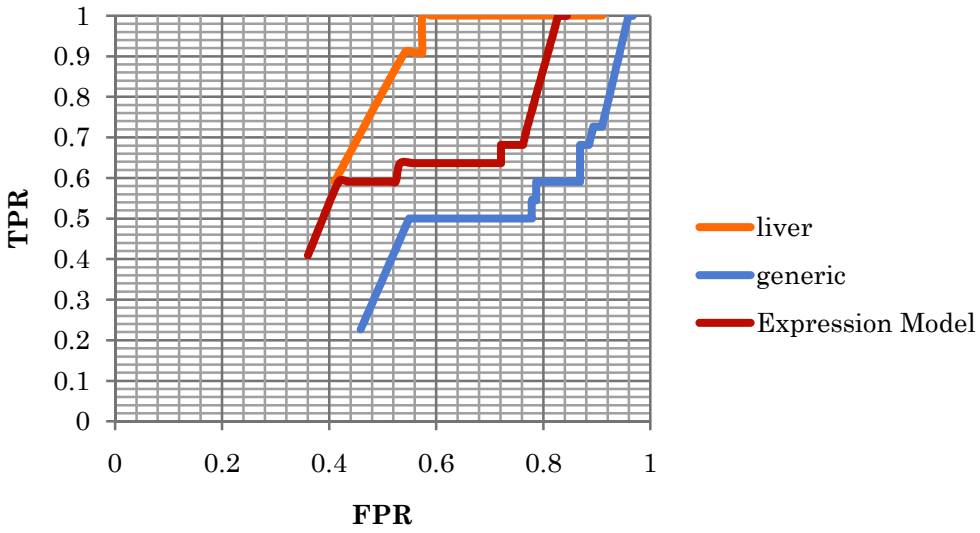
- ROC curve of predicting increasing fluxes:



- AUC
  - exp. Model 0.5321
  - Liver 0.672
  - Generic 0.4791

Flux Predictions

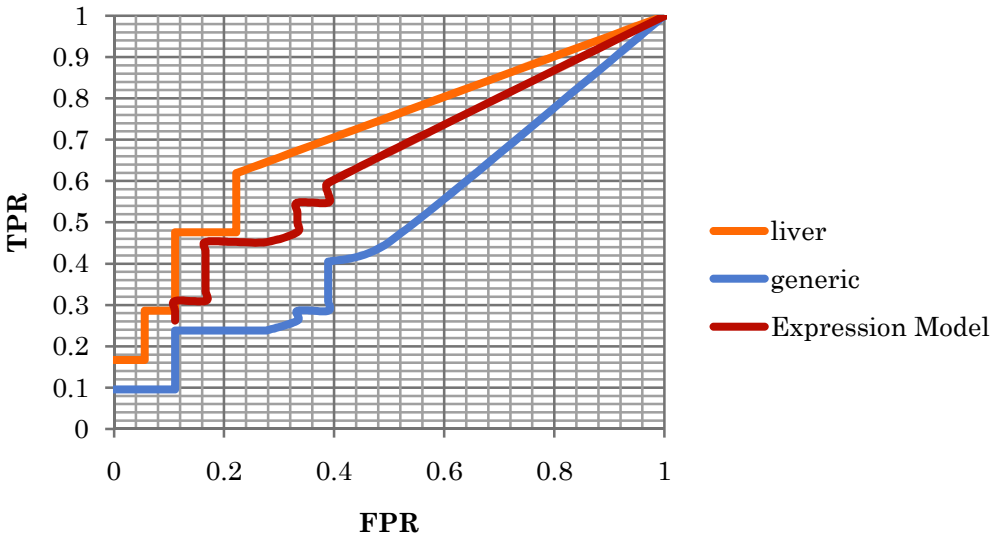
- ROC curve of predicting decreasing fluxes:



- AUC
  - Exp. Model 0.5458
  - Liver 0.7001
  - Generic 0.3797

Exchange Flux Predictions:

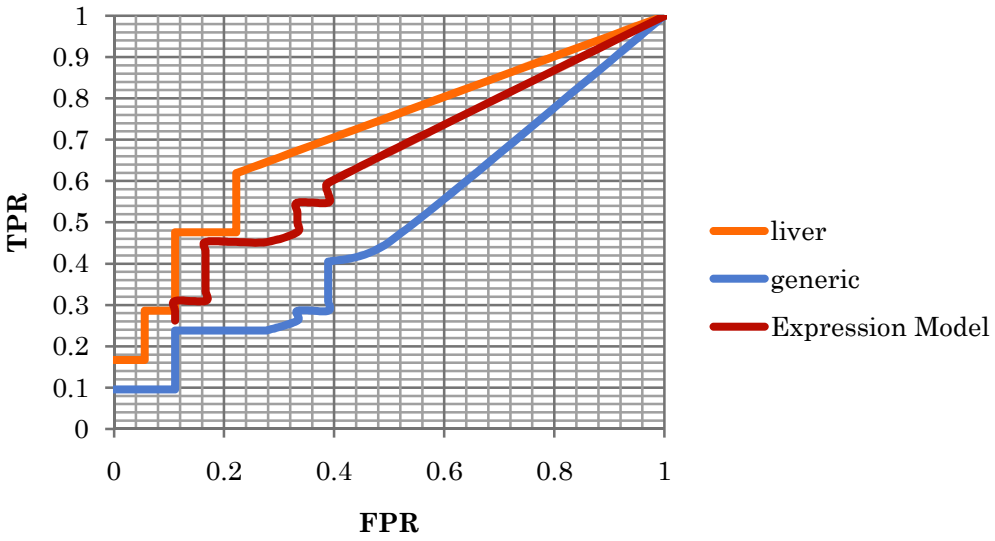
- ROC curve of predicting increasing fluxes (exchange fluxes):



- AUC
  - exp. Model 0.61376
  - Liver 0.6561
  - Generic 0.4874

Exchange Flux Predictions:

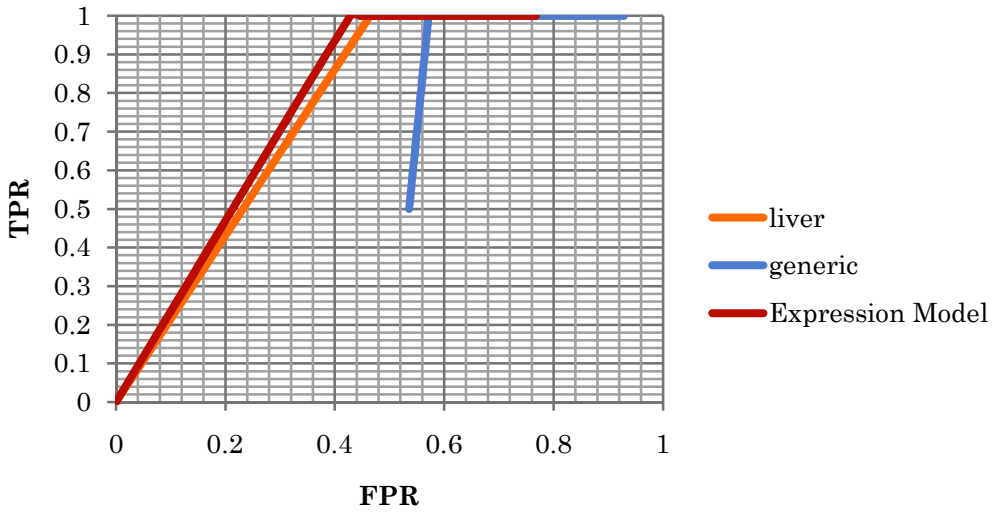
- ROC curve of predicting increasing fluxes (exchange fluxes):



- AUC
  - exp. Model 0.61376
  - Liver 0.6561
  - Generic 0.4874

Exchange Flux Predictions:

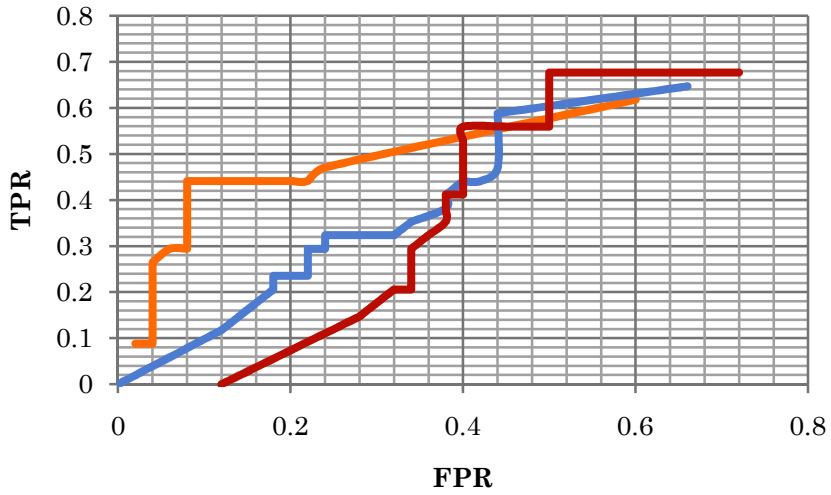
- ROC curve of predicting decreasing fluxes (exchange fluxes):



- AUC
  - Exp model 0.7857
  - Liver 0.7679
  - Generic 0.5893

Inner Flux Predictions:

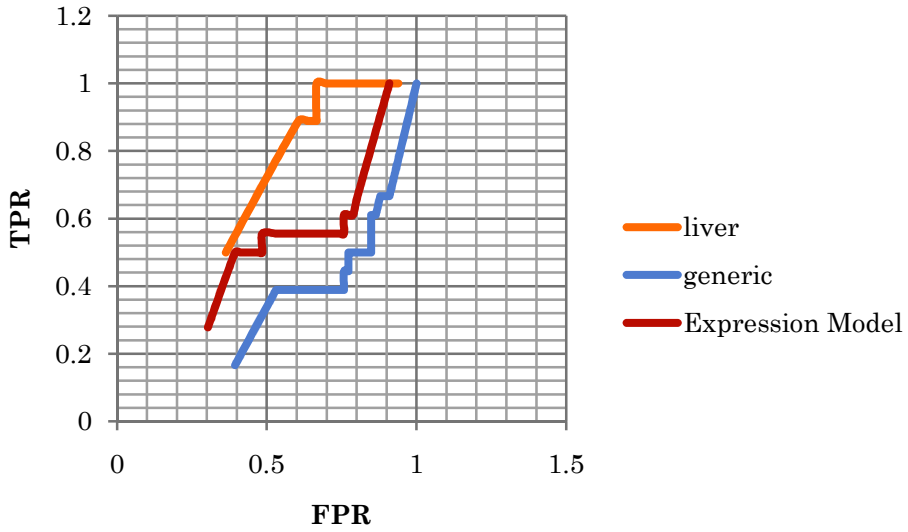
- ROC curve of predicting increasing fluxes (inner fluxes):



- AUC
  - Exp. Model 0.4806
  - Liver 0.6044
  - Generic 0.5103

Inner Flux Predictions:

- ROC curve of predicting decreasing fluxes (inner fluxes):



- AUC
  - Exp. Model 0.4819
  - Liver 0.6465
  - Generic 0.3186

Conclusions and Further Work

- Building and integrating tissue-specific models into a comprehensive *in-silico* model of human metabolism.
- Utilizing the model(s) to depict metabolic disorders and further investigate them.
- Apply the algorithm outside the scope of tissue-specific metabolism (e.g., given a metabolic archetype of bacterial metabolism, generating a model of specific bacteria species).