A generative statistical model implementing thermodynamic flux analysis

This study proposes a new method for incorporating thermodynamic information into an analysis of a metabolic network. We create a joint generative statistical model of measurements of four kinds of quantity: metabolic fluxes, metabolite concentrations, enzyme concentrations and Gibbs free energies of reaction. These are treated as interconnected, allowing for more precise estimates than would be possible with independent analyses.

Generative model

Fluxes are treated as determined by Gibbs free energies of reaction, enzyme concentrations and a latent parameter vector b, which can be interpreted as the amount of flux carried by each enzyme at steady state. The flux v_{ij} of reaction i in condition $\{\mathrm{j}\}$ is as follows:

$$v_{ij} = \Delta_r G_i \cdot e_{ij} \cdot b_{ij}$$

Gibbs free energies of reaction are treated as determined by metabolite concentrations and formation energies as follows:

$$\Delta_r G' = S^T(\Delta_f G' + RT \ln c)$$

Standard condition measurements of reaction gibbs free energies (as can be derived, for example, from the TECRDB database) and metabolic fluxes (as derived from fluxomics analysis) are represented using a standard linear regression model:

$$y_{\Delta_r G} \sim N(\Delta_r G, \sigma_{\Delta_r G})$$
$$y_v \sim N(v, \sigma_v)$$

Measurements of metabolite and enzyme concentrations, as derived from metabolomics and proteomics analyses, are represented using a lognormal generalised linear model:

$$y_c \sim LN(\ln(c), \sigma_c)$$
$$y_e \sim LN(\ln(e), \sigma_e)$$

This model is generative in the sense that, given an assignment of values to the unknown parameters $\Delta_f G$, c, e, b, $\sigma_{\Delta_r G}$, σ_v, σ_c and σ_e it is possible to simulate new values for the measured quantities $y_{\Delta_r G}$, y_v, y_c and y_e. The model therefore represents a theory as to how the observed data was generated. The theory can be tested both by comparing its predictions with real data and by assessing the plausibility of its parameters.

Contrast with traditional tfa

Traditional thermodynamic flux analysis (TFA) seeks to improve analyses of metabolic networks by taking advantage of information about the thermodynamic properties of the chemical reactions involved. TFA has historically been carried out within a constraint-based framework according to which the flux profile of a biological system is predicted by optimising an objective function representing the system's goals, subject to constraints imposed by the available information. For example, according to the mass balance constraint, metabolic fluxes must leave the system in a steady state.

In this framework thermodynamic information allows extra constraints to be imposed, representing the fact that the amount and direction of the flux a reaction carries is partly determined by its thermodynamic properties.