

doi:10.3969/j.issn.1672-5565.2016.03.06

整合转录组学数据的代谢网络研究进展

汪 慧,丁德武,孙 啸,谢建明*

(东南大学生物科学与医学工程学院,南京 210096)

摘要:高通量测序技术的快速发展催生了涵盖各层次细胞生命活动的组学数据,如转录组学数据、蛋白质组学数据和互作组学数据等。同时,全基因组代谢网络模型在不断完善和增多。整合组学数据,对生物细胞的代谢网络进行更深入的模拟分析成为目前微生物系统生物学研究的热点。目前整合转录组学数据进行全基因组代谢网络分析的方法主要以流量平衡分析(FBA)为基础,通过辨识不同条件下基因表达的变化,进而优化目标函数以得到相应的流量分布或代谢模型。本文对整合转录组学数据的FBA分析方法进行总结和比较,并详细阐述了不同方法的优缺点,为分析特定问题选择合适的方法提供参考。

关键词:代谢网络,转录组学,流量平衡分析,算法

中图分类号:Q493.2 **文献标志码:**A **文章编号:**1672-5565(2016)03-160-07

Development of integrating transcriptomic data into metabolic network analysis

WANG Hui, DING Dewu, SUN Xiao, XIE Jianming*

(College of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, China)

Abstract: With the advent of high-throughput technologies, the field of systems biology has amassed an abundance of developed metabolic network models and “omics” data, such as transcriptomic data, proteomic data and interactomic data. How to integrate omics data into metabolic network for further simulation analysis is becoming a hot spot of the microbial systems biology research. Several published studies have successfully demonstrated that the flux balance analysis(FBA), a constraint-based modeling approach, can be used to integrate transcriptomic data into genome-scale metabolic network model reconstructions to generate predictive computational models. In this review, we summarize such FBA-based methods for integrating expression data into genome-scale metabolic network reconstruction, highlighting the advantages as well as the limitations, and offer the suggestion to select appropriate method to a specific issue.

Keywords: Metabolic network ; Transcriptomics ; Flux balance analysis ; Algorithm

全基因组代谢网络(Genome-scale metabolic models, GEMs)是以基因组学数据为基础,在高通量测序技术支持下整合基因相关的组学信息,并结合文献中的知识,最终得到生物细胞中生化反应的全局网络。第一个全基因组代谢模型是细菌 *H. influenzae* 的代谢模型,构建于1999年,包括343个代谢物和488个代谢反应^[1]。截止2016年3月,在基因组代谢网络数据库(GSMNDB)中已发布了116个物种的135个全基因组代谢网络模型(<http://synbio.tju.edu.cn/GSMNDB/Pages/Models.htm>)。

全基因组代谢网络分析指利用相关代谢网络分析技术并结合相关生物学知识对代谢网络进行生物学分析的过程。其中预测细胞内特定功能生化反应流量分布并确定关键反应是全基因组代谢网络分析的重点。

代谢网络中代谢物质量数据获取的方法主要是通过碳同位素标记法,该方法对代谢物中的碳元素进行标记,使用质谱仪或者核磁共振仪对标记的代谢物进行检测,根据测量强度的变化得到细胞内的代谢流量分布^[2, 3]。此类实验数据的不断增多为

收稿日期:2016-04-18;修回日期:2016-06-07.

基金项目:国家自然科学基金项目(61472078)。

作者简介:汪慧,女,硕士研究生,研究方向:生物信息学;E-mail:m15150565077_2@163.com.

*通信作者:谢建明,男,副教授,研究方向:生物信息学;E-mail:xiejm@seu.edu.cn.

代谢网络流量的计算机模拟分析提供了基础。

1 代谢网络分析—流量平衡分析方法

代谢流量计算机模拟分析的经典方法是代谢流量平衡分析方法 (Flux Balance Analysis, 简称 FBA)^[4]。该方法假定代谢网络反应状态达到准稳态时每一个中间反应物产生与消耗的物质质量相等,如公式(1)所示,其中 x 代表代谢物的物质质量, t 代表特定时刻。

$$\frac{dx}{dt} = 0 \quad (1)$$

FBA 旨在找出准稳态条件下使细胞内某一目标函数 $Z = v_b$ 最优的可行解集。基于线性规划的原理,求解过程存在两种形式的约束条件,一种是等式约束,即在稳定状态下代谢物的输出与输入的代谢流量矢量和为 0。假设代谢网络中的每一个节点代表的是一个反应物,每一条线代表一个生化反应。设定 S 为网络的化学计量矩阵(以反应物为行,生化反应为列), v 为平衡时每一生化反应的流量。则平衡条件约束为公式(2)。

$$S \cdot v = 0 \quad (2)$$

另一种是不等式约束,即每一生化反应的代谢流量存在上下限 $v_{min} \leq v \leq v_{max}$ 。根据这两个约束条件可求得流量 v 的基本解空间,利用线性规划算法 (Linear Programming, LP) 或混合线性优化算法 (Mixed-integer linear programming, MILP) 计算 Z 达到最大值时整个网络的流量分布值。

FBA 方法自提出便成为研究代谢流量的基础方法,一般用来研究微生物的成长特性,解释生长表型以及预测特定环境下的关键反应^[5]。随着对代谢网络研究的日渐深入,研究者尝试在 FBA 方法中增加额外的生物学约束条件,以得到更为准确的预测结果。这方面的研究主要包括三个方面:一是改进目标函数,或者增加目标函数。典型的方法有最小化代谢调整方法 (Minimization Of Metabolic Adjustment, MOMA)^[6] 和最小调控控制方法 (Regulatory On /Off Minimization, ROOM)^[7];二是加入更多的约束条件(如,热力学信息),使得模拟过程更加接近于环境内部原则,如动态流量平衡分析 (Dynamic Flux Balance Analysis, DFBA)^[8] 方法, Willemsen 等人在 DFBA 的基础上进一步提出的修正方法——MetDFBA^[9]。三是多水平最优化,即同时改进目标函数和增加约束条件,如 OptStrain 方法和 OptKnock^[10] 方法等。

随着高通量测序技术的快速发展,对代谢网络

的认识越来越深刻,对代谢反应的调控机制分析也越来越深入。以全基因组代谢网络为基础,整合基因表达数据集对网络进行 FBA 分析的方法可以更精确的模拟细胞内代谢变化,提高代谢模型预测的准确性。

2 整合转录组学数据的代谢网络分析

基因表达在不同条件下呈现不同的表达状态^[11],而基因表达过程与代谢过程密切相关。基因表达数据反映基因的转录产物 mRNA 的相对或绝对丰度,有研究表明转录 mRNA 的浓度与生化反应催化剂酶的浓度有很强的相关性^[12, 13]。转录组学数据的获取方法主要有 cDNA 微阵列和寡聚核苷酸芯片技术等。目前,采用基因芯片技术已获得海量的转录组学数据。利用各种整合方法模拟基因表达与代谢反应之间的联系是该类方法研究的重点。

FBA 改进方法一般是在流量平衡约束基础上添加额外信息约束解空间或改变目标函数以更好的描述细胞代谢情况,并尝试通过多水平的最优化方法提高模拟的准确性。研究利用基因表达信息约束解空间一直是其中的焦点。2004 年第一次开始整合转录组学数据对代谢网络进行约束,近十年来研究者们提出各种与实际应用相结合的方法^[14-26],整合多组学数据对代谢网络进行 FBA 分析已经成为一种趋势。因而,本文总结了几种整合基因表达信息求解代谢流量的方法并分析了几种方法在判定基因表达状态时的不同策略。

2.1 设定先验阈值判定基因表达状态

Åkesson 等于 2004 年提出整合基因表达数据对全基因组代谢网络进行约束的方法,他们根据生化反应中酶所对应基因的表达情况判断反应是否发生^[27]。他所使用的表达数据来源于寡核苷酸芯片,检测到基因表达的信号则表示对应的生化反应发生,若反应相关的多个基因均不表达则对应反应不发生,反应流量值置为 0。通过这种方法进一步约束了 FBA 分析中流量 v 的解空间。该方法最早将多组学数据整合进代谢研究,为后续研究提供了一种新的思路。应用该方法,成功预测了葡萄糖为生长底物时酵母菌在无氧和有氧情况下的生长情况。但该方法判断基因是否表达的方式过于简单,同时因芯片的检测灵敏度较低而产生较多误差。

第一种整合方法由于数据处理粗糙,在判定基因表达状态时可能会遗漏表达量小但细胞完成功能必须的反应。根据特定条件下代谢功能的差异,Becker 提出了 GIMME (Gene Inactivity Moderated by

Metabolism and Expression)方法,模拟了细胞在特定条件下能够行使特定的功能^[28]。其思路是删除在特定条件下的不重要生化反应通路,同时增加一些功能反应通路,以得到特定条件下能完成特定功能的高效代谢网络模型。该方法可分为两个步骤:首先设定基因表达阈值,比较酶对应基因的表达值,将低于该阈值的酶催化的生化反应从代谢网络中删除,从而得到一个削减的代谢网络。第二步,将特定环境中细胞生长功能代谢模型(Required Metabolic Functionalities, RMF)中包含的反应加到第一步得到的模型中,得到初步的特定条件下的网络模型。使用积分函数计算不一致得分来评价不同表达阈值下所得到代谢流量分布是否合适,不一致得分值越小表示网络完成的功能越好。该方法的表达阈值需要用户提前设定并不断调整以得到最好的模拟结果。Schmid 等在此方法的基础上提出了改进的 GIM3E(Gene Inactivation Moderated by Metabolism, Metabolomics and Expression)方法^[29],同时考虑代谢网络中可逆生化反应的优化问题,并成功应用于预测沙门氏菌在不同生长环境下的代谢过程。

上述方法对基因表达阈值的选取进行评价并不断修正以拟合真实数据,Chandrasekaran 和 Price 对于酶基因是否表达的判断依据进行了改进,提出了代谢概率调控方法 PROM(Probabilistic Regulation Of Metabolism)^[30]。该方法分析全基因组基因表达微阵列数据,构建由转录因子及其调控基因构成的基因调控网络。由特定条件得到转录因子 B 的状态,并由先验知识设定表达阈值,计算同一条件下多个表达数据集得到基因 A 的状态,基因 A 表达且转录因子起作用的概率 $P(A = 1 | B = 1)$ 大于 0.5 则对应酶所催化的生化反应发生,如此判断各个反应的状态进而对代谢网络中的生化反应进行约束。该方法需要通过分析大量的实验数据来确定转录因子和靶基因表达的关系进而确定基因状态。与 Åkesson 等的方法相似,该方法本质上是在特定条件下删除网络中生化反应通路达到精简网络的目的。

上述方法共同点是均依赖于先验阈值的选取。其中 PROM 利用多个条件的均值增加阈值选择的可信度,GIMME 和 GIM3E 方法使用积分函数衡量阈值选取可信度。但初次选取的阈值对后续操作至关重要。另外一些方法使用数据之间的内在关系选取相对高表达和低表达的基因。这种方法可以在先验知识缺乏的情况下发挥作用,有一定的使用价值。

2.2 计算相对阈值判定基因表达状态

Hadas 等提出的 iMAT 方法比较全部基因表达数据并设置阈值(均值 \pm 方差)将基因表达值离散化

为低、中、高三种状态,分别用(-1/0/1)表示^[31]。并分别对三种状态的反应上下限增加约束,在提高高表达基因对应反应比例同时通过积分函数计算反应状态与表达数据相一致程度。最终找到与表达数据相一致的特定条件下的代谢网络。Rossell 等在 iMAT 方法基础上提出了 EXAMO(The EXploration of Alternative Metabolic Optima)^[32]方法,EXAMO 方法选定全部基因中表达量从高到低排列在前 15% 的为高表达基因,后 15% 的为低表达基因,其余为中等表达基因,并在 iMAT 方法基础上构建不同条件下的反应得分,尝试构建特定环境下的代谢模型。该方法还能最小化整体的流量之和并将对应流量都由高表达基因对应的生化反应所携带,进而预测特定环境下的流量。

上述两种方法对基因状态的判定是相对于整体基因表达情况,EXAMO 在先验知识缺乏的前提下构建特定代谢模型依赖于反应得分函数评价反应的判定与测得表达数据之间的相关性,期望找到最符合表达数据的表达情况。另外一些方法使用不同条件表达数据之间的比较关系得出不同条件下的可能代谢情况,例如 Jensen 等提出了 MADE 方法(Metabolic Adjustment by Differential Expression)^[33]。MADE 方法对不同条件下的表达数据进行 t 检验分析。设定必须发生反应对应状态为 1,利用不同条件下的表达差异显著性判断不同条件下的其他基因表达状态。得到基因表达状态后预测对应生化反应开启或者关闭,得到特定条件下的适应性网络模型,最终应用于 FBA 分析。本质上,MADE 方法通过考察不同条件下基因表达的差异程度判断基因的表达状态。使用该方法准确预测了酵母菌在葡萄糖生长环境下,在有氧和无氧生长条件的不同时间段基因表达的情况以及代谢模型的变化。

上述方法的共同点是使用基因表达数据之间相对比较的方法得到基因的表达状态,并进一步约束代谢网络。但也有差异,MADE 方法将计算得到的基因状态集状态为 0 的反应直接删除,可能会由于前期预处理过程导致后期删除反应的误差。iMAT 和 EXAMO 相对提高高表达基因集对应生化反应并降低低表达基因集对应生化反应,在一定程度上可以避免此类误差,但是准确性依赖于得分函数的选取与计算。

2.3 拟合函数法整合基因表达状态与代谢网络

最近几年新发展了一种使用判定函数的方法整合转录组学数据与代谢网络。有研究表明基因表达量与酶浓度之间的关系近似对数曲线^[13],使用特定函数描述基因表达水平与酶浓度的相关关系,将结

果用于约束代谢反应的流量上下限并对代谢网络的流量进行优化。Angione 等提出了 METRADE (MEtabolic and TRanscriptomics Adaptation Estimator)^[34]方法,该方法对于网络优化的工作比较复杂,在整合转录组学数据方面采用公式 3 表示的方法。

$$\varphi = \begin{cases} (1 + |\log(y_i)|)^{\text{sgn}(y_i-1)} & \text{if } y_i \in R^+ \setminus \{1\} \\ 1 & \text{if } = 1 \end{cases} \quad (3)$$

其中 $\text{sgn}(y_i - 1) = (y_i - 1) / |y_i - 1|$ 该函数中设定 y_i 为基因表达数据, φ 为相对应酶浓度,将函数值 φ 带入反应上下限 $v_i \geq v_i^{\min} \varphi$, $v_i \leq v_i^{\max} \varphi$ 并进行 FBA 分析得到特定环境下代谢流量的分布。Angione 等使用该方法精确模拟了大肠杆菌在不同环境下多个目标代谢物同时得到优化时的代谢流量分布情况。2016 年, METRADE 方法的提出者在原

方法基础上,在 $\log(y_i)$ 前面添加方差参数,进一步优化函数拟合的结果并运用该方法成功重建了大肠杆菌的代谢网络^[35]。

该类方法结合基因表达数据与代谢数据之间的关系构建函数尽力拟合真实的代谢情况,在大量数据的支持下能够实现高准确性的拟合,是一种新型的模拟方法,随着对转录与代谢之间模式的理解不断深刻,该方法将得到更广泛的应用。

2.4 方法总结与比较

整合基因表达数据进行代谢网络分析,其本质是对代谢网络的流量平衡分析添加约束条件。前面所述方法大部分是在进行 FBA 分析之前对网络进行调整,增加删除特定生化反应或调整网络结构以模拟特定条件下的真实情况。各种方法的简单示意如图 1 所示。

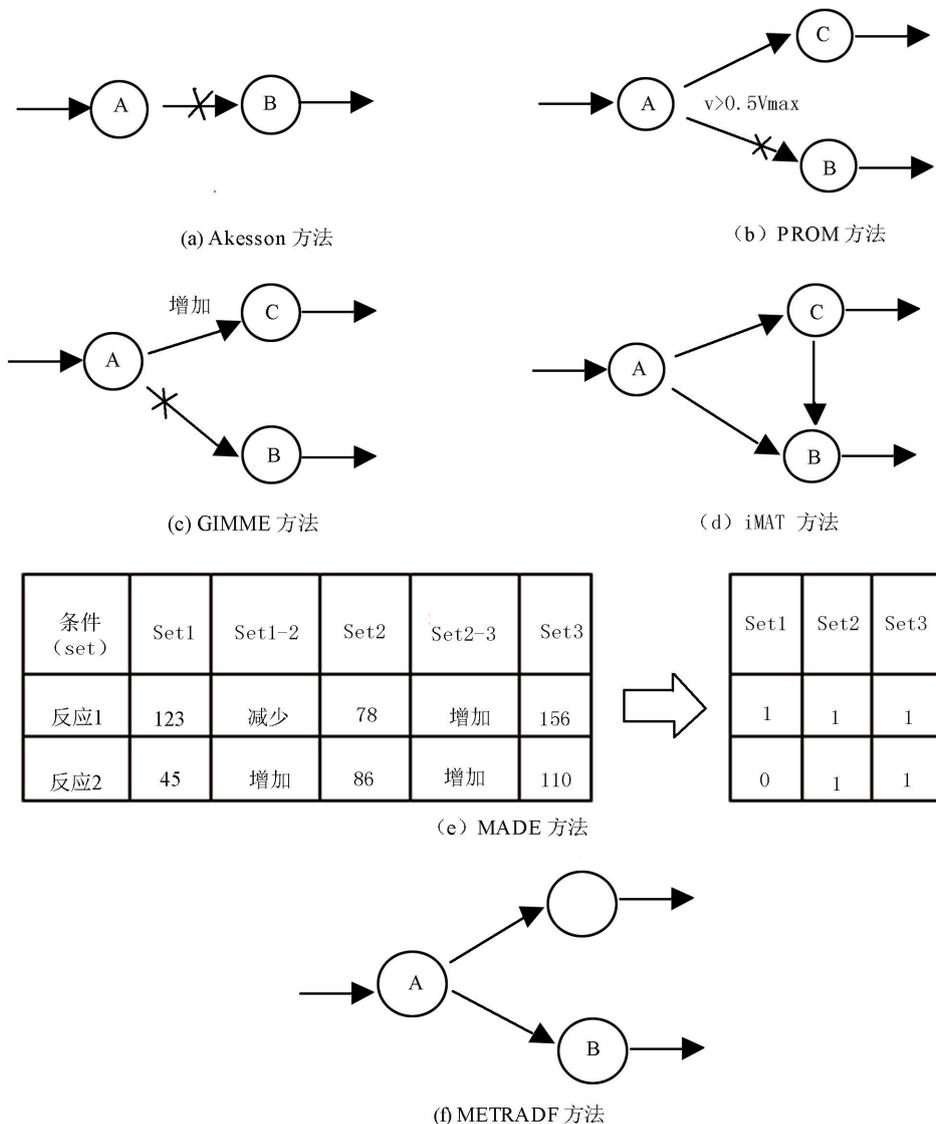


图 1 改进的 FBA 网络分析方法差异比较分析

Fig. 1 Comparative analysis diagram of improved FBA method

图中节点 A、B、C 代表代谢物,连接线代表生化反应。(a) Åkesson 方法,直接检测反应对应基因表达值, X 表示该反应从代谢网络中删除;(b) PROM 方法,计算多次表达值的平均值, $v > 0.5V_{max}$ 则反应发生,反之不发生;(c) GIMME 方法,设定阈值 y ,若 A→B 反应中酶对应基因表达值小于 y ,则删除该反应用 X 表示,并增加功能反应 A→C,得到功能网络模型,设定得分函数判断 y 的合理性;(d) iMAT 方

法,计算得到中-低表达阈值 y_1 ,中-高表达阈值 y_2 ,得到不同反应对应状态,高表达对应反应 rH,低表达对应反应 rL。提高 rH 比例,同时降低 rL 比例;(e) MADE,方法采用 t 检验方法得到不同条件下基因表达状态;(f) METRADE 方法,每一个生化反应的酶对应基因表达值不同时得到的酶的浓度不同,对应生化反应流量上下限不同,网络分析方法评价见表 1。

表 1 改进的 FBA 网络分析方法评价表
Table 1 Evaluation of improved FBA method

	表达数据	阈值设定区别基因状态	整合数据进行 FBA 分析	实现功能	matlab 算法包	特点	缺点
Åkesson et al	单一条件表达数据	基因表达大于 0, 基因表达状态即为 1。	基因表达状态为 0 则对应反应关闭	流量预测	无	首次提出	判定表达状态粗糙
GIMME GIM3E	单一条件表达数据	设定一个阈值并不断修正,大于阈值则基因状态为 1,阈值修订至积分函数最小	基因表达状态为 0 则将对应反应删除,加入功能反应	流量预测并构建功能模型	有	删除反应后添加功能通路构建功能模型	模型准确性依赖阈值选取
PROM	多个条件表达数据	多个条件下基因表达概率大于 0.5 则基因表达状态即为 1	基因表达状态为 0 则对应反应关闭	流量预测	有	概率模型判定反应状态	依赖阈值选取
iMAT EXAMO	单一条件表达数据	将基因表达值的均值±方差设定为区分基因表达状态的阈值,基因表达离散化为状态-1/0/1	提高基因表达状态为 1 的对应反应比例同时降低 -1 对应反应比例	流量预测并构建功能模型	无	数据之间比对待得出相对高低表达量	特定条件功能模型依赖得分函数
MADE	多个条件表达数据	比较不同条件下表达值并进行 t 检验,得到基因不同条件下状态变化情况	基因表达状态为 0 则对应反应关闭	流量预测并构建条件模型	有	T-test 区分不同条件表达差异	直接删除反应可能导致细胞功能障碍
METRADE	多个条件表达数据	无阈值设定	设定函数,输入表达数据,得到酶浓度,进一步约束 FBA 中反应流量	流量预测并构建功能模型	有	函数求解基因表达水平对代谢反应影响	准确度与数据量成正比

2.5 方法应用

本文介绍的方法代表整合转录组学研究代谢网络的大概发展思路,近年来各种整合转录组学数据进行代谢网络研究的实例涵盖微生物至人类的各个物种,2015 年 Shuyi 等利用 PROM 方法构建地核杆菌的转录调控代谢网络并指导实验操作^[36]。Tuulia 等利用 iMAT 方法研究人类肝脏肿大的代谢机制^[37]。Emrah 等使用 GIMME 和 MADE 方法构建胶

质母细胞瘤的特定全基因组代谢网络并预测癌症细胞代谢流量分布^[38]。在应用上述方法的同时,一些研究者也在探索整合多种组学数据进行代谢网络研究。如 Weihua 等在 FBA 基础上整合多个组学数据进行代谢网络研究并解释细胞表型^[39]。其他此类方面的应用还有很多^[40-47]。由此可推断以代谢网络为基础,用系统生物学观点整合多组学数据研究细胞内代谢过程正成为一种常用方法。

3 总结与展望

由于生物体内活动的复杂性,从系统生物学角度分析细胞内活动已经成为一种研究趋势,对细胞活动进行网络建模研究的方法也将愈加成熟。随着高通量数据的涌现,整合表达数据进入代谢网络已经成为改进 FBA 分析方法的主流思路。此类方法可分为两种模式,一种对代谢网络的结构进行直接调整,即通过表达数据判断相应的生化反应发生可能性,再进行 FBA 分析;另一种是在不改变原代谢网络的基础上,利用基因表达的数据和调控网络对优化过程进行限制。这些方法得出的结果大部分经过实验验证,有较大的应用价值。

但是本文所介绍方法整合数据相对单一,各方法均有其适用的局限性,方法准确度有待改进。目前的代谢网络研究也有待改进,首先,目前的代谢网络研究主体是整个细胞内网络,得到的是特定代谢途径的目标函数,细胞内的代谢子网络功能未被充分考虑;其次,基因表达与生化反应酶之间的关系比想象中复杂,多个基因共同调控生化反应的酶的机制目前尚不清楚。从 MATRADE 和 omFBA 方法^[39]看出,未来从系统生物学角度研究代谢过程将是一个研究热点,整合多个层面的组学数据将更精确模拟细胞内代谢活动,整合方法的研究也将不断发展。

参考文献(References)

[1] EDWARDS J S, PALSSON B O. Systems properties of the *Haemophilus influenzae* Rd metabolic genotype [J]. *Journal of Biological Chemistry*, 1999, 274(25):17410-17416.

[2] ZAMBONI N, FENDT S M, RÜHL M, et al. ¹³C-based metabolic flux analysis [J]. *Nature Protocols*, 2009, 4(6):878-892.

[3] GARCÍA MARTÍN H, KUMAR V S, WEAVER D, et al. A method to constrain genome-scale models with ¹³C Labeling data [J]. *PLoS Computational Biology*, 2015, 11(9):e1004363.

[4] ORTH J D, THIELE I, PALSSON B O. What is flux balance analysis? [J]. *Nature Biotechnology*, 2010, 28(3):245-248.

[5] CHAPMAN S P, PAGET C M, JOHNSON G N, et al. Flux balance analysis reveals acetate metabolism modulates cyclic electron flow and alternative glycolytic pathways in *Chlamydomonas reinhardtii* [J]. *Frontiers in Plant Science*, 2015(6):474. DOI: 10.3389/fpls.2015.00474. eCollection 2015.

[6] SHLOMI T, BERKMAN O, RUPPIN E, et al. Regulatory on/off minimization of metabolic flux changes after genetic perturbations [J]. *Proceedings of National Academy of Sciences of the United States of America*, 2005, 102(21):7695-7700.

[7] MAHADEVAN R, EDWARDS J S, DOYLE F R, et al. Dynamic flux balance analysis of diauxic growth in *Escherichia coli* [J]. *Biophysical Journal*, 2002, 83(3):1331-1340.

[8] GOMEZ J A, HOFFNER K, BARTON P I, et al. DFBAlab: a fast and reliable MATLAB code for dynamic flux balance analysis [J]. *BMC Bioinformatics*, 2014(15):409. DOI: 10.1186/s12859-014-0409-8.

[9] WILLEMSSEN A M, HENDRICKX D M, HOEFSLOOT H C, et al. MetDFBA: incorporating time-resolved metabolomics measurements into dynamic flux balance analysis [J]. *Molecular BioSystems*, 2015, 11(1):137-145.

[10] BURGARD A P, PHARKYA P, MARANAS C D, et al. Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization [J]. *Biotechnology and Bioengineering*, 2003, 84(6):647-657.

[11] 卢汀. 生物信息学基因表达差异分析 [J]. *生物信息学*, 2014, 12(02):140-144.
LU Ting. *Bioinformatics analysis for gene differential expression* [J]. *Chinese Journal of Bioinformatics*, 2014, 12(2):140-144.

[12] ANGIONE C, PRATANWANICH N, LIO P, et al. A hybrid of metabolic flux analysis and bayesian factor modeling for multiomic temporal pathway activation [J]. *ACS Synthetic Biology*, 2015, 4(8):880-889.

[13] VOGEL C, MARCOTTE E M. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses [J]. *Nature Reviews Genetics*, 2012, 13(4):227-232.

[14] MEGCHELENBRINK W, ROSSELL S, HUYNEN M A, et al. Estimating metabolic fluxes using a maximum network flexibility paradigm [J]. *PLoS One*, 2015, 10(10):e139665.

[15] PACHECO M P, JOHN E, KAOMA T, et al. Integrated metabolic modelling reveals cell-type specific epigenetic control points of the macrophage metabolic network [J]. *BMC Genomics*, 2015(16):809. DOI: 10.1186/s12864-015-1984-4.

[16] FERRAZZI F, MAGNI P, SACCHI L, et al. Inferring gene expression networks via static and dynamic data integration [J]. *Studies in Health Technology and Informatics*, 2006(124):119-124.

[17] SALAZAR M, VONGSANGNAK W, PANAGIOTOU G, et al. Uncovering transcriptional regulation of glycerol metabolism in *Aspergilli* through genome-wide gene expression data analysis [J]. *Molecular Genetics and Genomics*, 2009, 282(6):571-586.

[18] MOXLEY J F, JEWETT M C, ANTONIEWICZ M R, et al. Linking high-resolution metabolic flux phenotypes and transcriptional regulation in yeast modulated by the global regulator Gcn4p [J]. *Proceedings of National Academy of Sciences of the United States of America*, 2009, 106(16):6477-6482.

[19] COLIJN C, BRANDES A, ZUCKER J, et al. Interpreting expression data with metabolic flux models: predicting *Mycobacterium tuberculosis* mycolic acid production [J]. *PLoS Computational Biology*, 2009, 5(8):e1000489.

- [20] VAN BERLO R J, DE RIDDER D, DARAN J M, et al. Predicting metabolic fluxes using gene expression differences as constraints [J]. *IEEE/ACM Trans Computer Biology Bioinformatics*, 2011, 8(1):206–216.
- [21] LEE D, SMALLBONE K, DUNN W B, et al. Improving metabolic flux predictions using absolute gene expression data [J]. *BMC Systems Biology*, 2012(6):73. DOI: 10.1186/1752-0509-6-73.
- [22] KIM J, REED J L. RELATCH: relative optimality in metabolic networks explains robust metabolic and regulatory responses to perturbations [J]. *Genome Biology*, 2012, 13(9):R78.
- [23] COLLINS S B, REZNIK E, SEGRE D. Temporal expression-based analysis of metabolism [J]. *PLoS Computational Biology*, 2012, 8(11):e1002781.
- [24] TÖPFER N, JOZEF CZUK S, NIKOŁOSKI Z. Integration of time-resolved transcriptomics data with flux-based methods reveals stress-induced metabolic adaptation in *Escherichia coli* [J]. *BMC Systems Biology*, 2012, 3(6):148. DOI: 10.1186/1752-0509-6-148.
- [25] NAVID A, ALMAAS E. Genome-level transcription data of *Yersinia pestis* analyzed with a new metabolic constraint-based approach [J]. *BMC Systems Biology*, 2012, 3(6):150. DOI: 10.1186/1752-0509-6-150.
- [26] WANG Y, EDDY J A, PRICE N D. Reconstruction of genome-scale metabolic models for 126 human tissues using mCADRE [J]. *BMC Systems Biology*, 2012, 3(6):153. DOI: 10.1186/1752-0509-6-153.
- [27] AKESSON M, FORSTER J, NIELSEN J. Integration of gene expression data into genome-scale metabolic models [J]. *Metabolic Engineering*, 2004, 6(4):285–293.
- [28] BECKER S A, PALSSON B O. Context-specific metabolic networks are consistent with experiments [J]. *PLoS Computational Biology*, 2008, 4(5):e1000082.
- [29] SCHMIDT B J, EBRAHIM A, METZ T O, et al. GIM3E: condition-specific models of cellular metabolism developed from metabolomics and expression data [J]. *Bioinformatics*, 2013, 29(22):2900–2908.
- [30] CHANDRASEKARAN S, PRICE N D. Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in *Escherichia coli* and *Mycobacterium tuberculosis* [J]. *Proceedings of National Academy of Sciences of the United States of America*, 2010, 107(41):17845–17850.
- [31] SHLOMI T, CABILI M N, HERRGARD M J, et al. Network-based prediction of human tissue-specific metabolism [J]. *Nature Biotechnology*, 2008, 26(9):1003–1010.
- [32] ROSSELL S, HUYNEN M A, NOTEBAART R A. Inferring metabolic states in uncharacterized environments using gene-expression measurements [J]. *PLoS Computational Biology*, 2013, 9(3):e1002988.
- [33] JENSEN P A, PAPIN J A. Functional integration of a metabolic network model and expression data without arbitrary thresholding [J]. *Bioinformatics*, 2011, 27(4):541–547.
- [34] ANGIONE C, LIÓ P. Predictive analytics of environmental adaptability in multi-omic network models [J]. *Scientific Reports*, 2015(5):15147. DOI:10.1038/srep15147.
- [35] ANGIONE C, CONWAY M, LIÓ P. Multiplex methods provide effective integration of multi-omic data in genome-scale models [J]. *BMC Bioinformatics*, 2016, 17(Suppl 4): (83). DOI: 10.1186/s12859-016-0912-1.
- [36] MA S, MINCH K J, RUSTAD T R, et al. Integrated modeling of gene regulatory and metabolic networks in *Mycobacterium tuberculosis* [J]. *PLoS Computational Biology*, 2015, 11(11):e100454311. DOI: 10.1371/journal.pcbi.1004543.
- [37] HYÖTYLÄINEN T, JERBY L, PETÄJÄ E M, et al. Genome-scale study reveals reduced metabolic adaptability in patients with non-alcoholic fatty liver disease [J]. *Nature Communications*, 2016(7):8994. DOI: 10.1038/ncomms9994.
- [38] ÖZCAN E, ÇAKIR T. Reconstructed metabolic network models predict flux-level metabolic reprogramming in *Glioblastoma* [J]. *FRONTIERS IN NEUROSCIENCE*, 2016(10):156. DOI: 10.3389/fnins.2016.00156.
- [39] GUO W, FENG X. OM-FBA: integrate transcriptomics data with flux balance analysis to decipher the cell metabolism [J]. *PLoS One*, 2016, 11(4):e154188.
- [40] GARAY C D, DREYFUSS J M, GALAGAN J E. Metabolic modeling predicts metabolite changes in *Mycobacterium tuberculosis* [J]. *BMC Systems Biology*, 2015, 9(1):57. DOI: 10.1186/s12918-015-0206-7.
- [41] TÖPFER N, JOZEF CZUK S, NIKOŁOSKI Z. Integration of time-resolved transcriptomics data with flux-based methods reveals stress-induced metabolic adaptation in *Escherichia coli* [J]. *BMC Systems Biology*, 2012, 6(1):1–10.
- [42] AGREN R, BORDEL S, MARDINOGLU A, et al. Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT [J]. *PLoS Computational Biology*, 2012, 8(5):e1002518.
- [43] FANG X, WALLQVIST A, REIFMAN J. Modeling phenotypic metabolic adaptations of *Mycobacterium tuberculosis* H37Rv under hypoxia [J]. *PLoS Computational Biology*, 2012, 8(9):e1002688.
- [44] YIZHAK K, GABAY O, COHEN H, et al. Model-based identification of drug targets that revert disrupted metabolism and its application to ageing [J]. *Nature Communications*, 2013(4):2632. DOI:10.1038/ncomms3632.
- [45] YIZHAK K, GAUDE E, LE DEVEDEC S, et al. Phenotype-based cell-specific metabolic modeling reveals metabolic liabilities of cancer [J]. *eLIFE*, 2014, 21(3):e03641. DOI: 10.7554/eLife.03641.
- [46] STEMLER S, YIZHAK K, RUPPIN E. Integrating transcriptomics with metabolic modeling predicts biomarkers and drug targets for Alzheimer's disease [J]. *PLoS One*, 2014, 9(8):e105383.
- [47] AGREN R, MARDINOGLU A, ASPLUND A, et al. Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling [J]. *Molecular Systems Biology*, 2014(10):721. DOI: 10.1002/msb.145122.