

酿酒酵母糖酵解途径中酶量变化对乙醇浓度影响的模拟分析

Simulation and Analysis of Ethanol Concentration Response to Enzyme Amount Changes in *Saccharomyces cerevisiae* Glycolysis Pathway Model

孔德翀 杨雪莲 严明* 柳常青 许琳

KONG De-Chong, YANG Xue-Lian, YAN Ming*, LIU Chang-Qing and XU Lin

南京工业大学制药与生命科学学院 南京 210009

College of Life Science and Pharmaceutical Engineering, Nanjing University of Technology, Nanjing 210009, China

摘要 代谢组学是系统生物学的一个重要组成部分,应用相关方法获得了大量的数据。如何处理这些数据以及如何将这些数据与其他组学数据结合起来的问题不容忽视。在酶的反应动力学方程中引入“酶量倍数因子”能够解决其中的部分问题。如果反应动力学方程中酶的量发生变化,只需要改变相应的酶量倍数因子的数值。为了观察酿酒酵母糖酵解途径中酶量变化对乙醇浓度的影响,设定了高低两个酶量水平进行计算机模拟,对应的酶量倍数因子分别为 10 和 0.1。基于计算机模拟结果,使用聚类分析方法,12 种酶被分为两类。属于第一大类的四种酶 ADH、HK、PFK 和 PDC,均催化不可逆反应。第二大类 8 种酶中的 6 种 ALD、GAPDH、GlcTrans、IpPEP、PGI 和 TIM 均催化可逆反应。第二大类中另外两种酶 IpGlyc 和 PK 催化不可逆反应。按照这种方法,代谢组和蛋白质组数据能较容易地结合起来对系统作出较全面的分析。

关键词 代谢组学,酿酒酵母,糖酵解途径,酶量倍数因子,乙醇,计算机模拟,聚类分析

中图分类号 Q591 文献标识码 A 文章编号 1000-3061(2007)02-0332-05

Abstract Metabolome has become an important part of Systems Biology, and a large set of data has already gained by applying the methods of metabolome. How to deal with the data and how to combine data of metabolome with data of other omics are problems that can not be ignored. An Enzyme Amount Multiple Factor was imported into the enzyme kinetic equation. When the enzyme amount in the system changed, *in silico* model, it means to alter the Enzyme Amount Multiple Factor. In order to observe ethanol concentration response to enzyme amount changes in *S. cerevisiae* glycolysis pathway model, enzyme amount was separately set at high and low level, the corresponding Enzyme Amount Multiple Factor value was 10 and 0.1, relatively. Based on the result of simulation, twelve enzymes in pathway were separated into two classes, class I and class II by cluster analysis. The four enzymes belonging to class I, ADH, HK, PFK and PDC, all catalyze irreversible reactions. The six out of eight enzymes belonging to class II, ALD, GAPDH, GlcTrans, IpPEP, PGI and TIM, catalyze reversible reactions. The other two enzymes belonging to class II, IpGlyc and PK, catalyze irreversible reactions. Based on this method, data of metabolome and proteomics are easily integrated to accomplish relatively overall analysis of system properties.

Key words metabolomics, *Saccharomyces cerevisiae*, glycolysis pathway, Enzyme Amount Multiple Factor, ethanol, *in silico* simulation, cluster analysis

Received: September 6, 2006; Accepted: November 30, 2006.

This work was supported by the grants from National Major Fundamental Research Program of China (No.2003CB716000) and the National Natural Science Foundation of China Key Project (No.20336010).

* Corresponding author. Tel: +86-25-83587695; E-mail: yanming@njut.edu.cn

国家自然科学基金重点基金(No. 20336010)和国家重大基础研究项目基金(2003CB716000)资助

代谢组通常是指一个细胞、组织或物种内所有代谢物的总和^[1],包括分子化学类型的不同组合,即肽类、碳水化合物、脂类、核酸和外源物质的催化产物等^[2]。代谢组学是伴随着生物技术和计算科学的发展而形成的,并逐渐成为系统生物学的重要组成部分^[3]。目前,代谢组学的研究主要集中在揭示代谢网络和反应途径结构之间的关系与整合生化网络两方面工作上^[4]。传统的生物体代谢研究在模型构建上的简化使得生物体的本质特征并没有完全的表现出来,局限于代谢网络层面上进行分析。如果将酶在量上的改变作为一个因素考虑,那么可以把代谢组与蛋白质组的数据结合在一起,对整个体系做较为全面的分析。而酿酒酵母(*Saccharomyces cerevisiae*)作为模式生物,有着明晰的遗传信息,已经积累了大量的相关数据,适合作为研究的对象^[5]。

糖酵解途径作为生物化学研究中的核心部分之一,已经被广泛深入地进行了分析和讨论^[6]。通常人们将糖酵解途径看作一系列独立反应的集合来分别考虑,其中每一个生化反应都已经有了独立的详尽数学描述,同时这些数字和方程也与对应的生化反应的生物学意义关联了起来。在前人的工作已经积累了大量数据的前提下,如果将这些生化反应作为一个整体中的不同组成部分结合起来考虑,有可能会获得对整个糖酵解途径的一个新的认识。使用代谢组学的相关理论及方法能够解决其中的一部分问题。

在高通量技术盛行的今天,数据的获得已经不是一个难以逾越的鸿沟,面对着天文数字般的数据,传统的技术手段已经力不从心,采取新的思路和新的方法才能更有效的从中获得有价值的信息^[7]。聚类(clustering)这一名称广泛出现在各种新兴的方法中,主要描述使用简约模式将多变量数据进行分组的方法。聚类分析(cluster analysis)主要用来精简原始数据并将其可视化,或者用来预测新获得数据的大致分类。运用这一有力工具,研究人员便能够在短时间内处理大量的复杂数据^[8]。

本文描述的主要工作是将酿酒酵母糖酵解途径中催化反应的酶在量上的改变纳入到整个体系里,即在反应动力学方程中引入参数“酶量倍数因子”,模拟途径中12种酶在高低两种浓度水平下对最终产物乙醇产量的影响。提出了一种新的对酶进行分类的方法,并应用该方法对12种酶作了分类。

1 材料与方法

1.1 实验数据及软件工具

本文中模拟分析的酿酒酵母糖酵解途径原始实验数据来自 Hynne 等人的工作^[9]。酿酒酵母糖酵解途径模型的构建及糖酵解途径中酶量的变化对乙醇产量影响的模拟计算,使用免费软件 Cell DesignerTM v3.2^[10]、免费软件包 SBW (Systems Biology Workbench)^[11]和 SBML 常微分方程求解运行库 (SBML ODE Solver Library, SOSlib)^[12]。

1.2 SBML (Systems Biology Markup Language)

系统生物学标记语言(SBML)^[11]是一种能够将生化反应网络模型图形化的计算机可读格式。它可用于代谢网络、细胞信号网络、调控网络等各种生化网络类型的图形化显示。

1.3 酶量倍数因子(Enzyme Amount Multiple Factor, EAM Factor)的引入

酶量倍数因子用 E_k 表示,用来调节反应动力学方程中酶量的变化。

$$\text{米氏方程: } v = \frac{V_m \cdot [S]}{K_m + [S]}$$

$$\text{引入酶量倍数因子后: } v = \frac{E_k \cdot V_m \cdot [S]}{K_m + [S]}, 12 \text{ 种酶}$$

的反应动力学方程中 E_k 初始值均为 1。

1.4 聚类分析(cluster analysis)

1939年 Tryon 首先使用了“聚类分析”这个概念^[13]。聚类分析的目的是将类似的对象归入相应的类别中,它实际是一系列分类算法的总称。通常使用距离树图来表示各个对象之间的相互关系。

本文中用于聚类分析计算的欧几里德距离(Euclidean distance)是最为常用的距离计算方法。其本质是计算多维空间中的几何距离。计算方法如下: $\text{distance}(x, y) = [\sum_i (x_i - y_i)^2]^{1/2}$ 。

2 结果

2.1 模型的构建

构建的酿酒酵母糖酵解途径模型见图 1。模型中使用的缩略词具体含义见表 1。

2.2 模型的模拟计算

利用构建好的酿酒酵母糖酵解途径模型,考察途径中 12 种酶在高低两种浓度水平下对终产物乙醇胞外浓度的影响。模拟反应时间为 100min,初始模型的胞外乙醇积累的模拟数据曲线见图 2。由于

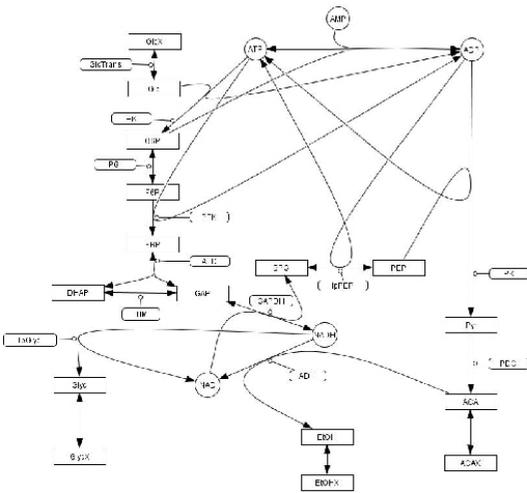


图1 重构的酿酒酵母糖酵解途径模型

Fig.1 Reconstructed *S. cerevisiae* glycolysis pathway model

初始模型是根据酿酒酵母处于稳态时的数据构建的,所以胞外乙醇的初始浓度并不为0。

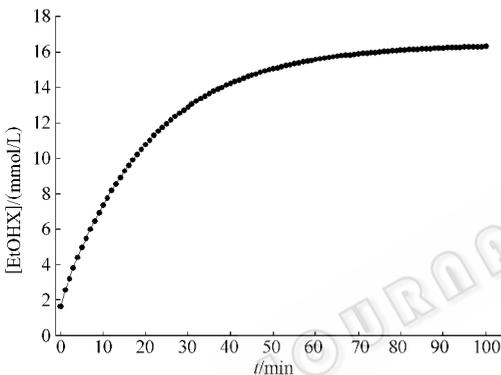


图2 酿酒酵母糖酵解途径模型乙醇积累曲线

Fig.2 Ethanol accumulation curve in original *S. cerevisiae* glycolysis pathway model
Simulation time was 100min.

分别将12种酶的浓度变为原浓度的0.1倍和10倍(即将酶量倍数因子 E_k 的值分别变为0.1和10)后进行模拟,得到胞外乙醇积累的数据,计算其胞外乙醇终浓度相比初始模型的胞外乙醇终浓度的变化值(见表2)。

酿酒酵母糖酵解途径中12种酶的量减少到原来的0.1倍时,ADH、HK、PDC和PFK这四种酶产生的影响最大,均使乙醇产量的降低超过了80%;ALD产生的影响最小,乙醇产量仅降低了0.26%。

酿酒酵母糖酵解途径中12种酶的量增加到原来的10倍时,lpGlyc对乙醇产量影响最大,使其降低了16.82%;ALD、lpPEP、PDC、PFK、PGI、PK和TIM对乙醇产量变化的影响均不超过0.5%。

表1 酿酒酵母糖酵解途径模型中使用的酶与组分名称缩略语列表

Table 1 A list of enzymes and compounds abbreviations in *S. cerevisiae* glycolysis pathway model

Abbreviation	Name
Enzyme	
GlcTrans	Glucose transport reaction
HK	Hexose kinase
PGI	Phosphoglucose isomerase
PFK	Phosphofruktokinase-1
ALD	Aldolase
TIM	Triose phosphate isomerase
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
lpPEP	lumped Phospho-glycerate kinase reaction, Phospho-glycerate mutase reaction, Enolase reaction
PDC	Pyruvate decarboxylase
ADH	alcohol dehydrogenase
lpGlyc	lumped glycerol 3-phosphate dehydrogenase reaction, glycerol phosphatase reaction
Metabolite	
GlcX	extracellular glucose
Glc	intracellular glucose
G6P	glucose-6-phosphate
F6P	fructose-6-phosphate
FBP	fructose-1,6-bisphosphate
DHAP	dihydroxyacetone phosphate
GAP	glyceraldehyde-3-phosphate
BPG	glycerate-1,3-diphosphate
PEP	phosphoenolpyruvate
Pyr	pyruvate
ACA	intracellular acetic acid
ACAX	extracellular acetic acid
EtOH	intracellular ethanol
EtOHX	extracellular ethanol
Glyc	intracellular glycerol
GlycX	extracellular glycerol

2.3 12种酶的分类

根据得到的胞外乙醇终浓度相比初始模型的胞外乙醇终浓度的变化值,计算欧几里德距离进行聚类分析,并根据结果绘制距离树(见图3酿酒酵母糖酵解途径模型中12种酶的单连接欧几里德距离树)。观察距离树可知12种酶被分为两大类。

ADH、HK、PDC和PFK属于第一大类;GAPDH、GlcTrans、PGI、TIM、lpPEP、ALD、lpGlyc和PK属于第二大类。

表 2 酿酒酵母糖酵解途径模型中乙醇浓度对酶量变化的响应

Table 2 The response of ethanol production to enzymes amount changes in *S. cerevisiae* glycolysis pathway

Enzyme amount multiple factor	ADH	ALD	HK	GAPDH	GlcTrans	lpGlyc	lpPEP	PDC	PFK	PGI	PK	TIM
0.1	-85.02%	-0.26%	-85.61%	-17.86%	-45.80%	2.59%	-1.18%	-88.78%	-81.94%	-3.86%	-44.04%	-25.95%
10	9.57%	0%	2.57%	1.83%	1.16%	-16.82%	0.10%	0%	0.20%	-0.24%	-0.44%	0.49%

When Enzyme Amount Multiple Factor of 12 enzymes in yeast glycolysis pathway were separately set to 0.1 and 10, the change of end product ethanol extracellular concentration were calculated, and were shown in percentage.

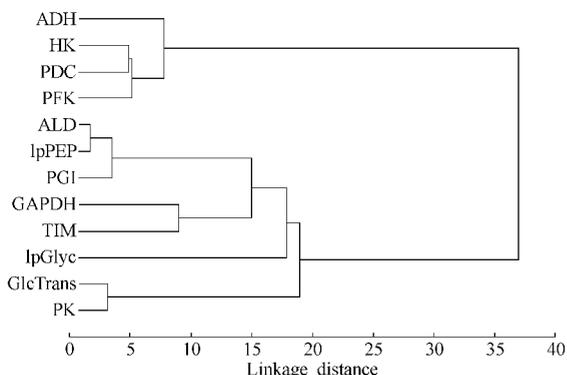


图 3 酿酒酵母糖酵解途径模型中 12 种酶的单连接欧几里德距离树

Fig. 3 Single linkage Euclidean distance tree of 12 enzymes in *S. cerevisiae* glycolysis pathway model

3 讨论

糖酵解途径中的 12 种酶按照对乙醇产量的影响的聚类分析分为两大类,结合这 12 种酶催化的反应类型,发现第一大类所包括的 4 种酶 ADH、HK、PDC、PFK 所催化的反应均为不可逆反应,第二大类总共 8 种酶中的 6 种, GAPDH、GlcTrans、PGI、TIM、lpPEP、ALD 所催化的反应均为可逆反应,特殊的 2 种酶是 lpGlyc 和 PK,它们所催化的反应是不可逆反应。

第二大类中催化不可逆反应的酶 lpGlyc 和 PK 在量上的改变对乙醇产量的影响,与第一大类中的 4 种酶在量上的改变对乙醇产量的影响不相同的原因可能是由于:

(1) lpGlyc 催化的是甘油生成途径上的关键反应,因此它在量上的变化必然导致甘油生成量的变化,同时需要消耗一定量的 DHAP,如果 lpGlyc 的量增高, DHAP 消耗增多,胞内浓度降低,在 TIM 的作用下,一部分的 GAP 会转化成 DHAP 来补偿 DHAP 的消耗,这就导致了乙醇生成途径上的物质流量降低,并最终使得乙醇终浓度降低,如果 lpGlyc 的量降低, DHAP 形成积累,在 TIM 的作用下,一部分 DHAP 会转化成 GAP,使得乙醇生成途径上物质流量升高,

乙醇的终浓度升高。

(2) PK 催化的反应有能量物质参与,与第一大类中催化反应有能量物质参与的酶 ADH、HK、PFK 不同的是,这三种酶参与的反应均为消耗能量物质的反应,是从相对较高的自由能变为相对较低的自由能,而 PK 参与的反应为生成能量物质的反应,是从相对较低的自由能变为相对较高的自由能,这可能是其对乙醇终浓度的影响与其他三种酶对乙醇终浓度的影响有差别的原因。

4 结论

代谢组学研究正在生命科学中扮演越来越重要的角色,如何充分利用代谢组学的相关数据,与其他组学研究相结合,是目前需要解决的一个重要问题。将“酶量倍数因子”参数引入反应动力学方程,通过观察酶在量上的改变对整个系统的影响,使得把代谢组数据和蛋白质组数据结合起来成为可能。

根据构建的酿酒酵母糖酵解途径模型,计算模拟得到了酿酒酵母糖酵解途径中 12 种酶在高低两种浓度水平(初始浓度的 10 倍, 0.1 倍)下对乙醇产量影响的数据。

对上述结果进行了数据分析,根据分析结果,12 种酶被分为了两个大类,第一大类为不可逆反应,第二大类中主要为可逆反应,在第二大类中有 2 种特殊的酶 lpGlyc 和 PK,它们均催化不可逆反应,对这种特殊现象做出了解释与讨论。通过改变酶的量来观察整个系统的变化,并使用聚类分析的方法对酶进行了分类,这是一种新的酶的分类方法,这种新方法对网络特性的分析提供了一种有价值的思路。

REFERENCES (参考文献)

- [1] Fiehn O. Metabolomics—the link between genotypes and phenotypes. *Plant Mol Biol*, 2002, 48(1–2): 155–171.
- [2] Saghatelian A, Cravatt BF. Global strategies to integrate the proteome and metabolome. *Curr Opin Chem Biol*, 2005, 9(1): 62

- [3] Rochfort S. Metabolomics reviewed: a new “ omics ” platform technology for systems biology and implications for natural products research. *J Nat Prod*, 2005, **68**(12): 1813 – 1820.
- [4] Weckwerth W. Metabolomics in systems biology. *Annu Rev Plant Biol*, 2003, **54**: 669 – 689.
- [5] Maraz A. From yeast genetics to biotechnology. *Acta Microbiol Immunol Hung*, 2002, **49**(4): 483 – 491.
- [6] Teusink B, Passarge J, Reijenga CA, et al. Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. *Eur J Biochem*, 2000, **267**(17): 5313 – 5329.
- [7] Shulaev V. Metabolomics technology and bioinformatics. *Brief Bioinform*, 2006, **7**(2): 128 – 139.
- [8] Slonim DK. From patterns to pathways: gene expression data analysis comes of age. *Nat Genet*, 2002, **32**(Suppl): 502 – 508.
- [9] Hynne F, Dano S, Sorensen PG. Full-scale model of glycolysis in *Saccharomyces cerevisiae*. *Biophys Chem*, 2001, **94**(1 – 2): 121 – 163.
- [10] Kitano H, Funahashi A, Matsuoka Y, et al. Using process diagrams for the graphical representation of biological networks. *Nat Biotechnol*, 2005, **23**(8): 961 – 966.
- [11] Hucka M, Finney A, Sauro HM, et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 2003, **19**(4): 524 – 531.
- [12] Machne R, Finney A, Muller S, et al. The SBML ODE Solver Library: a native API for symbolic and fast numerical analysis of reaction networks. *Bioinformatics*, 2006, **22**(11): 1406 – 1407.
- [13] Tryon RC. Cluster Analysis; Correlation Profile and Orthometric (factor) Analysis for the Isolation of Unities in Mind and Personality. Edwards brother, inc., litho printers and publishers, 1939.