《生物科学杂志》

Journal of Biological Sciences

ChengZhu Science™

江西省诚筑环保工程有限公司主办

2022 年 11 月刊物/Serial in November, 2022

出版人： 刘焕 香江出版社有限公司

Publisher: Liu Huan, Xiangjiang Publishing Company Ltd.

Copyrights Statements

Copying and Transferring is Forbidden!

版权申明

禁止复制、转载！

Article 3: Metabolomics (1) --- The Systematic Chemistry Fingerprints Between Genotype and Phenotype and its Application on the Conservation Genetics /新陈代谢组学---连接基因型和表现型的一项系统化学指纹识别技术与在保护遗传学中的应用

Author: Liu Huan (1983-), Master of Science (First Class Honours, 2009), The University of Auckland

DOI:10.58473/JBS0005 Retrieval from official database: www.crossref.org

Latest revised on 29/05/2023.

1. Introduction

DNA should be defined as the main carrier of material genetic information. Because cells are intelligent creatures, in addition to material genetic information, there are also memory of spirit passed on to the offspring cells in the genetic process. This has been discussed in my previous paper. After genetics, complex metabolic processes begins in cells, which can be divided into three categories in this article: primary metabolic molecules (such as tRNA), secondary metabolic biochemistry molecules (such as glutamic acid discussed in plant physiology paper), and the final metabolites (such as polysaccharide compounds). However, enzyme is the indispensable biochemicals in this metabolic process. Enzyme is the principal biochemical regulator to initiate the biochemical process of metabolism in cell biology. The primary and secondary metabolites can be significantly influenced by the cellular physiological environment and genetics, whereas the cell functions constantly relies on the regulator of isozyme family (there are various enzyme species within one isozyme family) initiating the pathways of various metabolite process regardless of environmental changes and genetics. For example, in my previous article, it is concluded that the isozyme spectrum correspondingly to the valid antibiotics of specific pathogens must be relatively specific and unique, regardless of DNA genetics and environmental changes between different host cells. The thinking ability of multi-cellular individuals of higher intelligent is limited to three-dimension, as the thinking ability of single-cell organisms should only stay in two dimension: the recognition of bio-signals and utilization of biological enzymes. The enzyme utilization capacity can be 'learned' by cells. The previous paper has discussed the intelligence of cells.

2. Key features of metabolomics in this paper:

1. Succession of study focus on the metabolic biochemistry pathways from primary

metabolites to final metabolites, but choose enzymes as the indicator of metabolic reaction pathways;
2.Compared with previous enzyme spectroscopy, this article designed the novel matrix calculation methods to systematically analyze the cell function regulation pathway, because the environment adaptability trait of cell (such as a specific pathogen resistance) is functioned by multiple metabolic pathways to regulate a number of sub-functions, and the multiple sub-functions and pathways are not independent of each other;
3. It connects genotypes with phenotypes to distinguish which functions of cells are inherited and which need to be cultivated.

译文：细胞结构生物中 DNA 分子应当准确定义为物质遗传信息的主要载体。因为细胞是有智力的，遗传过程中除了物质遗传信息，还有精神记忆机制。这在我之前的论文中已经论述。那么在此后，细胞中开始了复杂的新陈代谢过程，这里可以区分为初级代谢生化分子（如 tRNA），次级代谢生化分子（如本人在植物逆境生理中论述的 glutamic acid），和最终代谢生化分子（如多糖化合物）。其中不可缺少的就是酶。酶是细胞生物中对新陈代谢生化过程进行调节的首要生化指标。初级和次级代谢生化分子既可以受遗传 DNA 影响、也可以受细胞环境的影响而显著发生变异特性，但是细胞生理功能的调节，仅仅通过调节同工酶酶谱的合成和分泌作为生化反应链中首要途径（在同一同工酶家族中有许多不同的酶种类），不管 DNA 先天遗传和细胞环境变化因素。比如，在本人之前文章已经论述，针对特定病原体的抗体，不管宿主细胞 DNA 遗传变异特性，也不管细胞环境的变化，细胞内合成针对特定病原体的有效抗体的同工酶谱都是相对唯一性和特定性。正如本文已经论述多细胞高等智慧生物的思维局限于三维思维能力，而单细胞生物的智力思维能力应该仅仅停留在识别生物信号和运用生物酶两个维度而已。而细胞生物对生物酶的识别与合成可以是后天培养的。本文已经在细胞“智慧”一文中论述了细胞的思维。

本文中新陈代谢组学的特点：
1.继承从初级代谢物到最终代谢物中对新陈代谢生化反应途径/路径作为重点研究方向，但是选择以酶作为新陈代谢反应途径的指示性指标；
2.与之前酶谱学相比，本文设计了一个新的矩阵算式方法系统地对细胞生物功能调节途径进行分析，因为细胞环境适应性中的某一性状（比如对某一病原体实现抗病性），是由于细胞在多项功能、多个新陈代谢途径进行调节的结果，而且细胞各功能、各个新陈代谢途径之间一定是相互关联，非独立的。
3.连接了基因型和表现型，区分细胞哪些功能是先天遗传获取的，哪些功能需要后天培养的。

3. Genotype and Genetic Diversity Conservation

The feasibility of large-scale application of DNA markers on biodiversity assessment has been discussed by Liu et al.,(2014)[1]. However, the DNA markers suit not only for the classification of plant sub-populations for biodiversity assessment, but also provide the faster and convenient tool to identify the suitable plant varieties (genotype)

from wild ecosystem for ecological restoration. The suitable environmental conditions for each variety growth (phenotype) can be identified by the analysis of both community and species interactions with environment as discussed by Liu et al.,(2015)[2]. According to the Environmental Standard on Classifying the Categories of Genetic Resources (HJ 626-2011) in Mainland China, three kinds of DNA molecular methods have been listed to rank genetic resources (or endangered species) between categoryIand categoryII, including assessment of genetic diversity, evolutionarily significant unit (ESU), or genetic contribution rate, which have been substantially discussed by Liu et al.,(2015)[2]. However, it is advised that assessment of genetic diversity would be the first choice in ranking genetic resources (or endangered species), when the total SSR primers are calculated [3]; assessment of genetic variation would be the best method to select the suitable varieties for restoration of endangered species (or other important constructive species as well), when only polymorphic SSR primers are calculated [3]. The optimalization of both sampling units and polymorphic SSR primers, which allows to well present the genetic diversity for each variety at reasonable cost, has been pointed out as well [3].

4. Metabolomics and Environmental Adaptivity

However, the supplementary test of biochemical variation in enzyme species among different varieties collected in field, as the indicator for different varieties to adjust metabolism pathways in different environmental conditions, is advised for the conclusion of environmental adaptability between genotype and phenotype (metabolomics analysis). To be more comparable, the biochemical variation in enzyme species within one isozyme family, which catalyze the same metabolism substances, is analyzed according to the similarity coefficient. The function of each enzyme family in plant resistance to different environmental stress is summarized in table 1 below, which can be used for the development of isozyme primers initiating the isozyme test. The experiment procedure of biochemical test is listed in isozyme chapter [4]. To minimize the inaccurate conclusion between genotype and phenotype, the comparison between different varieties should be conducted on the basis of bio-samples collected in the same tissue and development phase of a plant species during the same season. In principle, the higher variation in enzyme species among varieties, the better environmental adaptability for restoration. This can be attributed into two reasons: firstly, the activity of an enzyme species only responds to the specific environmental conditions, and consequently the higher enzyme species variation of an isozyme family would result in the broader environmental conditions triggering the activity of the whole isozyme family; secondly, the gene expression of an enzyme species would be regulated by the specific environmental conditions only, which also explains the higher environmental adaptivity caused by the higher enzyme species variation of an isozyme family due to the broader environmental conditions for the regulation of gene expression as the whole isozyme family. Both reasons can result in the variation in isozyme electrophoretogram. The enzyme function in plant resistance to environmental stress is summarized below, and the chemical functional

group of these enzyme molecules become the indicator synthesizing the isozyme primers for metabolomics test.

Table 1. The Enzyme Function in Plant Resistance to Environmental Stress.

Isozyme Families	**Function in Plant Resistance to Environmental Stress**
Peroxidase (POD)	Disease Infection[5]
Catalase (CAT)	Drought Stress; Temperature Stress (both cold and hot); Salinity Stress; Disease Infection; Ozone; Radiation Stress.[6]
Malate dehydrogenase (MDH)	Acid soil; Aluminum toxicity[7]
Alcohol dehydrogenase (ADH)	Waterlogging Stress; Salinity Stress; Cold Stress; Drought Stress; Anaerobic Stress.[8]
NAD-dependent isocitrate dehydrogenase （ICDH）	Drought Stress; Salinity Stress; Heavy Metal Stress; Anti-Oxidation;[9]
Lactate dehyderogenase (LDH)	Heavy Metal Stress;[10]
Glucose-6-phosphate dehydrogenase (G6PDH)	Anti-Oxidation; Drought Stress; Salinity Stress; Cold Stress;[11]
Glutamate dehydrogenase (GDH)	Drought Stress; Salinity Stress; Cold Stress; Disease Infection[12]
Malic Enzyme (ME)	Drought Stress; Salinity Stress; Cold Stress; UV-B Radiation Stress; Physical Injury;[13]
Beta-Amylase(BAM）	Drought Stress; Temperature Stress (both cold and hot); Salinity Stress;[14]

The calculation of similarity coefficient between zymogram of different varieties is performed in one isozyme family[4]. However, the overall similarity coefficient among different isozyme families is calculated, on the basis of matrix for PCA analysis designed below, to reveal the systematics of environmental adaptability, as metabolomics analysis. The comparison of enzyme species variation between different seasons is required to reveal some resistance characteristics during specific environmental stress (such as cold stress). Compared with other article of this journal, the simulated environmental conditions of microbial cultivation are not suitable for botany. There are two reasons: firstly, the metabolic enzymes of botanical species is usually less sensitive to environmental conditions in comparison to microbes; secondly, the life cycle of constructive species for ecological restoration of botany communities can be hardly simulated in the controlled Lab. A novel matrix is designed below to conduct PCA analysis on the basis of comparisons between different isozyme families:

In the electrophoretogram of zymogram, if the electrophoresis pipes are the vertical ones, different electrophoresis pipe contains different isozyme family for comparison. Each isozyme family is labeled as 1, 2, 3..., and E in each electrophoresis pipe; It is hypothesized that the electrophoresis bands at the same horizontal line across different isozyme families are the enzyme species at the same locus (due to the same status of relative molecular weight), named as enzyme 'species i' (i = 1, 2, ..., I), and each isozyme family has the same amount (I) of enzyme species. The underlying theory making the bands comparable across different isozyme families is presented below: a locus in DNA / gene molecule should be defined as the proportion of a specific gene sequence segment to the total information of the whole genome, rather than as a specific physical position in the whole genome, which is more accurate. Gene mutation at a certain locus is defined as the gene sequence alteration at a specific physical position in the genome, which is not accurate. PCR electrophoresis is a reflection of the relative molecule mass among various gene sequences. Therefore, the relative position of a specific PCR electrophoresis band is not only a reflection of its quantitative gene among all the gene sequences tested in this experiment, but also a reflection of the proportion of the specific gene sequence information to the total information of all gene sequences tested in this experiment. Correspondingly, the electrophoretic bands in the electrophoretogram also reflect the relative molecular weight of the enzyme molecules expressed by different quantitative genes tested in this experiment. The electrophoretic bands at the same locus are also comparable among different isozyme species of an individual. The electrophoretic bands on the same locus among different isozyme species are just the reflection of different gene sequences with the same quantitative gene status (or the same proportional gene information to the total gene information tested in a experiment).译文：对应的矩阵算式所蕴含的理论观点如下：DNA/基因分子结构中的某个位点（locus）应当定义为特定基因序列片段所蕴含的基因信息在整个基因组序列中信息总量的比例，不应当定义为在整个基因组上的特定物理位置，这样定义更为准确一些。基因在某个位点上的突变，定义为发生在基因组的某个特定物理位置上，这样不是很准确，

仅仅是便于理解而已。PCR 电泳条段是各种基因序列片段之间一种相对质量的反映。因此特定 PCR 电泳条段的相对位置是其自身在该实验中所有测定基因组序列中数量型基因的反映，同时也反映该特定基因序列在该实验中所有测定基因序列信息总量中的比例。对应的，酶化学电泳图谱中的电泳条段也反应各种不同数量型基因所表达出酶分子的相对分子质量。对于一个个体的多种类同工酶酶谱，在同一位点上的电泳条段，也具有可比性。对于不同同工酶种类之间的在同一位点上的电泳条段，正好就是该个体在数量型基因具备相等地位（基因信息比例具备相同地位）的不同基因序列之间的反映。

Then there is a 3-dimension (I× E × N) matrix presented in this research. I is the total amount of enzyme species within a isozyme family; E is the total amount of isozyme families; N is the total amount of zymograms among different simulated moisture conditions:

X= ⎪Xien⎪(i = 1, 2, ...I; e = 1, 2, ... E; n= 1, 2, ... N)
Xien is the occurrence of enzyme ‘species i’ in the isozyme ‘family e’ during simulated moisture condition Tn. The value of Xien is one or zero. If the electrophoresis band occurs at this locus, the value is one;otherwise it is zero. The matrix X is below:

$$\begin{vmatrix} X111 & X112 & ..... & X11n & X121 & X122 & ..... & X12n & ..... & X1e1 & X1e2 & ..... & X1en \\ X211 & X212 & ..... & X21n & X221 & X222 & ..... & X22n & ..... & X2e1 & X2e2 & ..... & X2en \\ ..... & ..... & ..... & ..... & ..... & ..... & ..... & ..... & ..... & ..... & ..... & ..... & ..... \\ Xi11 & Xi12 & ..... & Xi1n & Xi21 & Xi22 & ..... & Xi2n & ..... & Xie1 & Xie2 & ..... & Xien \end{vmatrix}$$

Matrix Se = Xe × (Xe)T , where Xe = ⎪Xin⎪(i = 1, 2, ...I; n= 1, 2, ... N); (Xe)T is the transpose of the matrix Xe. The matrix Xe is below:

$$\begin{vmatrix} X11 & X12 & ..... & X1n \\ X21 & X22 & ..... & X2n \\ ..... & ..... & ..... & ..... \\ Xi1 & Xi2 & ..... & Xin \end{vmatrix}$$

The Principal Components Analysis (PCA) method of matrix X is specified [1]. However, the overall matrix X can be divided by sub-factors: PCA is firstly conducted on the basis of matrix Se, revealing the biochemical dynamics of a isozyme ‘family e’ among different simulated moisture conditions. In matrix Se, it is hypothesized that the variable in PCA represents the biochemistry dynamics of each enzyme ‘species i’.

Matrix S = $\sum_{E}^{1} Se$ (e = 1, 2, ..., E)

PCA is further conducted on the basis of matrix S, revealing the biochemical dynamics among different isozyme families over the whole simulated moisture conditions. In matrix S, it is hypothesized that the variable in PCA represents the biochemistry dynamics of each enzyme 'species i' across all the isozyme families. Further application has been discussed in later articles of this journal.

5.Phenotype and Gene Mapping for Genetic Breeding

Environmental adaptivity is definitely one of the main considerations for plant genetic breeding in restoration work. Nevertheless, as discussed in other article of this journal, gene expression traits as higher environmental adaptivity are usually associated with the gene traits of lower biomass productivity (or carbon sink), which means that both gene traits would be located in the same linkage group of genome. However, the gene trait of plant drought tolerance would increase the capacity of water & soil conservation due to the advantageous partitioning for root system, which results in higher ratio of root biomass to the total biomass. For the conservation of endangered birds, the gene traits as the partitioning of more branches for habitats or suitable fruits would become the major consideration in variety selection as well.

According to the results sourcing from the 'Crop Science' course instructed by Lincoln University NZ in 2007, yield components were also significantly affected by genotypes. The highest values of pods/plant, seeds/pod, and mean seed weight were achieved from genotype Aragorn, genotype PRO, and genotype Midichi, respectively. However, the total seed yield was not affected by pea genotypes. This result indicated the interdependent compensation mechanism among yield components. Wilson (1987) and Taweekul (1999) also suggested that large variation in one yield component might not lead to changes in total seed yield, due to the 'plasticity' of yield components [16-21]. However, my article here further points out that the experiments above are conducted under the 'comfortable environment' with sufficient growth conditions, which does not reveal the environmental adaptiveness under environmental stress in the field. This is firstly explained as the inter-dependent compensation mechanism among these yield components. However, my article would also explain this inter-dependent compensation mechanism by the theory that the sets of gene, underlying the expression as these yield component traits above, should locate in the same linkage group of genome. This plastic inter-dependent compensation mechanism leads some agriculture scientists to announce that the gene traits of yield components are not useful in breeding selection. However, this article hypothesizes that the gene expressed as partitioning more branches would locate in the same linkage group as some gene traits of environmental adaptivity (such as drought tolerance and higher

capacity of nitrogen fixation in root system), which becomes the objectives of my future study. The infection between microbes of biological nitrogen fixation and botanical roots must be quite specific[15], so the thinner root skin, usually associated with the partitioning of more root branches, would benefit the parasitic infection of microbes, enhancing the biological nitrogen fixation in root system. Additionally, the gene trait of partitioning more branches should result in higher radiation use efficience (RUE) as well, an environmental adaptivity trait in shading side of hills. This gene trait provides not only more suitable shelters for endangered birds, but also higher sustainability of habitats for food.

This is the revised materials in book “Proceedings for Degree of Postgraduate Diploma in Environmental Science (3rd Edition).” Published in 2016. The ‘chapter’ content mentioned in this article is in previous book. Secondly revised on 05/01/2021. Thirdly revised on 17/09/2021. Fourthly revised on 29/11/2021. Fifth revised on 20/12/2021. This journal article is previously published as: Liu Huan. (2021). Article 7. Metabolomics (1) --- The Systematic Chemistry Fingerprints Between Genotype and Phenotype and its Application on the Conservation Genetics. Journal of Environment and Health Science (ISSN 2314-1628), 2021(02)., which is converted into Journal of Biological Sciences (ISSN 2958-4035). Both Journals belong to the same publisher, Liu Huan. The previous journal article is closed to the public, but the previous reference is still valid. Latest revised on 29/05/2023.

Acknowledge:
Advice on the section 5 (yield components and nitrogen fixation content) of this article: Hill, G. D. & Mckenzie, B. A. Lincoln University, New Zealand. 2007.

References:

[1]. Review of Conservation Genetics and Its Application on Biodiversity Monitoring and Assessment. 刘焕，张洪初与唐秋盛，保护遗传学方法在生物多样性监测和评价领域的应用研究．科技视界，2014(8). DOI: 10.3969/j.issn.2095-2457.2014.08.203
[2]. Liu, H, Ouyang T. L, Tian Chengqing (2015). Review of Evolutionary Ecology Study and Its Application on Biodiversity Monitoring and Assessment. Science & Technology Vision (6) 2015.
[3]. Liu Huan. 2022 (12).Epidemiology------Comparison between indoor and outdoor air quality at three representative sites in Auckland Center. Journal of Environment and Health Science (ISSN 2314-1628). https://doi.org/10.58473/JEHS0010
[4]. 周延清，张改娜与杨清香，生物遗传标记与应用, 2008，化学工业出版社.
[5].梁艳荣，胡晓红，张颖力，刘湘萍(2003).植物过氧化物酶生理功能研究进展.内蒙古农业大学学报．第 24 卷第 2 期.
[6].陈金峰，王宫南，程素满(2008).过氧化氢酶在植物胁迫响应中的功能研究进展.西北植物学报, 2008, 28(1):0188-0193.
[7].王晓云,毕玉芬(2006). 植物苹果酸脱氢酶研究进展.生物技术通报.2006 年第 4 期.
[8].张计育，王刚，黄胜男，宣继萍，贾晓东，郭忠仁 (2015).乙醇脱氢酶基因家族在植物抵抗非生物胁迫过程中的作用研究进展.中国农学通报 2015,31(10):246-250.
[9].郝兆丰 袁进成 刘颖慧.(2012).异柠檬酸脱氢酶在植物抗氧化胁迫中的作用.生物技术通报.2012 年第 6 期.
[10]. 段昌群 王焕校.(1998).Pb+ 2、Cd+ 2、Hg+ 2 对蚕豆(Vicia fabaL.)乳酸脱氢酶的影响. 生态学报.第 18 卷第 4 期.
[11].于定群，汤浩茹，张勇，罗娅，刘泽静(2012).高等植物葡萄糖-6-磷酸脱氢酶的研究进 展.July 25, 2012, 28(7): 800−812.
[12].黄国存,田波.(2001). 高等植物中的谷氨酸脱氢酶及其生理作用.植物学通报 2001,18(4):396～ 401.
[13].周滈，杨传平，柳参奎(2011).植物苹果酸酶在抵御逆境中的作用.第 36 卷第 6 期.2011 年 11 月.
[14].杨泽峰，徐暑晖，王一凡，张恩盈，徐辰武.(2014).禾本科植物β- 淀粉酶基因家族分子进化及响应非生物胁迫的表达模式分析.科技导报 2014，32（31）.
[15].Liu Huan. Biodiversity Conservation & Management Studies in New Zealand. September, 2021.Journal of Environmental & Health Science.https://doi.org/10.58473/JEHS0008
[16]. Ayaz, S., Mckenzie, B. A., Hill, G. D., & McNeil, D. L., (2004). Variability in yield of four grain legume species in a sub-humid temperate environment I. Yields

and harvest index. Journal of Agricultural Science (2004), 142, 9-19.https://doi.org/10.1017/S0021859604004101
[17]. Ayaz, S., Mckenzie, B. A., Hill, G. D., & McNeil, D. L., (2004). Variability in yield of four grain legume species in a sub-humid temperate environment II. Yield components. Journal of Agricultural Science (2004), 142, 9-19.https://doi.org/10.1017/S0021859604004101
[18]. Moot, D. J., & McNeil, D. L., (1995). Yield components, harvest index and plant type in relation to yield differences in field pea genotypes. Euphytica 86, 31-40.https://doi.org/10.1007/BF00035936
[19]. Moot, D. J. (1997). Theoretical analysis of yield of field pea crops using frequency distributions for individual plant performance. Annals of Botany 79: 429-437. 1997.https://doi.org/10.1006/anbo.1996.0373
[20]. TAWEEKUL, N. (1999). Factors affecting seed vigour in field peas (Pisum sativum). Ph.D. thesis, Lincoln University,Canterbury.
[21]. WILSON, D. R. (1987). New approaches to understanding the growth and yield of pea crops.Agronomy Society of New Zealand 6, 23–28.