Review

Chinese Science Bulletin

SPECIAL TOPIC
Omics in Marine Biotechnology

September 2012 Vol.57 No.25: 3253–3261
doi: 10.1007/s11434-012-5237-5

Metabolomics: Concept, methods and potential prospect in marine biology

ZHANG Xuan & CHEN Hao*

Key Laboratory of Marine Biological Active Substances, The First Institute of Oceanography, State Oceanic Administration, Qingdao 266061, China

Received October 19, 2011; accepted April 23, 2012; published online July 11, 2012

The term "omics" refers to the comprehensive analysis of a specific biological system. With the development of omics, a number of omics subdisciplines have emerged, which play an important role in system biology. As a subdiscipline of omics, metabolomics provides a comprehensive analysis of the metabolome and has been widely applied to various fields of biology. In this paper, we introduce the concept, approaches, applications, and promising prospects of metabolomics.

metabonomics, metabolomics, biosynthetic pathway, NMR, LC/GC-MS

Citation: Zhang X, Chen H. Metabolomics: Concept, methods and potential prospect in marine biology. Chin Sci Bull, 2012, 57: 3253–3261, doi: 10.1007/s11434-012-5237-5

The complete sequencing of the human genome has introduced a new field of systems biology, known as omics. In biology, the term "omics", which means the totality of sorts, refers to the comprehensive analysis of a specific biological system. For example, "genomics" is the study of all the genes of a cell or tissue at the DNA (genotype), mRNA (transcriptome) or protein (proteome) level [1]. Omics is an integrated concept that combines all the information acquired from the different levels at which biological systems are controlled to investigate the nature of biological phenomena at the molecular level. Compared with traditional biological ideology, omics places more emphasis on the integration of diverse bioinformation. Thus, it provides a general view of processes rather than detailed information on the independent behaviors of specific cells or tissues. To date, a range of omics subdisciplines have emerged, including genomics and proteomics, each of which has its own set of instruments, techniques, reagents, and software. Omics technology has driven the development of new research approaches, including DNA and protein chips, mass spectrometry, and instruments that enable high-throughput analysis. Omics not only offers the advantage of understanding biological processes, but the prospect of more accurate diagnoses, improvements in nutrition and other new applications in biological fields.

Nicholson first propounded the concept of metabonomics in 1999. The primary definition of metabonomics is "the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification" [2]. Metabonomics focuses on understanding the global biological system by the study of metabolic compounds. The early study of metabonomics was focused primarily on the application of nuclear magnetic resonance spectroscopy (NMR) to analyze the changes of metabolic compounds in biological fluids, cells, and tissues, and was "aimed at the augmentation and complementation of the information provided by measuring the genetic and proteomic responses to xenobiotic exposure" [2].

Another important subject in the study of metabolic compounds is metabolomics. The focus of metabolomics is the compositional analysis of low molecular-weight substance in biological samples. This approach traditionally uses MS (Mass Spectrometry) in combination with a separation technique such as gas chromatography (GC) or liquid chromatography (LC) to analyze information from low molecular-weight compounds. To avoid interaction between

*Corresponding author (email: hchen@fio.org.cn)

the compounds being measured in the matrix, sample preparation and analysis is focused on a series of compounds that share similar properties. The comprehensive identification and quantification of the metabolic compounds synthesized by an organism means that metabolomics is one of the most important approaches to study genetic expression and proteomics.

Although there is conceptual similarity between the fields of metabonomics and metabolomics, there are some significant differences in the approaches and objectives of the research. The former is dedicated to comprehensive metabonomics profiling at a given scale, whereas the latter is primarily concerned with the history of time-dependent metabolic profiles in an integrated system [3]. However, associated with the development of bioinformatics, the two concepts will likely be integrated. For convenience, we use the term metabolomics in the following discussion.

An organism is an integrated and networked system with a bioinformatics transport chain consisting of genes-proteins-metabolic compounds. Genes and proteins provide information about what will happen, whereas metabolic compounds provide information about what has happened. In comparison with other omics subdisciplines, metabolomics is defined by the following unique characteristics: (1) Metabolic data do not provide the information produced by the independent behaviors of specific cells or tissues but, instead, documents the global changes of metabolome under the external interference, which reflect the metabolic response to exogenous stimuli in a general view. (2) Metabolic compounds are the endpoint of functional gene activity. According to Johnson's research, metabolic throughput is not only controlled by genetic expression, but also adjusted by the exogenous environment [4]. The changes in concentrations of metabolic compounds illustrate the biological response to perturbation from outside. Therefore, metabolic compounds are more sensitive compared with genes and proteins. (3) During a biochemical reaction, the concentration of enzymes and metabolic throughput is generally unchanged, whereas the concentration of some low-molecular weight compounds will change significantly. Being the endpoint of the bioinformatic transport chain, the signal generated by alteration of the metabolome is magnified and can provide insight and accurate characterization of the cells' function [5].

1 Progress in metabolomics

1.1 Pharmacotoxicology

Pharmacotoxicology involves studying the toxicity of drugs, their effect on target organs, their latency, and the relationship between structure and toxicity. After being taken into an organism, a heterologous substance acts directly on the metabolome, but has no apparent effect on the genome or proteome. Thus, there is little value in using genomics and proteomics to evaluate the mechanisms of drug toxicity. Traditional drug screening depends primarily on biological screening, which is reliant on large sample sizes and is time consuming. The advent of high throughout metabolomics technologies means it is now possible to analyze large numbers of samples in a relatively short period of time. Using metabolomics technologies, we are able to detect abnormal changes in trace materials in biofluids such as urine, serum, and tissue biopsies following exposure to a heterologous substance. In addition, these technologies allow for simultaneous quantification of total metabolic compounds. For example, Zhou et al. [6] injected dexamethasone into a group of mice at day 17.5 of pregnancy then collected ^{1}H-NMR spectra data for plasma metabolites. Using this data, Zhou et al. [6] calculated the rate of uptake in embryos with/without cleft palate. Based on principle component analysis (PCA) and Soft Independent Modeling of Class Analogy (SIMCA), the result showed that the rate of uptake differed significantly in embryos with cleft palate. Zhou et al. [6] demonstrates the power of the metabolomics approach in detecting abnormalities in maternal metabolites during gestation, and supports the author's hypothesis that the presence of dexamethasone in maternal metabolites contributes to the occurrence of cleft palate.

Metabolomics uses biochemical phenotypes to analyze the integrated functioning of an organism, while traditional Chinese medicine combines the alteration of the outward manifestation and abstract immanence to study the function of a specific system. Thus, both approaches share a similar philosophy [7]. Metabolomics can be used to evaluate the pharmacotoxicology of traditional Chinese medicine. For example, Ma et al. [8] fed mice an extract of Morning Glory seed (MGS). Urinary samples from control and MGS treated rats were analyzed by ultra-performance liquid chromatography/mass spectrometry (UPLC-MS). Blood biochemistry and histopathology were examined for signs of renal damage. Treatment with MGS caused a significant change in eight metabolic compounds, regarded as biomarkers, demonstrating that MGS causes nephrotoxicity. This example illustrates the potential of UPLC-MS based metabolomics for mapping the metabolic response to toxicant exposure.

1.2 Clinical diagnosis

Using metabolomics technologies, we can acquire metabolic data from patient and control populations to isolate biomarkers for disease. This information aids in the understanding of the pathogenic mechanisms of a specific disease and provides important clues for clinical diagnosis and treatment. Traditional clinical diagnosis depends on the manifestation of early symptoms, which almost always delays the prognosis and treatment. The biological liquid circulating in an organism is in a state of homeostasis. Therefore, any alteration of specific metabolic compounds produced by

the disturbance of homeostasis provides evidence to diagnose illnesses. To date, metabolomics has been applied to the clinical diagnosis of diabetes, tumors, cardiovascular and cerebrovascular diseases, and the assessment of organ graft quality and function before transplantation. Marchesi et al. [9] combined high resolution ^{1}H-NMR spectroscopy and multivariate pattern recognition to analyze the fecal extracts of both Crohn's disease (CD) and Ulcerative colitis (UC) patients. The fecal extracts of both CD and UC patients were characterized by reduced levels of butyrate, acetate, methylamine, and trimethylamine in comparison with a control population, and elevated quantities of amino acids. This example demonstrates a number of advantages of metabolomics, namely that it allows for noninvasive diagnoses and rapid and accurate analysis of low quality samples.

A number of factors, including stress, changes in the environment, and radiation have no apparent effect on people's health. These factors can cause dysfunction and subclinical injury in organs and are hard to diagnose. However, abnormal changes in biomarkers in a number of metabolic pathways can be measured to determine the effect on an individual. For example, Chen et al. [10] used fluorescence difference gel electrophoresis combined with matrix assisted laser desorption ionization/time of flight MS (MALDI-TOF-MS) and ^{1}H-NMR to monitor the intracellular processes in the rat liver at the proteomic and metabonomic levels following restraint stress for 8 weeks. The authors observed elevated levels of lactate, fatty acid, glucose, and homocysteine in the livers of treated rats. Thus, the data confirm that restraint stress can cause subclinical hepatic injury, including inhibition of glycolysis and gluconeogenesis and dysfunction of fatty acid β-oxidation.

1.3 Hyphenate with genomics

With the development of genomics, the emphasis of genomic research has shifted gradually from the sequencing of genomes to the expression of the functional genome. The functions of the hereditable information in the genome, the expression and regulation of the proteome, and metabolic perturbations reflect the hierarchy of the bioinformation transport chain. As the endpoint of genomic expression, metabolomics reflects the intact state of functional genetic activities. This information aids biomolecular scientists in the comprehensive study of biological phenomena and the essence of life. Currently, such research is focused primarily on plant metabolomics. The objective is to measure the metabolic response caused by genovariation and alterations in the environment to determine the function of silent genes and the relationship between genotype and phenotype, and collect more information about the metabolic pathway of plant. Taylor et al. [11] and Fiehn [12] used GC-MS to detect 443 metabolic compounds in different genotypes of *Arabidopsis* and PCA and Artificial Neural Network (ANN) to analyze and classify the data. The authors found 4 metabolic compounds (malic acid, citric acid, glucose, and fructose) that could be used as criteria for classification. Using antisense mediated down-regulation and overexpression of the *Rpd3* gene (ZmRpd3) in the genome of an inbred maize line, Castro et al. [13] identified differences in metabolites between sample classes that could be directly integrated into the evolution of metabolic perturbation over time in the model. Moreover, the authors noted that major differences existed at the beginning of development, validating the hypothesis that ZmRpd3 transcription and proteins accumulation occurs during the initial stage of development, suggesting a role for this gene in cell cycle control.

Metabolomics has also been used to analyze the metabolic response after knocking out specific genes with the purpose of exploiting the relationship between the genotype and phenotype in transgenic animals. For example, Rull et al. used ^{1}H-NMR spectroscopy to assess the metabolic function of the *MCP-1* gene in transgenetic mice [14]. The data suggest that MCP-1 plays an important role, not only in macrophage migration, but also in lipid and glucose metabolism.

1.4 Nutriology

The current focus of nutrition science is on improving the health of individuals through diet. Thus, the goal of most research is to promote health, prevent disease, and improve performance [15]. Conventional tools for measuring dietary exposure depend primarily on food composition tables to allow estimation of intake of energy, nutrients, and non-nutrient food constituents. The accuracy of this method is always affected by an unknown degree of misreporting. Metabolomics allows the simultaneous monitoring of multiple and dynamic components of biological fluids, so can be used to determine which metabolic signals are influenced by food intake [16]. The complexity of the metabolic response to the intake of different foods makes it challenging to study the relationship between the metabolome and patterns of food intake. However, metabolomics allows a researcher to monitor the metabolic response under a range of nutritional conditions and food intake. Furthermore, this approach promotes chemometrics as a range of factors, including food type, geography, and race can be simplified by choosing a suitable algorithm. Hence, we can cluster the distinctive personal food patterns to simplify the data process [17]. For example, Peré-Trepat et al. [18] analyzed ^{1}H-NMR plasma metabolic profile data with PCA and partial least squares discriminant analysis (PLS-DA) to analyze the metabolic profile of a population under 5 different food patterns. The authors demonstrated that the metabolic phenotypes associated with these dietary patterns were based primarily on differences in lipid and amino acid profiles in the plasma. This new approach to assess the relationship between dietary intake and metabolic profile data will allow greater steps to be made in merging nutritional epidemiology with metabolomics. Pedersen et al. [19] used ^{1}H-NMR

plasma metabolic data to compare the metabolite profile of patients suffering from Irritable Bowel Syndrome who ate a non-probiotic acidified milk product to ones that ate a probiotic acidified milk product. Both diets resulted in elevated levels of blood serum *L*-lactate and 3-hydroxybutyrate, compounds that can be used as biomarkers after taking in high levels of lactate.

1.5 Microbial metabolome

The application of metabolomics to microbiology focuses primarily on microbiological assays, the metabolic pathways used by microorganisms, and the analysis of microbial biochemical reactions. Compared with the traditional microbiological assay technologies, metabolomics does not consider the phenotype and specific biochemical reactions, but instead quantifies the changes in metabolic compounds in microorganisms. This approach offers several advantages, including high speed, high throughput, and so on. Furthermore, the technology allows for excellent discrimination of strains with similar genetic information and distinctive pathogenicity. Bundy et al. [20] combined multivariate pattern-recognition methods with ^{1}H-NMR to analyze the metabolite profiles derived from six different *Bacillus cereus* strains. The two different ecotypes were clearly separated on the basis of their metabolite profiles, illustrating the utility of using metabolomics methods to classify pathogens on the basis of their expressed physiology, even when candidate genes could not be isolated to unambiguously distinguish between the two groups.

2 Use of metabolomics in marine biology

2.1 The study of viable but nonculturable marine microorganisms

Colwell et al. [21,22] discovered a survival strategy, known as the viable but nonculturable (VBNC) state, that is adopted by bacteria when exposed to environmental stresses to prevent rapid cell death. In the VBCN state, microorganisms are unable to form colonies, but the cells remain toxic [23,24] and transcription, which yields specific mRNA products [25,26], is coupled with the modification of the cell wall [27,28]. Furthermore, a proportion of the VBCN population exhibited anabiosis given the proper environmental conditions [29]. The VBCN state can be described as the self-protective response to adverse environmental conditions such as high salinity, high temperature, and oligotrophy. Although the widespread occurrence of VBCN in nature, particularly in the marine environment, provides abundant opportunities to study this special state, we still have not ascertained the mechanism(s) mediating this phenomenon. Metabolomics can be used to document the metabolic response to environmental perturbation. This approach may be used to conduct metabolic profiling of normal and VBCN strains and identify biomarkers for VBCN. In turn, the data may be used to investigate the role of biosynthetic pathways in controlling the VBCN state. Moreover, we can use metabolomics to detect the presence of the compounds that induce the VBCN state. Thus, metabolomics is likely to play a prominent role in this field in the future.

2.2 Marine active compounds (MACs)

Marine organisms have special structures and functions to adapt to aspects of their environment, including high salinity, high temperature, high pressure, and oligotrophic conditions. Some MACs produced during growth and metabolism exhibit a special function for adapting to these extreme conditions, and show promise for use as pharmaceuticals. A large number of MACs have been discovered that exhibit antineoplastic activity [30], antiviral activity [31], antioxidant activity [32] and anti-inflammatory activity [33], and some have been used in clinical research [34]. However, the low volume of active compounds that can be extracted from marine organisms limits their application to the production of new drugs. Furthermore, the discovery of novel active compounds generated by marine organisms is time-consuming. The concentration of specific compounds and the relationship with other compounds that can be determined using high throughput metabolomics technologies can help optimize the culture conditions to enhance productivity. Meanwhile, metabolic profiling of organisms under different conditions will provide information about the precursors of specific active compounds, which can be used to understand and control the processes generating the active compounds. Traditional active compound screening relies on labor and time intensive approaches, such as biological screening and high throughput molecular modeling and docking. By using high throughput metabolomics, we can observe the changes in metabolic compounds in an organism under distinctive environments and the metabolic response of experimental animals treated with active compounds. This data can then be used to aid in the choice of novel MACs.

Besides exploiting potential MACs, metabolomics also provides a new way to study the mechanism of action of MACs. Because of their potent cytotoxicity, structural diversity, and mechanistic complexity, MACs are rarely recognized as clinically relevant antineoplastic agents. The data obtained using metabolomics provides insight into the molecular mechanism of drugs and the enzyme or metabolic pathway that are targets for MACs. This information can be used to help modify MACs to produce new drugs. For example, Bayet-Robert et al. [35] used proton NMR spectroscopy-based metabolomics to unravel biochemical disorders induced in human MCF7 breast cancer cells by three lead candidate anticancer MACs. The response of MCF7 cells to these agents involves a number of metabolic pathways. Taken together, these examples illustrate the potential utility

of metabolomics to help investigate metabolic targets and develop new drugs.

2.3 Research on the biosynthesis pathway of potent marine natural products

Tetrodotoxin (TTX) 1 (Figure 1) is a potent neurotoxin that is responsible for fatal food poisoning caused by puffer fish. It is also an important probe in neurophysiological experiments as it is a highly specific sodium-channel blocker. Despite a long history of TTX research, little is known about the biosynthetic origin of this unique molecule. Metabolomics provides an ideal tool for us to determine the origin of TTX. Recently, we put forward a hypothetical biosynthetic pathway for tetrodotoxin (Figure 2). Glycine was supposed as an initial precursor, which was enzymatic catalyzed finally to form TTX. In our experiments, [^{14}C]glycine was fed to TTX-producing bacteria culture medium. Following exhaustive purification by paper chromatography and HPLC, we isolated the radioactive TTX. Furthermore, the involvement of other precursors in the TTX biosynthetic pathway may be revealed using this approach.

Marine invertebrates, particularly sponges and ascidians, are well known for their production of bioactive natural products, several of which are currently undergoing clinical trials [36,37]. Numerous synthetic approaches have been devoted to this class of compounds. However, because of the complexity of the structures, these approaches are not economically profitable. Bryostatin 1 (Figure 3), a potent anticancer agent, is a proposed metabolite of the *B. neritina* bacterial symbiont "*Candidatus* Endobugula sertula". Although this c-proteobacterium has not been cultivated, its reduction or elimination upon antibiotic treatment of laboratory cultured bryozoans leads to a reduction and eventual loss of bryostatin production in new generations of larvae without an effect on colony growth. Furthermore, bryostatins resemble bacterial type I PKS products, leading to speculation that these macrolides are derived from a microbial symbiont [38]. Based on the analysis of the secondary metabolites, a hypothetic biosynthetic pathway for bryostatin 1 was proposed.

2.4 Toxicology of marine heavy metal contamination

Heavy metal pollution has become a serious threat to the marine environment because of the persistence, ease of transport, and high toxicity of these metals. Marine bivalves can accumulate heavy metals in direct proportion to the environmental levels [39]. Thus, mussels and oysters are often used as sentinel organisms when monitoring for heavy metal pollution in the marine and coastal ecosystems. The majority of this research focuses on the response of sentinel organism following exposure to a specific heavy metal. Traditional methodologies involve measurement of specific metabolic compounds (e.g., acetylcholinesterase) to document neurotoxicity or antioxidant enzyme levels to document oxidative stress induced by heavy metals [40,41]. However, it is difficult to elucidate the effect on the whole metabolic pathway using this approach. Using metabolomics affords the ability to detect large numbers of endogenous low molecular weight metabolites in an organ or cells, so is

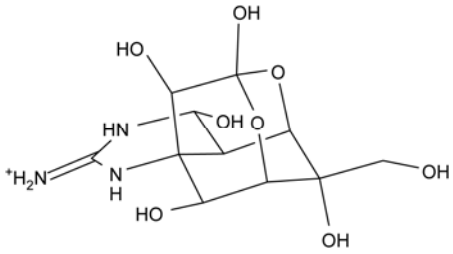

Figure 1 The structure of tetrodotoxin.

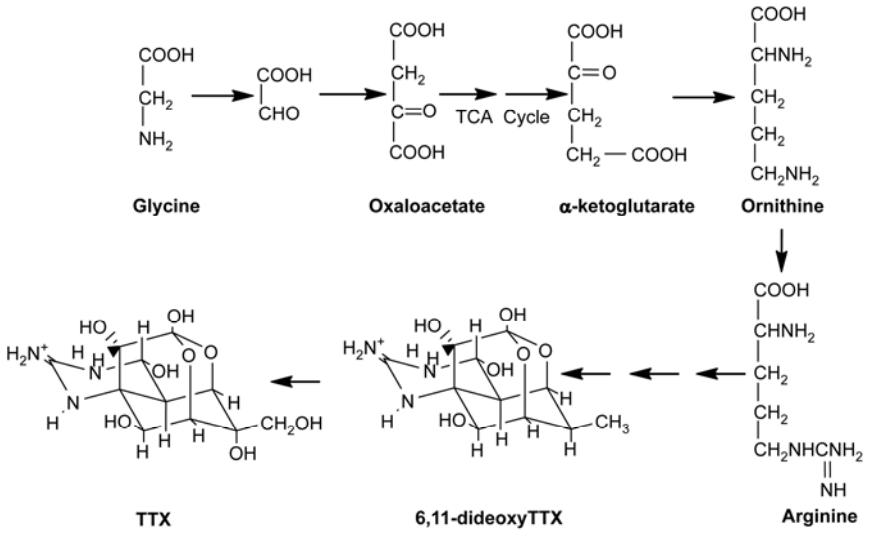

Figure 2 Hypothetical biosynthetic pathway of tetrodotoxin.

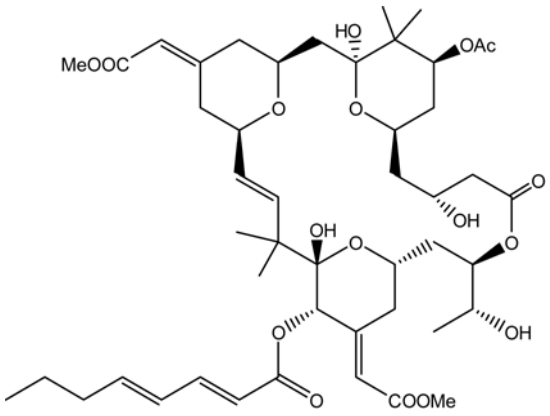

Figure 3 The structure of bryostatin 1.

a better option when investigating the mechanism(s) controlling the metabolic response to marine heavy metal contaminants. Zhang et al. [42] studied the metabolic responses of *Manila clam* to environmentally relevant copper exposures using NMR-based metabolomics. The authors showed that osmotic regulation and energy metabolism were disturbed by exposure to copper. This example illustrates the value of metabolomics in the discovery of metabolic biomarkers that can be used to understand the mechanistic pathways underlying marine heavy metal contaminant toxicology.

Other research has focused primarily on the tissue-special toxicological effects of heavy metal on indicator organisms. By investigating the distinctive metabolic profiles of different organs, we can establish time-dependent assessments for nutritional challenges and functional adaption of indicators to the environment. Andrey et al. [43] identified 32 major compounds and provided a basis for future metabolomics studies of the oyster in a comprehensive NMR analysis of the Eastern oyster. Increased concentrations of biomarkers were used as indicators of the main metabolic pathway in different organ blocks, providing reference information for tissue-specific toxicity. Similarly, Wu et al. [44] investigated the toxicological effects of Cadmium (Cd) in both the digestive gland and adductor muscle tissue of the green mussel (*Perna viridis*). The adductor muscle tissue was characterized by elevated levels of acidic amino acids and reduced levels of branched chain amino acids. These observations suggest that the adductor muscle tissue of mussels is a suitable supplemental biomarker for exposure to toxicants.

2.5 Cyanobacteria energy production

The drawback of traditional fossil fuels, including oil, coal, and natural gas, is their finite supply, which will be depleted in the relatively near future. The management and storage of nuclear wastes limit the worldwide use of nuclear power. Given this, scientists have begun to turn to new resources to exploit renewable and clean energy. Among the candidate resources, cyanobacteria have drawn increasing attention because of the ease with which they can be cultured and genetically modified. Cyanobacteria, being photosynthetic organisms, use solar energy, H_2O, and CO_2 to synthesize their energy storage components (i.e., carbohydrates, lipids, and proteins). The carbohydrates may be transformed to ethanol by fermentation under dark, anoxic conditions [45]. The carbohydrates, lipids, and proteins produced by cyanobacteria can be converted to CH_4 gas when they are cultured in conditions favoring anaerobic digestion [46]. Furthermore, many strains of cyanobacteria can produce hydrogen by the reversible activity of hydrogenase [47,48]. The focal point of current research is the identification of new strains with specific metabolic traits for optimizing cultivation conditions in bioreactors. Thus, the comprehensive information provided by metabolomics can also be used as supplementary data to improve renewable energy production. Using these techniques, researchers can identify specific enzymes or biomarkers to modify energy metabolic pathways or change the culture conditions according to the response of cyanobacteria to external stimuli. Although metabolomics has not been widely used in this field, there is considerable potential for its use in the study of cyanobacteria energy production.

2.6 The treatment and prevention of aquatic diseases

Poor management and breeding conditions are generally associated with an increase in the incidence of disease, which directly affects the growth and survival of cultured organisms. Cross-infection is common with aquatic diseases and infections are often difficult to treat. Thus, the key to sustainable production is in the prevention of disease. However, there remain difficulties in diagnosis due to the ambiguous nature of early symptoms. In this instance, metabolomics may be used to identify differences in the metabolic profiles of healthy and sick animals, in a relatively non-invasive procedure.

3 Metabolomics approaches

3.1 Analytical methodology for metabolomics

Metabolomics relies primarily on NMR and MS. The latter method is combined with LC or GC with precolumn derivatization as a supplementary means to separate the metabolic compounds. MS may also be combined with other separation methods such as capillary electrophoresis (CE), Fourier transform infrared spectroscopy (FT-IR), and array electrochemical detection.

(i) NMR. High throughput NMR yields information on hundreds of compounds with each measurement. The distribution of an individual signal is dependent on the chemical atmosphere around the nucleus. Thus, NMR provides a large volume of structural information to describe metabolic

compounds. However, the conditions of the solution (e.g., pH and ionic strength) have a significant influence on the observed spectrogram. NMR is the best choice for metabolic profiling because of the presence of different isotopes and the ubiquitous hydrogen bonds in biological molecules. A range of isotopes can be detected, including ^{1}H, ^{13}C, ^{31}P and ^{15}N, though ^{1}H-NMR is used most commonly. Some researchers have already used "inverse" methods of detection to detect compounds with low natural abundance or poor magnetic sensitivity. For example, the information generated by two or more dimensional data from different nuclei on the same molecule can be combined [49,50].

The decay of the NMR signal is highly sensitive to molecular dynamics. Lindon et al. [51] used the elimination of resonances to show the slowly tumbling proteins and lipoproteins in plasma or serum. Although NMR has lower sensitivity in practical applications when compared with LC-MS, the limitation of detection and the sample volume demanded for analysis is very low. This provides an advantage for analyzing mass limited circumstances or a large quantity of samples. Nowadays, automatic sample handling systems can be coupled with NMR to analyze hundreds of samples automatically with high accuracy and reproducibility per day [52].

Magic Angle Spinning NMR (MAS-NMR) involves the detection of intact tissues and cells with little or no preparation [53,54]. MAS-NMR significantly reduces the effect of magnetic field inhomogeneity in ordered materials. However, the peaks obtained from MAS-NMR are always broad and distorted. Thus this approach is not able to provide detailed information about the metabolic compounds in samples. However, the metabolic profile detected by MAS-NMR can be used to document the changes in metabolic compounds in intact tissue following exposure to toxins under different cellular conditions [55]. The ability to localize biochemical changes to a specific tissue, cell type, or organelle provides valuable information associated with the mechanism of toxicity and supplementary interpretation of biofluid analyses [3]. In conclusion, MAS-NMR provides an approach to validate and study the internal metabolic response to toxins *in vitro*.

(ii) GC-MS and LC-MS. In addition to having higher sensitivity than NMR, the distribution of signals in an MS spectrogram is dependent on the molecular mass. Thus, the MS spectrogram provides excellent discrimination and complete structural information. FT-MS (Fourier Transform MS) can be used for accurate mass determination to control the precision under one per million. To obtain additional structural information for a specific molecule, we can increase the energy of electron bombardment to generate more fragments of the parent molecule. The molecular structure can be rebuilt by comparison of the information from fragments and the standard spectrogram. However, the relationship between the concentration of metabolic compounds and the strength of peaks in the MS spectrogram is not linear, and because of ion suppression, the sensitivity is dependent on the presence of other compounds in the matrix. Some common separation approaches, such as GC and LC, are used to reduce the matrix effect. Recently, researchers used capillary electrophoresis as a supplementary approach to reduce the matrix effect [56]. This increases the chemical discrimination of different metabolic compounds and provides convenience when investigating metabolic profiles.

Most of the samples prepared for MS require derivatization because of the phase state of the mobile phase. This requires additional time and may cause transformation of the sample. The derivatization of MS limits the application to the analysis of heat sensitive substances and high-molecular-weight metabolic compounds. However, GC-MS has high analytical speed and separation efficiency. Comparatively speaking, LC-MC is suitable for the detection of unstable, non-volatile, and high-molecular-weight compounds without derivatization. However, it is also time-consuming and has low separation efficiency. The identification of compounds always requires further analysis because there is no standard database for searching and comparison.

(iii) Hyphenation. When combined with other separation technologies, MS has the advantages of rapid identification and detection of a range of compounds. However, the requirement that the output needs to be compared with standard database limits its application to detect unknown compounds. NMR can provide an atraumatic and reproducible approach to identify and quantify metabolic compounds, although it may be difficult to identify overlapping signals in complex compounds. As a result, the physical combination of multiple analytical technologies, or "hyphenation", has become a popular means of overcoming the disadvantages of each of these technologies. Since the Bruker interface combines NMR and MS, both methods have online analytical software that is able to offer the comprehensive structural data. High pressure Liquid Chromatography–diode array detector-solid phase extraction–NMR-MS (HPLC-DAD-SPE-NMR-MS) using Cryoflow Probes™ have also become more important for the analysis of complex samples [57]. In addition to being used to identify the structure of natural products, HPLC-DAD-SPE-NMR-MS can also be used to detect the metabolic response to drugs [58] and identify the compounds present in biofluid [59]. Besides the online system, integrated data can also be obtained by using different technologies in parallel. Chan et al. [60] used HR-MAS NMR and GC-MS to identify 31 marker metabolites in biopsied colorectal tumors and their matched normal mucosa obtained from 31 colorectal cancer (CRC) patients. The orthogonal partial least squares discriminant analysis (OPLS-DA) model generated from metabolic profiling can further differentiate colon from rectal cancers, providing new biomarkers for the diagnosis of CRC. Smilde et al. [61] combined GC-MS and LC-MS data to analyze the metabolic compounds of *E. coli* NST74 cultured under different nutritional conditions. This example illustrates how

this technology can be used to provide a comprehensive view of the metabolome of a microorganism [61].

3.2 Strategies for metabonomic data analysis

The key point of strategies for metabolomics data analysis lies in how to use "pattern recognition (PR)" to process the "raw" data [62]. Assuming the metabolic profile of two samples generated from a series of measurement is the same; the two samples are analogous. The central application of PR is the determination of the appropriate index of similarity. After determining the appropriate index of similarity, we can identify normal and abnormal metabolic profiles of samples under distinctive dysfunction and experimental control. Pattern recognition algorithms differ widely in both a general way in which variables are combined to generate a final output or model and in approaches to how a specific model is derived [3]. Unsupervised methods, such as PCA and self-organizing maps (SOMs), primarily rely on clustering and visible data to reveal the internal relationship and structure of the data. Conversely, supervised methods such as discriminant function analysis (DFA) and nonlinear regression take advantage of a given metabolic profile to establish a standard mode to predict the existence of specific metabolic data.

Beyond pattern recognition, there are other approaches to generate a functionally relevant model. For example, development modeling approaches reveal information about a living system with respect to the flux through metabolic pathways (metabolic flux/control analysis or MFA/MCA) [63]. These methods have been used successfully to analyze prokaryotic systems and eukaryotic microorganisms.

4 Conclusions

With the completion of the "Human Genome Project", researchers have obtained a large volume of information associated with a complex biological system. As the endpoint for genomics and proteomics, metabolomics combines the messages from higher levels of the hierarchy into an integrated system. Metabolomics offers noninvasive, rapid, accurate, and high throughput analysis, so is ideal for use in a range of biological fields. The processing of complex metabolic data improves the development of stoichiometry and bioinformation. Thus, metabolomics will play an increasingly important role in the future.

This work was supported by the National Natural Science Foundation of China (*20872022*).

1 US Environmental Protection Agency. Potential Implications of Genomics for Regulatory and Risk Assessment Applications at EPA. Science Policy Council. EPA Publication No. EPA 100/B-04/002, 2004
2 Nicholson J K, Lindon J C, Holmes E F. Metabonomics: Understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica, 1999, 29: 1181–1189
3 Keun H C. Metabonomic modeling of drug toxicity. Pharmacol Therapeut, 2006, 109: 92–106
4 Johnson H E, Broadhurst D, Goodacre R, et al. Metabolic fingerprinting of salt-stressed tomatoes. Phytochemistry, 2003, 62: 919–928
5 Urbanczuk-Wochniak E, Luedemann A, Kopka J. et al. Parallel analysis of transcript and metabolic profiles: A new approach in systems biology. EMBO Rep, 2003, 4: 989–993
6 Zhou J L, Xu B, Shi B, et al. A metabonomics approach to analyze the dexamethasone-induced clef palate in mice. J Biochem Biotech, 2011, 2011: 509043
7 Wang G J, Zha W B, Hao H P, et al. Metabonomics and its application prospect in TCM study. Chin J Nat Med, 2008, 6: 89–97
8 Ma C, Bi K S, Zhang M, et al. Metabonomic study of biochemical changes in the urine of Morning Glory Seed treated rat. J Pharmaceut Biomed Anal, 2010, 53: 559–566
9 Marchesi J R, Holmes E, Khan F, et al. Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. J Proteome Res, 2007, 6: 546–551
10 Chen M, Wang Y Q, Zhao Y, et al. Dynamic Proteomic and metabonomic analysis reveal dysfunction and subclinical injury in rat liver during restraint stress. Biochim Biophys Acta, 2009, 1794: 1751–1765
11 Taylor J, King R D, Altmann T, et al. Application of metabonomics to plant genotype discrimination using statistics and machine learning. Bioinformation, 2002, 18: 241–248
12 Fiehn O. Metabonomics—the link between genotypes and phenotypes. Plant Mol Biol, 2002, 48: 155–171
13 Castro C, Manetti C. A multiway approach to analyze metabonomics data: A study of maize seeds development. Anal Biochem, 2007, 371: 194–200
14 Rull A, Vinaixa M, Angel Rodríguez M, et al. Metabolic phenotyping of genetically modified mice: An NMR metabonomic approach. Biochimie, 2009, 91: 1053–1057
15 Rezzi S, Ramadan Z, Fay L B, et al. Nutritional metabonomics: Applications and perspectives. Proteome Res, 2007, 6: 513–525
16 Favé G, Beckmann M E, Draper J H, et al. Measurement of dietary exposure: A challenging problem which maybe overcome thanks to metabonomics? Genes Nutr, 2009, 4: 135–141
17 Newby P K, Weismayer C, Akesson A, et al. Long-term stability of food patterns identifies by use of factor analysis Swedish Women. J Nutr, 2006, 136: 626–633
18 Peré-Trepat E, Ross B A, Martin F, et al. Chemometric strategies to asses metabonomic imprinting of habits in epidemiological studies. Chemomet Intell Lab Syst, 2010, 104: 95–100
19 Pedersen M S, Nielsen C N, Andersen H J, et al. Probiotic acidified milk in irritable bowel syndrome patients is indistinguishable from that of non-probiotic acidified milk by ^{1}H-NMR-based metabonomics analysis. Nutrients, 2010, 2: 1141–1155
20 Bundy J G, Willey T L, Castell R S, et al. Discrimination of pathogenic clinical isolates and laboratory strains of *Bacillus cereus* by NMR-based metabolomic profiling. FEMS Microbiol Lett, 2005, 242: 127–136
21 Carroll J W, Mateescu M C, Chava K, et al. Response and tolerance of toxigenic *Vibrio cholerae* O1 to cold temperatures. Antonie Van Leeuwenhoek, 2001, 79: 377–384
22 Colwell R R. Global climate and infectious disease: The cholera paradigm. Science, 1996, 274: 2025–2031
23 Jians X, Chai T. Survival of *Vibrio parahaemolyticus* at low temperatures under starvation conditions and subsequent resuscitation of viable nonculturable cells. Appl Environ Microbiol, 1996, 62: 1300–1305
24 Kaprelyants A S, Kell D B. Dormancy in stationary-phase cultures of *Micrococcus luteus*: Flow cytometry analysis of starvation and resuscitation. Appl Environ Microbiol, 1993, 59: 3187–3196
25 Fisher-Le Saux M, Hervio-Heath D, Loaec S, et al. Detection of cy-

totoxin-hemolysin in nonculturable populations of environmental and clinical *Vibrio vulnificus* strains in artificial seawater. Appl Environ Microbiol, 2002, 68: 5641–5646
26 Heim S, Lleo M M, Bonato B, et al. The viable but nonculturable state and starvation are different stress responses of *Enterococcus faecalis*, as determined by proteome analysis. J Bacteriol, 2002, 184: 6739–6745
27 Signoretto C, Lleò M M, Canepari P. Modification of the peptidoglycan of *Escherichia coli* in viable but nonculturable state. Curr Microbiol, 2002, 44: 125–131
28 Signoretto C, Lleò M M, Tafi M C, et al. Cell wall chemical composition of *Enterococcus faecalis* in the viable but nonculturable state. Appl Environ Microbiol, 2000, 66: 1953–1959
29 Oliver J D, Bockian R. *In vivo* resuscitation and virulence towards mice of viable but nonculturable cells of *Vibrio vulnificus*. Appl Environ Microbiol, 1995, 61: 2620–2623
30 Fayette J, Coquard I R, Alberti L, et al. ET-743: A novel agent with activity in soft tissue sarcomas. Oncologist, 2005, 10: 827–832
31 Chill I, Rudi A, Aknin M, et al. New sesterteipenes from *Madagascan lendenfeldia* sponges. Tetrahedron, 2004, 60: 10619–10626
32 Mendis E, Rajapakse N, Byun H G, et al. Investigation of junbo squid skin gelatin peptides for their *in vitro* antioxidant effects. Life Sci, 2005, 77: 2166–2178
33 Shen Y C, Chen Y H, Huang T L, et al. Four new briarane diterpenoids from the gorgonian coral *Juceella fragilis*. Helv Chim Acta, 2007, 90: 1391–1398
34 Ajani J A, Jiang Y, Faust I, et al. A multi-center phase II study of sequential paclitaxel and bryosatin-1 in patients with untreated, advanced gastric or gastroesophageal junction Adenocarcinoma. Invest New Drugs, 2006, 24: 35–37
35 Bayet-Robert M, Lim S, Barthomeuf C, et al. Biochemical disorders induced by cytotoxic marine natural products in breast cancer cells as revealed by proton NMR spectroscopy-based metabolomics. Biochem Pharmacol, 2010, 80: 1170–1179
36 Newman D J, Cragg G M. Marine natural products and related compounds in clinical and advanced preclinical trials. J Nat Prod, 2004, 67: 1216–1238
37 Simmons L T, Andrianasolo E, McPhail K, et al. Marine natural products as anticancer drugs. Mol Cancer Ther, 2005, 4: 333–342
38 Davidson S K, Allen S W, Lim G E, et al. Evidence for the biosynthesis of bryostatins by the bacterial symbiont "Candidatus Endobugula sertula" of the bryozoan *Bugula neritina*. Appl Environ Microbiol, 2001, 67: 4531–4537
39 Roesijadi G. Influence of copper on the clam *Protothaca staminea*: Effects on the gills and occurrence of copper-binding proteins. Biol Bull, 1980, 158: 233–247
40 Matozzo V, Tomei A, Marin M. Acetylcholinesterase as a biomarker of exposure to neurotoxic compounds in the clam Tapes philippinarum from the Lagoon of Venice. Mar Poll Bull, 2005, 50: 1686–1693
41 Elbaz A, Wei Y Y, Meng Q, et al. Mercury-induced oxidative stress and impact on antioxidant enzymes in *Chlamydomonas reinhardtii*. Ecotoxicology, 2010, 19, 1285–1293
42 Zhang L, Liu X, You L, et al. Metabolic responses in gills of Manila clam *Ruditapes philippinarum* exposed to copper using NMR-based metabolomics, Mar Environ Res, 2011, 72: 33–39
43 Andrey P T, Christopher B J, Haakil L, et al. Metabolomic investigations of american oysters using ^{1}H-NMR spectroscopy. Mar Drugs, 2010, 8: 2578–2596
44 Wu H F, Wang W X. Tissue-Specific toxicological effects of cadmium in green mussels (*Perna Viridis*): Nuclear magnetic resonance-based metabolomics study. Environ Toxicol Chem, 2011, 30: 806–812
45 Stal L J, Moezelaar R. Fermentation in cyanobacteria. FEMS Microbiol Rev, 1997, 21: 179–211
46 Sialve B, Bernet N, Bernard O. Anaerobic digestion of microalgae as a necessary step to make microalgal biodiesel sustainable. Biotechnol Adv, 2009, 27: 409–416
47 McKinlay J B, Harwood C S. Photobiological production of hydrogen gas as a biofuel. Curr Opin Biotechnol, 2010, 21: 244–251
48 McNeely K, Xu Y, Bennette N, et al. Redirecting reductant flux into hydrogen production via metabolic engineering of fermentative carbon metabolism in a cyanobacterium. Appl Environ Microbiol, 2010, 76: 5032–5038
49 Dumas M E, Canlet C, André F, et al. Metabonomic assessment of physiological disruptions using ^{1}H-^{13}C HMBC-NMR spectroscopy combined with pattern recognition procedures performed on filtered variables. Anal Chem, 2002, 74: 2261–2273
50 Keun H C, Beckonert O, Griffin J L, et al. Cryogenic probe ^{13}C-NMR spectroscopy of urine for metabonomic studies. Anal Chem, 2002, 74: 4588–4593
51 Lindon J C, Nicholson J K, Holmes E, et al. Contemporary issues in toxicology the role of metabonomics in toxicology and its evaluation by the COMET project. Tox Appl Pharmacol, 2003, 187: 137–146
52 Keun H C, Ebbels T M D, Antti H, et al. Analytical reproducibility in ^{1}H-NMR-based metabonomic urinalysis. Chem Res Toxicol, 2002, 15: 1380–1386
53 Garrod S, Humpfer E, Spraul M, et al. High-resolution magic angle spinning ^{1}H-NMR spectro-scopic studies on intact rat renal cortex and medulla. Magn Reson Med, 1999, 41: 1108–1118
54 Garrod S, Humpher E, Connor S C, et al. High-resolution 1H-NMR and magic angle spinning NMR spectroscopic investigation of the biochemical effects of 2-bromoethanamine in intact renal and hepatic tissue. Magn Reson Med, 2001, 45: 781–790
55 Griffin J L, Pole J C M, Nicholson J K, et al. Cellular environment of metabolites and a metabonomic study of toxin in endometrial cells using gradient high resolution magic angle spinning ^{1}H-NMR spectroscopy. BBA Gen Subj, 2003, 1619: 151–158
56 Soga T, Ohashi Y, Ueno Y, et al. Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. J Proteome Res, 2003, 2: 488–494
57 Spraul M, Freund A S, Nast R E, et al. Advancing NMR sensitivity for LC-NMR-MS using a cryoflow probe: Application to the analysis of acetaminophen metabolites in urine. Anal Chem, 2003, 75: 1536–1541
58 Kammerer B, Scheible H, Zurek G, et al. *In vitro* metabolite identification of ML3403, a 4-pyridinylimidazole-type p38 MAP kinase inhibitor by LC-Qq-TOF-MS and LC-SPE-cryo-NMR/MS. Xenobiotica, 2007, 37: 280–297
59 Rezzi S, Vera F A, Martin F P J, et al. Automated SPE-RP-HPLC fractionation of biofluids combined to off-line NMR spectroscopy for biomarker identification in metabonomics S. Chromatogr B, 2008, 871: 271–278
60 Chan E C Y, Koh P K, Cheah P Y, et al. Metabolic profiling of human colorectal cancer using high-resolution magic angle spinning nuclear magnetic resonance (HR-MAS NMR) spectroscopy and gas chromatography mass spectrometry (GC/MS). Proteome Res, 2009, 8: 352–361
61 Smilde A K, van der Werf M J, Bijlsma S, et al. Fusion of mass spectrometry-based metabonomics data. Anal Chem, 2005, 77: 6729–6736
62 Lindon J C, Holmes E, Nicholson J K. Pattern recognition methods and applications in biomedical magnetic resonance. Prog Nucl Magn Reson Spectrosc, 2001, 39: 1–40
63 Wiechert W. Modeling and simulation: Tools for metabolic engineering. Biotechnology, 2005, 94: 37–63