Biochemical and Biophysical Research Communications 386 (2009) 268–272

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

A novel application of metabolomics in vertebrate development

Shunsuke Hayashi [a], Shinichi Akiyama [b], Yutaka Tamaru [b], Yasuaki Takeda [c], Toshinobu Fujiwara [d,e], Kunio Inoue [c], Akio Kobayashi [a], Shingo Maegawa [f,*], Eiichiro Fukusaki [a,*]

[a] Department of Biotechnology, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita 565-0871, Japan
[b] Department of Life Sciences, Faculty of Bioresources, Mie University, Tsu 514-8507, Japan
[c] Department of Biology, Graduate School of Science, Kobe University, 1-1 Rokkodaicho, Nadaku, Kobe 657-8501, Japan
[d] Department of Chemical Science and Engineering, Kobe University, 1-1 Rokkodaicho, Nadaku, Kobe 657-8501, Japan
[e] Precursory Research for Embryonic Science and Technology, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan
[f] Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University, Yoshida-Honmachi, Sakyo, Kyoto 606-8501, Japan

ARTICLE INFO

Article history:
Received 31 May 2009
Available online 11 June 2009

Keywords:
Metabolomics
Zebrafish
Development
Fingerprinting
PLS

ABSTRACT

Many studies have demonstrated the functions of individual genes associated with embryogenesis and have determined the genome sequences of several organisms. Despite the availability of enormous amount of genetic information, dynamic changes that occur during embryogenesis have not yet been completely understood. In order to understand the dynamic processes involved in embryogenesis, we employed the metabolomic approach. The results of our study indicated that there is a close correlation between metabolomes and developmental stages. Our method enables the identification of embryonic stages using metabolomes as "fingerprints." In this manner, we could successfully predict embryonic development on the basis of metabolomic fingerprints. This is the first report describing a model for predicting vertebrate development by using metabolomics.

Introduction

The genomes of several organisms have been sequenced and published in an attempt to understand mechanisms underlying biological activities such as evolution and early embryogenesis. Although the amount of genomic information has been increasing rapidly, the mechanism of biological activities has not yet been completely understood. It is difficult to determine the time and site of activation of individual genes involved in vertebrate development and to understand cell behavior associated with gene function solely on the basis of genomic information. In order to understand the transmission of information encoded in the genome, systematic approaches such as transcriptome analysis using microarray and proteome analysis using mass spectrometry have been developed [1,2]. Both the approaches are important for examining the spatiotemporal pattern of transfer of genetic information via mRNAs and proteins. However, both methods are inadequate for elucidating the dynamic changes occurring during embryogenesis because they offer low resolution and are not robust enough.

Metabolomics is one of the latest "-omic" sciences, which is based on the metabolome—an exhaustive profile of metabolites [3,4]. Thus far, metabolomics was typically applied for the determination of unknown gene functions, in association with genomics, transcriptomics, and/or proteomics. The most important independent application of metabolomics is the identification of biomarkers in specific cell types or tissues [5,6]. The identified biomarkers can be indirectly used for elucidating disease mechanisms, diagnostic targets, etc. [4,6,7]. In principle, the metabolome, particularly the unbiased metabolome, would be more diverse and dynamic in terms of chemical and physical properties of metabolites than the transcriptome and proteome. Therefore, the analysis of the metabolome would be suitable for describing the dynamic changes that occur during embryogenesis. However, there have been no reports on the practical application of metabolomics for determining the mechanisms underlying specific biological processes in higher organisms.

In the present study, we attempted to use metabolomics instead of transcriptomics and proteomics for understanding the mechanisms underlying complicated biological processes occurring during embryogenesis in zebrafish. Metabolites are known to play an important role during early development. For instance, retinoic acid is a key metabolite involved in many developmental processes such as neural differentiation and patterning [8]. In addition, several studies on glucose and lipid metabolism have revealed that alterations in the levels or types of metabolites affect embryonic development and/or cell fate determination [9,10]. Therefore, early embryogenesis was a suitable period for determining whether metabolomics can be used to understand complex biological

* Corresponding authors.
E-mail addresses: shingo@i.kyoto-u.ac.jp (S. Maegawa), fukusaki@bio.eng.osaka-u.ac.jp (E. Fukusaki).

0006-291X/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2009.06.041

processes. We first identified and profiled 63 types of metabolites from 24 developmental stages, i.e., from 1-cell stage to 48 h post-fertilization (hpf), of zebrafish embryos by using gas chromatography/mass spectrometry (GC/MS) method [11]. Analysis of the GC/MS data with partial least square (PLS) regression clearly indicated a good correlation between metabolomes and developmental stages. Next, we developed a model for predicting embryonic stages on the basis of the metabolome. Our novel model is a practical tool to analyze the biological processes in early development.

Materials and methods

Maintenance of zebrafish. Danio rerio adults were obtained from a local pet shop and maintained under the following conditions: at 28.0 ± 0.5 °C and a 14-h/10-h light/dark cycle. They were fed a commercially available artificial diet (TetraMin™ flakes) twice daily.

Embryo collection. Fertilized eggs were collected immediately after spawning, and the embryos were washed several times in system water, maintained at 28.5 °C in the same water, and staged by hpf according to the standard morphological criteria. Embryos were collected at 24 time points (1, 2, 3, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 39, 42, 45, and 48 hpf), immediately frozen using liquid nitrogen, and stored at −80 °C. The 24 time points were classified into the following developmental stages: cleavage period (1 and 2 hpf), blastula period (3 and 4 hpf), gastrula period (6–10 hpf), segmentation period (12–24 hpf), and pharyngula period (26–48 hpf).

Sample preparation. For GC analysis, we collected 50 embryos at each time point. We mixed 1000 μl of 80% methanol diluted in water with 30 μl of 0.2 mg/ml solution of ribitol and used it as an internal standard. The embryos were homogenized with MM 301 mixer mills (Retsch GmbH & Co., KG, Haan, Germany). The homogenate was shaken for 30 min at 37 °C and centrifuged at 16,000 *g* for 3 min at 4 °C. An 800-μl aliquot of the supernatant was transferred to an Eppendorf tube with a pierced cap. The samples were dried in a vacuum centrifuge dryer. For derivatization, 50 μl of methoxylamine hydrochloride in pyridine (10 mg/ml)—the first derivatizing agent—was added to the samples. The mixture was incubated at 30 °C for 90 min. The second derivatizing agent, i.e., 50 μl *N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA), was added, and the mixture was incubated at 37 °C for 30 min. A 1-μl sample was injected in the split mode (10:1, v/v).

GC/MS analysis. Chromatography was performed on the 6890 N system (Agilent Co., Palo Alto, CA) equipped with a fused-silica capillary column (internal diameter (i.d.), 30 m × 0.25 mm) coated with a CP-SIL 8 CB low bleed film (thickness, 0.25 μm; Varian Inc., Palo Alto, CA) and was connected to a Pegasus III TOF mass spectrometer (LECO, St. Joseph, MI) equipped with an autosampler 7683B series injector (Agilent Co., Palo Alto, CA). The injection temperature was 230 °C. The helium gas flow rate through the column was 1 ml/min. The column temperature was isothermally

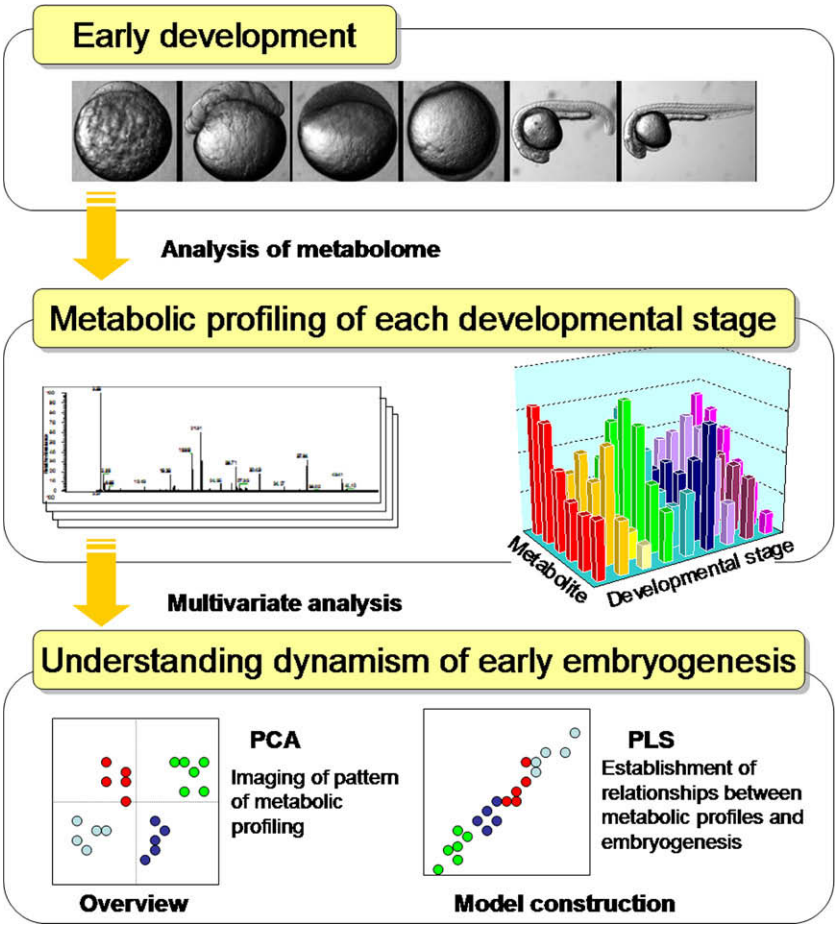

Fig. 1. Paradigm of our novel method developed to understand the dynamic changes involved in early vertebrate development by using metabolic profiling. We believe that alterations in the quality and/or quantity of metabolites reflect the dynamic changes in cell fate and pattern in an embryo during early development. Here, we profiled the metabolomes for each developmental stage by using GC/TOF-MS, as described in the text. Multivariate analysis of the metabolome data indicated that metabolic changes reflected progress in embryonic development, suggesting that biological processes can be analyzed and predicted.

maintained at 80 °C for 2 min and then raised to 330 °C at the rate of 15 °C/min; the temperature was isothermally maintained at 330 °C for 6 min. The temperatures of the transfer line and the ion source were 250 °C and 200 °C, respectively. Ions were generated at an electron impact (EI) energy of 70 kV, and 20 scans/s were recorded over the mass range of 85–500 *m/z*. The acceleration voltage was applied after a 245 s solvent delay.

Data processing and multivariate analysis. For processing GC data, raw chromatographic data (Pegasus file, *.peg) were converted into ANDI files (Analytical Data Interchange Protocol, *.cdf). The ANDI formatted data could be converted and transferred between different mass spectral equipments. The spectral intensity of each compound was determined after detecting the peaks by using the MZmine software program [13]. The software had an in-built Centroid peak detector tool and the parameter settings for the same were as follows: the *m/z* bin size was set to 1.0; the chromatographic threshold level, to 0%; the noise level, to 5; the minimum peak height, to 50; the minimum peak duration, to 0.5; the tolerance for *m/z* variation, to 0.5; and the tolerance for intensity variation, to 10%. The distinctive *m/z* peak for each compound was normalized on the basis of the ribitol peak intensity. Each metabolite was identified on the basis of an in-house chemical library, and then each intensity of same compound was multiplied by 100. Those *m/z* peaks were selected as distinctive *m/z* peaks in the mass spectrogram whose intensities were not affected by neighboring peaks (Supplementary Table 1). PCA was initially applied to understand the relationships among the various groups of multivariate data in terms of similarities or differences. The commercially available Pirouette software (Informetrix, Inc.) was used for this purpose. The projections to latent structures by means of partial least square (PLS) (SIMCA-P version 11.0; Umetrics, Umeå, Sweden) were used to validate the data and create a prediction model. PLS determines the association between two sets of variables: observations and responses.

Results

The metabolome analysis reflects dynamic changes in zebrafish embryogenesis

Zebrafish was chosen as the model animal because it has several advantages [12]. First, we separated the metabolites from 24 different developmental stages of zebrafish by using the GC/MS system (Fig. 1). Data thus obtained were preprocessed using the MZmine software [13]. The data were normalized using an internal standard, and each metabolite was identified on the basis of an in-house chemical library. The integral amount of each metabolite was set to 100 (see Supplementary Table 1). The results of the GC/MS analysis revealed that the quantity and type of metabolites were considerably altered during early development (Supplementary Table 1). In order to identify the entire spectrum of metabolic changes occurring during early development, the abovementioned data were examined by principle component analysis (PCA) (Fig. 1). Currently, PCA is one of the most useful methods for exploratory multivariate analysis and enables the analysis of the internal structure of the data by decreasing the dimensions of the data matrix (Supplementary Figure 1). Because of its advantages, PCA was used to process the metabolome data obtained from each developmental stage. The PCA score plot of factor 1 (eigenvector), which showed the highest contribution (69.4%), clearly revealed a positive separation among metabolite groups corresponding to each developmental stage (Fig. 2A). The data for the majority of the metabolites showed positive values in loading plot 1, suggesting that the amount of most of the metabolites increased as embryogenesis proceeded (Fig. 2B). The analysis of factor 2, which showed a high contribution (5.8%) in the second component, revealed three clusters; this indicated different patterns of alterations in the metabolite composition during development (Fig. 2A). In fact, the three clusters corresponded to three developmental stages (i.e., from 1 to 8 hpf, from 10 to approximately 24 hpf, and from approximately 30 to 48 hpf); this suggested that the composition of metabolites was considerably altered during zebrafish embryogenesis. Therefore, we concluded that a composition of the metabolome is a "fingerprint" of each developmental stage, indicating that dynamic alterations in metabolites coincide with developmental processes.

Our model based on the metabolome obtained from each developmental stage

Next, we attempted to study the relationship between the metabolome and the developmental stage. Since we observed distinct metabolic profiles corresponding to each developmental stage, we hypothesized that we could develop a model for predicting the developmental stage by using metabolomics. To validate

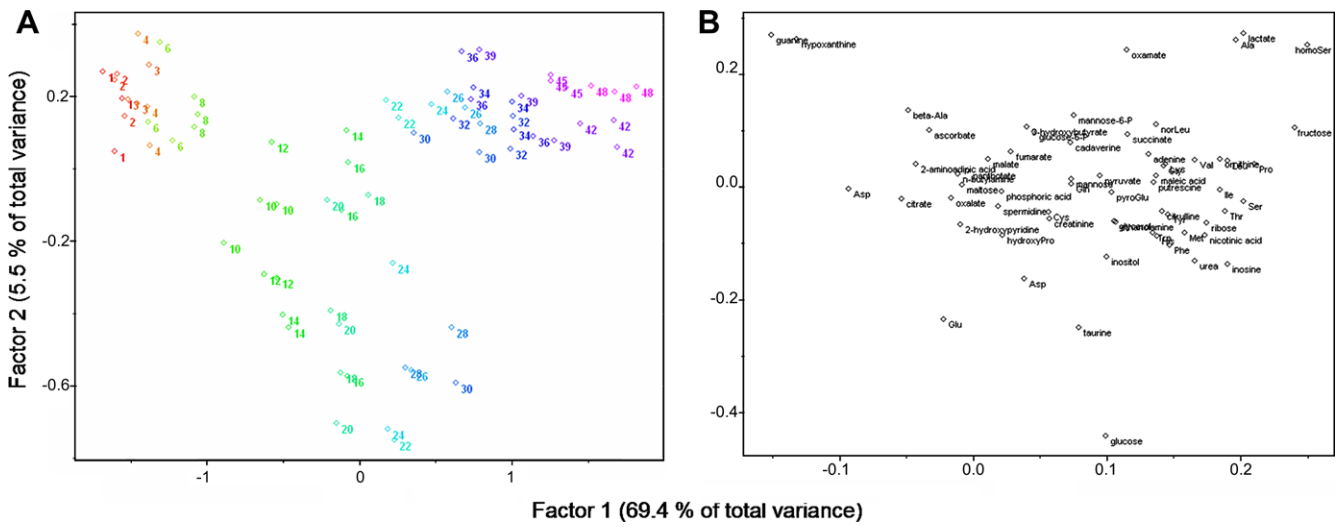

Fig. 2. Cluster of each metabolome during development. (A) Principle component analysis (PCA) revealed distinct metabolome clusters corresponding to each developmental stage. The score plot was created using PCA. The number of spots at each point represents the hours after fertilization. (B) The loading plot was calculated on the basis of the score plot. The loading plot indicates the impact of each metabolite on the results of the cluster analysis.

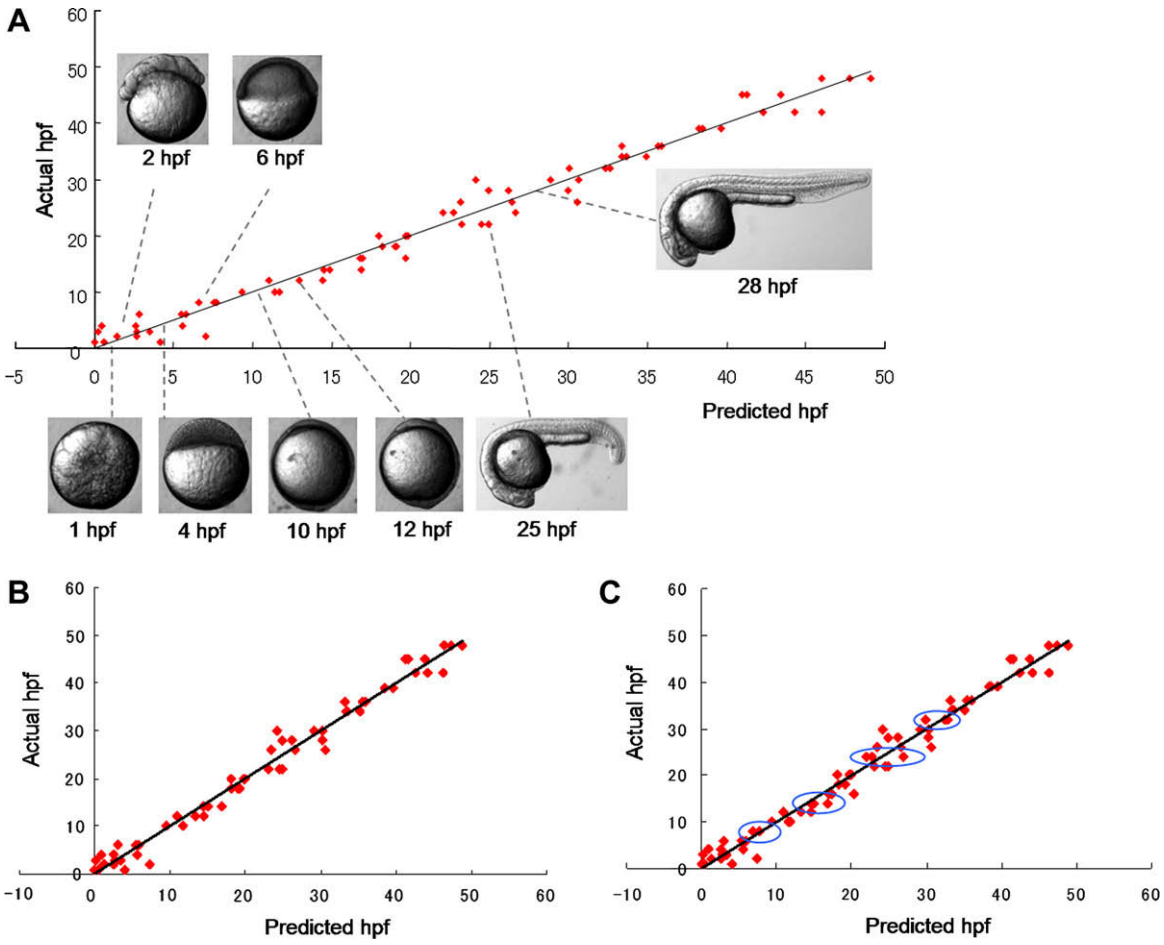

Fig. 3. A model constructed using partial least square (PLS) regression reveals a good correlation between the composition of the metabolome at a particular time and actual embryogenesis. (A) Relationship between actual developmental stages and the metabolic fingerprints. The model was constructed by using the entire gas chromatography (GC) data and demonstrates an obvious correlation between the developmental stages and the metabolomes. Each panel shows a photograph of an embryo at a particular developmental stage. (B) The prediction model including the training set. This model was calculated using the data for all developmental stages except the 8-, 16-, 24-, and 32-h post-fertilization (hpf) stages. (C) Cross-validation of the model. The data for the 8-, 16-, 24-, and 32-hpf stages were fitted onto the prediction model presented in (B). Blue circles correspond to the data for the 8-, 16-, 24-, and 32-hpf stages.

this hypothesis, we analyzed the data by projections to latent structures using partial least square (PLS) regression. PLS is one of the most reliable methods for the regression of multivariate data [14] because decreasing the dimensions of the data matrix is analogous to determining the principal components using PCA (Supplementary Figure 2). We could construct a regression model with high linearity because R^2 was 0.98 (Fig. 3A). In order to determine the accuracy of our model, we excluded the data from 8-, 16-, 24-, and 32-hpf stages and used the remaining data as the training set for PLS (Fig. 3B). For cross-validation, we used the data for 8-, 16-, 24-, and 32-hpf stages as the test set. The PLS results for the training set revealed high linearity without the test set, and the test set perfectly fit into the predicted regression line (Fig. 3C). Furthermore, there was no difference between the root mean square error for the entire data set (RMSEE) and that for the training data set (RMSEP) (Table 1). On the basis of these results, we could finally conclude that our prediction model was valid and that the linearity of our model was high enough to predict the developmental stages.

Table 1
Validated prediction model using partial least square (PLS) regression. RMSEE and RMSEP are root mean square errors of the fit for observations using the training set and of the fit of the prediction for observations using the test set, respectively.

	Latent factors	R2Y	RMSEE	RMSEP
Identified compounds				
Entire data	2	0.98	2.094	–
Cross-validation[a]	2	0.98	2.159	1.887

[a] Test set were 8-, 16-, 24-, and 32-hpf samples, and the other samples were included in the training set.

Discussion

In the present study, we attempted to use the metabolomics approach for elucidating early developmental processes. The results demonstrated a novel application of metabolomics in understanding the embryogenesis in zebrafish. With the progression of embryogenesis, the metabolome exhibited dynamic changes, indicating that the types and/or quantities of metabolites are correlated with the biological activities occurring during development. On the basis of the good correlation between the metabolome and embryogenesis, we concluded that the metabolome can be used as a fingerprint of a particular developmental process.

Thus far, several studies have attempted to analyze vertebrate embryogenesis using the transcriptomic or the proteomic approach [15,16]. Although these approaches are partially useful in elucidating the complex developmental processes, they cannot be used for the complete representation of dynamic changes occurring during embryogenesis because of low resolution and poor

robustness. Here, we successfully demonstrated the dynamic changes involved in early vertebrate development with sufficient accuracy and precision by using the metabolomic approach. Moreover, our results clearly indicated that the metabolomic approach has sufficient resolution and robustness to describe dynamic biological events and is a powerful tool to analyze early embryogenesis as well as other complicated biological processes. In the previous studies conducted on zebrafish as well as other vertebrates, morphology and gene expression patterns have been used as phenotypic parameters to analyze developmental events and mutants [17]. In addition to these traditional phenotypic parameters, metabolomics offers a new class of phenotypic parameters for analyzing early developmental processes. Metabolomics approach will enable identification of mutants exhibiting no visible defects and clarification of novel aspects of known mutants. The new phenotypic parameter—the metabolome—may reveal unknown genetic networks involved in early development.

Acknowledgments

This work was supported by a Grant-in-Aid for Scientific Research for Priority Areas (Life Surveyor). S.M. was supported by a Grant-in-Aid for Young Scientists (Start-up) (No. 20870024).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2009.06.041.

References

[1] W.J. Longabaugh, E.H. Davidson, H. Bolouri, Computational representation of developmental genetic regulatory networks, Dev. Biol 283 (2005) 1–16.
[2] A. Bachi, T. Bonaldi, Quantitative proteomics as a new piece of the systems biology puzzle, J. Proteomics 71 (2008) 357–367.
[3] S.G. Oliver, M.K. Winson, D.B. Kell, F. Baganz, Systematic functional analysis of the yeast genome, Trends Biotechnol. 16 (1998) 373–378.
[4] J.C. Lindon, E. Holmes, J.K. Nicholson, Metabonomics in pharmaceutical R&D, FEBS J. 274 (2007) 1140–1151.
[5] J.K. Nicholson, J.C. Lindon, E. Holmes, "Metabonomics": understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data, Xenobiotica 29 (1999) 1181–1189.
[6] N. Kaplowitz, Idiosyncratic drug hepatotoxicity, Nat. Rev. Drug Discov. 4 (2005) 489–499.
[7] D.I. Ellis, W.B. Dunn, J.L. Griffin, J.W. Allwood, R. Goodacre, Metabolic fingerprinting as a diagnostic tool, Pharmacogenomics 8 (2007) 1243–1266.
[8] M. Maden, Retinoic acid in the development, regeneration and maintenance of the nervous system, Nat. Rev. Neurosci. 8 (2007) 755–765.
[9] M. Barceló-Fimbres, G.E. Seidel Jr., Effects of either glucose or fructose and metabolic regulators on bovine embryo development and lipid accumulation in vitro, Mol. Reprod. Dev. 74 (2007) 1406–1418.
[10] R. Dumollard, Z. Ward, J. Carroll, M.R. Duchen, Regulation of redox metabolism in the mouse oocyte and embryo, Development 134 (2007) 455–465.
[11] W. Pongsuwan, E. Fukusaki, T. Bamba, T. Yonetani, T. Yamahara, A. Kobayashi, Prediction of Japanese green tea ranking by gas chromatography/mass spectrometry-based hydrophilic metabolite fingerprinting, J. Agric. Food Chem. 55 (2007) 231–236.
[12] W. Driever, S. Dtemple, A. Schier, L. Solnica-Krezel, Zebrafish: genetic tools for studying vertebrate development, Trends Genet. 10 (1994) 152–159.
[13] M. Katajamaa, M. Oresic, Data processing for mass spectrometry-based metabolomics, J. Chromatogr. A 1158 (2007) 318–328.
[14] L.P. Ammann, M. Merritt, A. Sagalowsky, P. Nurenberg, Peak-finding partial least squares for the analysis of ^{1}H NMR spectra, J. Chemometrics 20 (2006) 231–238.
[15] M.G. Katz-Jaffe, D.W. Linck, W.B. Schoolcraft, D.K. Gardner, A proteomic analysis of mammalian preimplantation embryonic development, Reproduction 130 (2005) 899–905.
[16] S. Mathavan, S.G. Lee, A. Mak, L.D. Miller, K.R. Murthy, K.R. Govindarajan, Y. Tong, Y.L. Wu, S.H. Lam, H. Yang, Y. Ruan, V. Korzh, Z. Gong, E.T. Liu, T. Lufkin, Transcriptome analysis of zebrafish embryogenesis using microarrays, PLoS Genet. 1 (2005) 260–276.
[17] T. Sakaguchi, A. Kuroiwa, H. Takeda, A novel sox gene, 226D7, acts downstream of Nodal signaling to specify endoderm precursors in zebrafish, Development 107 (2001) 25–38.