Accepted Manuscript

The relations between metabolic variations and genetic evolution of different species

Zhishui Li, Chenghong Lin, Jingjing Xu, Huifeng Wu, Jianghua Feng, Heguang Huang

Subject category: Metabolite Determination

The relations between metabolic variations and genetic evolution of different species

Short title: **Metabolic variations of different animal species**

Zhishui Li[a], Chenghong Lin[a], Jingjing Xu[a], Huifeng Wu[b], Jianghua Feng[a*], Heguang Huang[c]

[a]Department of Electronic Science, Fujian Provincial Key Laboratory of Plasma and Magnetic Resonance, Xiamen University, Xiamen, 361005, China

[b]Key Laboratory of Coastal Zone Environment Processes, CAS, Shandong Provincial Key Laboratory of Coastal Zone Environment Processes, Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai, 264003, China

[c]The Basic Surgical Department, Fujian Medical University Union Hospital, Fuzhou, 350001, China

[*] Correspondence to: Jianghua Feng

Email: jianghua.feng@xmu.edu.cn, Tel: +86-592-2186393, Fax: +86-592-2189426

Abstract

Metabonomics has been applied into many bio-related scientific fields. Nevertheless, some animal researches are proved to be failed when they are extended to human beings. Therefore it's essential to figure out suitable animal modeling to mimic human metabolism so that animal findings could serve human beings. In this study, two kinds of commonly-selected body fluids (serum and urine) from human being and various experimental animals were characterized by integration of NMR spectroscopy with multivariate statistical analysis to identify the interspecies metabolic differences and similarities at a baseline physiological status. Our results highlight that dairy cow and piggy may be an optimal choice for the transportation and biodistribution study of drug and KM mouse model may be the most effective for the excretion study of drug, while SD rat could be a most suitable candidate for animal modeling under overall consideration. The biochemical pathways analyses further provide inner connection between genetic evolution and metabolic variations, where species evolution most strongly affects microbial biodiversity, and consequently takes effects on the species-specific biological substances biosynthesis and corresponding biological activities. Knowledge of the metabolic effects from species difference will enable the construction of better models for the disease diagnosis, drug metabolism and toxicology research.

Keywords: NMR; metabonomics; physiological variation; laboratory animals; genetic evolution

Introduction

With increasing frequency of application of metabonomics in the bio-related fields such as disease early detection, drug metabolism and toxicology research [1-3], it is necessary and important to characterize and compare the metabolic similarities and dissimilarities between humans and various laboratory animals in the control and baseline status so that the results from animal models could serve the human beings. Animal modeling does help human beings to understand the mechanism of the disease development, drug screening and some other biological functions before the research can be extended to clinical application. Among a great amount of animal models, some are successful and they represent excellent agreement with human studies and thus exert a positive impact on human community [4, 5]. Unfortunately, it is a fact that some animal models fail when their results are used in the human body. Therefore it is vital to select the proper types of experiment animals to improve the chance of success for extension of the findings from animals to the human.

Usually, the primary concerns about the animal selections focus on the experimental budget, research objects, the availability of the animal and some other reality factors such as environment facility control [6] in the traditional animal experiments. Further, a few researchers would think about the similarity of some certain immune reaction between animals and the human from the genetic level and analyze the animal order when considering animal selection [7]. But these considerations often lack the fundamental basis from the metabolite level and further experimental validation although it is very important in the metabonomics-based

3

studies. Commonly, the objective animals in the modeling construction will be among mouse, rat, rabbit, piggy, guinea pig, fish, canine, ferine and rhesus monkey [8]. The animal species should be chosen according to the purpose of mimicking drug stimulation, disease processes and gene function well in the humans, and it is quite necessary to characterize the baseline physiological variations and ensure the boundaries of metabolic normality between human being and different animal species in order to be confident of and accurately understand drug- or disease-induced response. Considering the above facts and the fact that serum and urine samples are the mostly used biological samples in metabonomics study, we collected the serum samples from various species including human, dairy cow, guinea pig, KM mouse, piggy and SD rat and the urine samples from human, Balb/c mouse, KM mouse, rabbit, SD rat and Wistar rat to characterize and interpret the metabolic and biochemical correlations of interspecies by NMR spectroscopy.

The NMR spectrum of a biological sample enables the simultaneous identification and monitoring of hundreds of low molecular weight endogenous metabolites, thus providing a biochemical profile of an organism [9]. However, the NMR resonance of some metabolites may be overlapped with each other, which will lead the difficulty in analyzing the metabolite of interest. Multivariate statistical analysis facilitate the comparison of NMR spectra of serum or/and urine of human and animals, enabling the establishing of variations in metabolic profiles and highlighting their metabolic relations at metabolite level [10]. Thus metabonomic method based on information-rich NMR spectroscopic data can be used to evaluate

4

normal physiological variations and similarity and further understand their evolution relations between human beings and laboratory animals [11]. Therefore in this paper, we try to compare the metabolic similarity and dissimilarity between human and these commonly used experimental animals to study their metabolic and evolution relation at the baseline status by using NMR-based metabonomics. The purpose of this study is to identify the most suitable animal model for the bio-related especially metabonomics-related research community.

Materials and methods

Animals and biofluid sample collections

The animal study protocol was in accordance with the Guidelines for the Care and Use of Laboratory Animals of the Chinese Animal Welfare Committee and approved by the Ethical and Research Committee of Xiamen University. The healthy experimental animals including rats, mouse, guinea pig and rabbit were obtained from Xiamen University Laboratory Animal Center, and piggies and dairy cows were obtained from Shanghai Biotree Biotechnique Co. Ltd. Animals were allowed to acclimate for at least one week before initiating experiment. As two characteristic biofluids that are easily collected and the most frequently used in metabonomic analysis, serum and urine were collected respectively from human beings and laboratory animals. The serum samples were collected from dairy cow, guinea pig, Kunming [3] mouse, piggy and Sprague-Dawley (SD) rat (n=8), respectively. And the urine samples were collected from Balb/c mouse, KM mouse, rabbit, SD rat and Wistar rat (n=8), respectively. The serum and urine samples of human being were

5

selected from our previous work [12]. The animals and human beings were fasted overnight before collecting serum samples and all the urine samples were collected between 8-12 am to avoid metabolic changes caused by the circadian rhythms.

Preparation of biofluid samples

Serum samples were prepared by mixing 400 μL of serum with 200 μL of a 90 mM phosphate buffer (pH 7.4) in 0.9% deuterated saline solution. The serum-buffer mixture was centrifuged at 10000 rpm under the condition of 4 °C for 10 min. And then 550 μL of the supernatant was transferred into a 5 mm NMR tube.

Urine samples were prepared by mixing 400 μL of urine with 200 μL of 1.5 M of deuterated phosphate buffer (NaH_2PO_4 and K_2HPO_4, containing 0.05% of TSP (sodium 3-(trimethylsilyl)-2,2,3,3-$^2H_4$ propionate), pH 7.4). The urine-buffer mixture was left to stand for 5 min at room temperature and then centrifuged at 10000 rpm at 4 °C for 10 min to remove suspended debris. The supernatant (550 μL) was then transferred into a 5 mm NMR tube.

^{1}H NMR spectroscopic analysis

The serum or urine samples were analyzed randomly using proton (^{1}H)-NMR spectroscopy, respectively. All spectra of the serum samples were acquired at 298 K on a Varian INOVA-600 NMR spectrometer (operating at 599.92 MHz for ^{1}H) equipped with an inverse detection probe with a shielded Z-gradient. And all spectra of the urine samples were analyzed at 298 K using a Bruker AMX-600 spectrometer equipped with a TXI CryoProbe operating at 600.13 MHz.

NMR spectra of the serum samples were acquired using a water-suppressed

6

Carr-Purcell-Meiboom-Gill (CPMG) spin-echo pulse sequence (cpmgt2pr) $(RD-90°-(\tau-180°-\tau)_n-Acq)$. Spin-echo loop time $(2n\tau)$ of 70 ms was applied with a relaxation delay 2.0 s. A total of 96 transients were collected into 16384 points over a spectral width of 8,000 Hz. A standard one-dimensional spectrum using the first increment of the NOESY pulse sequence $(RD-90°-t_1-90°-t_m-90°-Acq)$ with water suppression (noesypr1d) was acquired for the urine samples. The spectra were acquired with a relaxation delay of 2.0 s at a fixed interval t_1 of 4 μs. The water resonance is irradiated during relaxation delay and the mixing time t_m of 100 ms. A total of 96 transients were collected into 16384 data points with a spectral width of 12000 Hz.

NMR spectral processing

The NMR spectra were processed by using MestReNova (V7.1). All free induction decays (FIDs) were multiplied by an exponential weighting function with a 1.0 Hz line broadening factor prior to Fourier transformation. All of the ^{1}H-NMR spectra were then phased by manual correction and baseline corrected by using the "Bernstein Polynomial Fit" method with a respective polynomial order for the purpose of best baseline correction effect. ^{1}H-NMR spectra of serum and urine samples were respectively referenced to the internal lactate CH_3 resonance at 1.33 ppm and the single peak of TSP at 0 ppm. After the above operation, the chemical shift ranges including the resonances from water and urea were manually removed to eliminate baseline effects of imperfect water signal and urea signal. The spectra over the ranges of 0.5-9.0 ppm for serum samples and 0.5-10.0 ppm for urine samples were

7

integrated automatically into regions with equal width of 0.005 ppm. The integrated data were finally normalized to the total area of the NMR spectrum to compensate for the overall concentration differences before statistical analysis.

Multivariate statistical analysis

Multivariate statistical analyses were performed by using SIMCA-P$^+$ software (version 11.0.0.0, Umetrics, Sweden). Principal component analysis (PCA) was initially carried out by using a par-centered approach for both serum and urine samples to examine class clustering and find possible outliers. The samples were visualized by using the principal component (PC) score plots, where each point represented an individual sample.

After PCA, orthogonal partial least square discriminant analysis (OPLS-DA) was separately applied to the samples from human being and animals with unit-variance (UV) scaling to identify metabolic differences between them. The interpretation of each model was facilitated by a back-transformation of the loadings incorporating color-coded correlation coefficients of the metabolites responsible for the discrimination. The coefficient plots were generated with MATLAB(version 7.7.0.471) scripts with some in-house modifications and were color-coded with the absolute value of coefficients (r). In the plots, a hot color (i. e., red) corresponds to the metabolite being highly different between classes, while a cool color (i. e., blue) corresponds to no obvious differences. In our case, the correlation coefficient of $|r| >$ 0.900 was used as cutoff value for the close relation based on the discrimination significance at the level of p=0.05. To check the validity of the models and avoid the

8

over-fitting, the OPLS-DA models were evaluated by 8-fold cross-validation parameters R^2 and Q^2 and random permutation tests (permutation number = 200) of the corresponding PLS-DA models. The R^2 values describe how well the data fit with the derived models, whereas Q^2 values describe the predictive ability of the derived models and provide a measure of the model quality. The random permutation test is a way that challenges a model by using multiple misclassifications to decide whether a model could have been developed by random chance. According to the overview plots, some pretty important messages could be dug out.

Results and discussion

Overall metabolic profiles of serum and urine from human and different laboratory animals

Various physiological factors, such as differences in gender, age, rearing history, hormonal status, geographic origin and alterations in physical space, diurnal variation, diet, rate of metabolism, as well as stress, will take influences on the metabolic profiles of experimental animals [13], however the effects of inter-species or inter-strains play a vital and most distinguishing role in the variations of biofluids composition. Thus the difference at the level of gene expression inevitably causes the proteomic especially metabonomics changes.

Serum and urine are the most commonly used and easily collected biological fluids in metabonomics study especially for the clinical use. All kinds of metabolites containing in serum and urine from healthy laboratory animals could reflect various

9

intrinsic physiological factors such as genetic drift, well-being, hormonal differences, and rate of metabolism [14].

In this study, we selected a representative NMR spectrum from each kind of animals and human and demonstrated their serum and urinary profiles in Fig.1. Resonance assignments were performed according to the related literature [15, 16, 17] and confirmed by public (Human Metabolome Database V3.0, see www.hmdb.ca) and in-house developed NMR databases. The assignable metabolites were marked in the spectra to give a direct visual comparison of metabolic profiles between human and the different animals.

The typical NMR spectra of serum obtained from dairy cow, KM mouse, piggy, SD rat, guinea pig, and human were demonstrated in Fig.1A. And the NMR spectra of urine obtained from Balb/c and KM mouse, rabbit, SD and Wistar rat, and human were demonstrated in Fig.1B. From the comparison, although the spectral profiles of the same biofluids from difference species or strains were broadly similar, some obvious differences were observed in certain spectral regions among them. For example, the spectra of serum from human and dairy cow contained similar resonances from various metabolites, however, higher levels of acetate, creatine, and phenylalanine were observed in dairy cow serum. Similarly, the spectra of urine from same species but different strains (such as Balb/c and KM mouse, SD and Wistar rat) showed remarkably similar profiles, but obvious differences were also visible in the spectra, for example, KM mouse was characterized by higher levels of homogentisate and phenylacetylglycine and lower levels of butyrate than Balb/c mouse, while higher

10

levels of N-methylnicotinamide were observed in Wistar Rat than SD rat.

Obviously visual comparison could not provide too much information on the similarity and difference between human and various animals, and multivariate statistical analysis will be helpful to give more details of metabolic information on species variations.

Interspecies differences and similarities in serum and urine profiles via PCA analysis

The principal components analysis (PCA), an unsupervised pattern recognition technique, decreases the original super-dimension data matrix to a nearly no-side-effect level such as 2D or 3D to establish new component axis called principal components (PCs) containing as much information as the original data [18]. PCA is very commonly used to see data sets with multiple descriptors, such as genomics and metabonomics. The spectral data from biological samples can be visualized by plotting the PC score. In the PCA score plot, the closer spatial distance between two classes, the closer their biochemical relation.

In our case, PCA was performed on NMR data of serum and urine samples from 6 different kinds of animals. The scaling for PCA is par method which is denoted as equation below

$$par = \frac{x_{ij} - \overline{x_j}}{\sqrt{s_j}} \tag{1}$$

Where $x_{ij}, \overline{x_j}, s_j$ respectively denote the original matrix data, column mean value and column standard deviation value. By such scaling, the data will be centred and

11

standardized which suggests the small signals will have the same projection importance with the strong signals. The metabolic relation can be obtained from the PCA score plot, and in our work the first three primary components contain 78.2% and 53.5% original data information for serum and urine samples, respectively. According to the 3D PCA score plot of serum (Fig.2A), it is clear enough that the piggy, dairy cow, and rat have a relatively closer space distance with the human being, which implied that they have closer metabolic similarity. And the order of space distance from animals to human being (Fig.2A) was in accordance with the animal selection in traditional animal modeling. This result conformed to the previous related result in genetic research that the piggy and rodents, rat for instance, have many close connections with the human being during the human evolution history [19; 20]. However, in the 3D PCA score plot of urine (Fig.2B), KM mouse, Balb/c mouse, and SD rat, instead of the piggy, dairy cow and rat, have a relatively closer distance with the human being.

The PCA score plot (Fig.2) based on the serum and urinary NMR spectral data from various species illustrated not only the similarities between species but also the degree of inter-species variation in serum and urine metabonomes. By metabonomic analysis in a 3D reduction of original biological data (Fig.2), it comes to the conclusion that the dairy cow, pig and rat have the most similar metabolic control profiles of serum to the humans, while KM and Balb/C mouse and SD rat are the most similar to the humans in the urinary baseline metabonomes, which indicate we should choose the proper animals according to the different research aim. For example, dairy

12

cow and piggy may be an optimal choice for the transportation and biodistribution study of drug, and KM mouse model may be the most effective for the excretion study of the drug. Notwithstanding that, SD rat could be a more proper and suitable candidate for animal modelling as the possible replacement of human being at metabonome level in metabonomic research if taking general factors including metabolic relations, experimental budget, animal availability, daily administration, laboratory facility and easier handling of animals into account.

Metabolic variations between different species

To get a deeper understanding of the significant variations and similarities between different species, orthogonal partial least square discrimination analysis (OPLS-DA), a supervised method, was constructed on the NMR data of selected pair-wise species. Based on the genetic differences, the OPLS-DA score plots (left panels in Fig.3) give a clear separation of serum (Fig.3A) and urine (Fig.3B metabonomes between different species or strains, and the corresponding loading plots (middle and right panels in Fig.3) offer an insight into the types of metabolites responsible for the separation. And the significant species-discriminating metabolites corresponding to these pair-wise groups are tabulated in Table S1 in the supplemental materials. To check the validity of the models and avoid the over-fitting, the OPLS-DA models were evaluated by 8-fold cross-validation parameters R^2 and Q^2 and random permutation tests of the corresponding PLS-DA models. The parameters of R^2X, R^2Y and Q^2 (Fig.3) and the results of permutation tests (Fig.S1 in the supplemental materials) suggested that all the models are valid and robust enough,

13

implying that genetic difference plays an important role in differentiating species.

Nine species pairs (human-mouse, human-rat, human-piggy, mouse-rat, rat-guinea pig from serum metabonomes and Balb/c mouse-human, rabbit-human, Wistar rat-human and KM mouse-human from urine metabonomes) and two strain pairs (KM mouse-Balb/c mouse and SD rat-Wistar rat from urine metabonome) were selected from the serum and urine metabonomes in order to get the detailed metabolic information of species variations. A higher correlation coefficient of more than 0.900 (i. e. a high correlation) was set as the cutoff value for the statistical significance to achieve the metabolic pathways highly related with species variations. As shown in Fig.3, the metabolites marked on the positive direction of the loading plots represent that these metabolite levels are higher in the groups located in the positive direction of first PCs in the score plots, and vice versa. For example, serum metabonome of rat was highlighted by higher levels of lactate, glucose, creatine and methionine and lower levels of N-acetylglycoproteins (NAS), N, N-dimethylglycine (DMG), phosphocholine (PC), glycerol, 3-hydroxybutyrate, *myo*-inositol, methanol, ethanol, formate, some lipids and amino acids including glutamate, glutamine, histidine, leucine, tyrosine and valine compared with that of human being (Fig.3A). Higher level of N-acetylglutamate [1] and lower levels of methylmalonate, N-methylnicotinamide (NMN), phenylacetylglycine (PAG), betaine, ethanolamine and taurine were presented in the urine metabonome of SD rat than that of Wistar rat (Fig.3B).

Metabolic pathways involved in species variations and species evolution

Pathways of these differential metabolites were analyzed using KEGG database. According to the KEGG pathway analysis, these metabolites in the serum corresponding to pair-wise species differences were basically involved in microbial metabolism in diverse environments, biosynthesis of biological substance such as secondary metabolites, amino acids, alkaloids and aminoacyl-tRNA, ABC transporters, protein digestion and absorption, and carbon metabolism and carbon fixation pathways (Table 1A). Some specific amino acids, carboxylic acid metabolisms and degradation were also involved in the different pair-wise species, for example, the differences in glycine, serine and threonine metabolism and degradation of aromatic compounds were demonstrated in the serum metabonomes of human and mouse, and the difference in glyoxylate and dicarboxylate metabolism was shown in the serum metabonomes of human and piggy, while more specific variations in metabolic pathways, including arginine and proline metabolism, alanine, aspartate and glutamate metabolism, and glyoxylate and dicarboxylate metabolism, were displayed in the serum metabonomes of rat and guinea pig.

With the help of KEGG database, the metabolic pathways for the above 6 pair-wise species models could be also derived from their differential metabolites in the urine metabonome (Table 1B). Similarly, almost of all species variations involved in microbial metabolism in diverse environments except the model between SD rat and Wistar rat. The most obvious differences were found in the human-rabbit species pair, in which many metabolic pathways were affected including biosynthesis of secondary metabolites, ABC transporters, carbon metabolism and carbon fixation,

15

alanine, aspartate and glutamate metabolism, glyoxylate and dicarboxylate metabolism, and degradation of aromatic compounds, whereas minor difference in metabolic pathways were observed between human and mouse (including KM mouse and Balb/c mouse), and SD rat and Wistar rat. And surprisingly, major differences in metabolic pathways were found between KM mouse and Balb/c mouse, which include biosynthesis of biological substance including secondary metabolites and alkaloids, carbon fixation pathways, degradation of aromatic compounds, and a series of amino acid metabolisms such as phenylalanine metabolism, tyrosine metabolism, and alanine, aspartate and glutamate metabolism.

Although the differentiated metabolites are involved in a good number of biological pathways in the pair-wise animals we concerned, the most primary variations are microbial metabolism in diverse environment and biosynthesis of secondary metabolites except the model of SD rat and Wistar rat (Table 1B). It's obvious that the metabonome from different species demonstrated species-specific levels of the metabolites associated with microbial environment and animal evolution [21]. Given the diversity in animal species found from the different evolution background, their distinct microbial communities suggest a different biological role in animal development [22]. Further, secondary metabolites are unevenly distributed in diverse microbial environment and seem closely associated with their biological activity that endows the producer with increased fitness at the stage of species evolution [23]. The characteristic secondary biosynthetic pathways and biosynthesis of other species specific biological compounds such as amino acids, alkaloids and

16

aminoacyl-tRNA should be consistent with their evolution in the different environmental challenges [24]. Consequentially, the need of energy and nutrients (e. g. carbon) in such challenges will definitely involve carbon metabolism and carbon fixation as demonstrated in Table 1B [25]. It should be realized that metabolic similarity between Wistar rat and SD rat and KM mouse and Balb/c mouse (Fig.3B and Table S1B) are closely correlated with their evolution background [26, 27]. Especially, the metabolic differences between SD rat and Wistar rat have not involved microbial metabolism in diverse environment and biosynthesis of secondary metabolites, implying the cultivation and continuity from Wistar to SD rat [28].

The metabolic variations of different animal species also include ATP-binding cassette transporters (ABC transporters) and protein digestion and absorption (Table 1). They are correlated with providing energy and utilizing the energy of ATP to transport a wide variety of substrates across extra- and intracellular metabolic products, lipids and sterols [13, 29, 30, 31]. Consequently, species specific amino acids and carboxylic acids metabolisms and degradation are induced (Table 1). These species specificity suggests that distinctive biological and chemical reactions may be of recent origin. The diversity of products may reflect many roles for which specific biological compounds have been selected at their recent origin and evolution.

The correlation coefficients of metabolites and species variations

In the Table S1 in the supplemental materials, the correlation coefficient of │r│ is an indicator used as statistical significance [32]. The positive and negative coefficients indicate relatively increasing or decreasing levels of the corresponding

17

metabolites in one species group compared with another one. Thus in a model, the lower the coefficient is, the less statistical significance the metabolites behaves in the discrimination. In another word, if the species pair has a relatively closer metabolic relation, the corresponding coefficient of those metabolites should be relatively lower. However, as demonstrated in Fig.4A, not the coefficient of any a metabolite in serum is greater from human-mouse than from human-rat although the metabolic variation between mouse and human is larger than that between rat and human as displayed in Fig.2A. According to Fig.4A, the coefficient line in the H-R model is not lower than in the H-M model completely, but overall, the coefficient line still shows us that the metabolic difference in human-mouse model is larger than in human-rat model. It should be more reasonable and effective to use the summed value of coefficients corresponding to all metabolites as a criterion parameter to sort the models. Accordingly, it comes to the conclusion that rat and piggy have a closer relation with human beings in serum metabonome than mouse and further, rat is closer to human than piggy (Fig.4B). Meanwhile summation of coefficient value in M-R model is close to that in R-G model (Fig.4B), which represents the species variations between mouse and rat are similar with those between rat and guinea pig. This conclusion complies with the space distance relations in metabolic profile (Fig.2A).

The same strategy was employed to the urine metabonomes of species pairs. The MB-H and MK-H models display a relatively small summed coefficient │r│ values (Fig.4C), which means Balb/c mouse and KM mouse have a close connection with human beings compared with Wistar rat and rabbit. A relative small but close

summation of coefficient between SD rat-Wistar rat and Balb/c mouse-KM mouse were observed (Fig.4C), which implies a less intra-species metabolic variation and similar inter-strain metabolic difference.

We understand that the metabolic profiles of animal can be changed by many external factors, but variations based on intrinsic genomic factors, i. e. the difference in animal strains or species are generally much more than the variations in response to external stress. We got to accept that there are some weaknesses in this work. First, the number of each species is not great enough to get an exact comparison, however, we chose a higher cutoff value (here is 0.900) for the statistical significance based on the discrimination significance at the level of $p=0.05$ (Table S1 in the supplemental materials) in order to get accurate and detailed discriminating information. Second, the sum processing algorithm for $|r|$ in every model seems a bit simple, but there are some valuable clues to get the related biological significance rather than space distances of different species. It should be appreciated that this is only a gross result to study the metabolic difference and similarity of different species, however, our results based on the analysis of multivariate statistical analysis and biochemical pathways analysis provide insight into which species evolution most strongly influence microbial biodiversity, and further, it will be helpful to select a suitable modeling animal in the bio-related research community especially metabonomics-related ones.

Acknowledgements

References

[1] G.A.N. Gowda, S. Zhang, H. Gu, V. Asiago, N. Shanaiah, and D. Raftery, Metabonomics-based methods for early disease diagnostics. Expert Rev. Mol. Diagn. 8 (2008) 617-633.

[2] M. Li, J. Wang, Z. Lu, D. Wei, M. Yang, and L. Kong, NMR-based metabonomics approach to study the toxicity of lambda-cyhalothrin to goldfish (Carassius auratus). Aquat. Toxicol. 146 (2014) 82-92.

[3] T. Ramirez, M. Daneshian, H. Kamp, F.Y. Bois, M.R. Clench, M. Coen, B. Donley, S.M. Fischer, D.R. Ekman, E. Fabian, C. Guillou, J. Heuer, H.T. Hogberg, H. Jungnickel, H.C. Keun, G. Krennrich, E. Krupp, A. Luch, F. Noor, E. Peter, B. Riefke, M. Seymour, N. Skinner, L. Smirnova, E. Verheij, S. Wagner, T. Hartung, B. van Ravenzwaay, and M. Leist, Metabonomics in toxicology and preclinical research. Altex-Altern Anim. Ex. 30 (2013) 209-225.

[4] A.M. Bertolotti, F.A. Alvarez, S. Defranchi, M. Alvarez, R.P. Laguens, and R.R. Favaloro, Successful circumferential free tracheal transplantation in a large animal model. J. Invest. Surg. 25 (2012) 227-234.

[5] S. Duty, and P. Jenner, Animal models of Parkinson's disease: a source of novel treatments and clues to the cause of the disease. Brit. J. Pharmacol. 164 (2011) 1357-1391.

[6] X. Liu, Z. Yan, M. Luo, L.N. Zhang, C. Trygg, C. Smith, K. Munson, W. Guggino, B. Bunnell, and J.F. Engelhardt, Species choice for animal modeling cystic fibrosis and gene therapy to the airway. Mol. Ther. 13 (2006) S1-S1.

[7] R.T. Prehn, and L.M. Prehn, Is an immune reaction required for malignant transformation and cancer growth? Cancer Immunol. Immun. 61 (2012) 963-966.

[8] C.M. Bagi, E. Berryman, and M.R. Moalli, Comparative bone anatomy of commonly used laboratory animals: implications for drug discovery. Comparative Med. 61 (2011) 76-85.

[9] S. Kim, and T. Szyperski, GFT NMR, a new approach to rapidly obtain precise high-dimensional NMR spectral information. J. Am. Chem. Soc. 125 (2003) 1385-1393.

[10] Y.-S. Shin, K.-H. Bang, D.-S. In, J.-S. Sung, S.-Y. Kim, B.C. Ku, S.-W. Kim, D. Lee, and H.-K. Choi, Fingerprinting differentiation of astragalus membranaceus roots according to ages using H-1-NMR spectroscopy and multivariate statistical analysis. Biomol. Ther. 17 (2009) 133-137.

[11] H.J. Issaq, Q.N. Van, T.J. Waybright, G.M. Muschik, and T.D. Veenstra, Analytical and statistical approaches to metabonomics research. J. Sep. Sci. 32 (2009) 2183-2199.

[12] J. Xu, S. Yang, S. Cai, J. Dong, X. Li, and Z. Chen, Identification of biochemical changes in lactovegetarian urine using H-1 NMR spectroscopy and pattern recognition. Anal. Bioanal. Chem. 396 (2010) 1451-1463.

[13] J. P. Capitanio, R. C. Kyes, L. A. & Fairbanks, Considerations in the selection and conditioning of old world monkeys for laboratory research: Animals from domestic sources. ILAR J. 47 (2006) 294-306.

[14] M.E. Bollard, E.G. Stanley, J.C. Lindon, J.K. Nicholson, and E. Holmes, NMR-based metabonomic approaches for evaluating physiological influences on biofluid composition. NMR Biomed. 18 (2005) 143-162.

[15] J. Li, Z. Zhao, J. Feng, J. Gao, and Z. Chen, Understanding the metabolic fate and assessing the biosafety of MnO nanoparticles by metabonomic analysis. Nanotechnology 24 (2013).

[16] F. Dieterle, A. Ross, G. Schlotterbeck, and H. Senn, Metabolite projection analysis for fast identification of metabolites in metabonomics. Application in an amiodarone study. Anal. Biochem. 78 (2006) 3551-3561.

[17] J.M. Fonville, A.D. Maher, M. Coen, E. Holmes, J.C. Lindon, and J.K. Nicholson, Evaluation of Full-Resolution J-Resolved H-1 NMR projections of biofluids for metabonomics information retrieval and biomarker identification. Anal. Chem. 82 (2010) 1811-1821.

[18] U. Demsar, P. Harris, C. Brunsdon, A.S. Fotheringham, and S. McLoone, Principal component analysis on spatial data: an overview. Ann. Assoc. Am. Geogr. 103 (2013) 106-128.

[19] K. Chen, T. Baxter, W.M. Muir, M.A. Groenen, and L.B. Schook, Genetic resources, genome mapping and evolutionary genomics of the pig (Sus scrofa). Int. J. Biol. Sci. 3 (2007) 153-165.

[20] S. Hu, H. Wang, A.A. Knisely, S. Reddy, D. Kovacevic, Z. Liu, and S.M.G. Hoffman, Evolution of the CYP2ABFGST gene cluster in rat, and a fine-scale comparison among rodent and primate species. Genetica 133 (2008) 215-226.

[21] E.B. Alsop, E.S. Boyd, and J. Raymond, Merging metagenomics and geochemistry reveals environmental controls on biological diversity and evolution. BMC Ecology. 14 (2014).

[22] S. Mitri, and K.R. Foster, The genotypic view of social interactions in microbial communities. Annu. Rev. Genet. 47 (2013) 247-273.

[23] R.D. Firn, and C.G. Jones, The evolution of secondary metabolism - a unifying model. Mol. Microbiol. 37 (2000) 989-994.

[24] L.C. Vining, Roles of secondary metabolites from microbes. Ciba Found. Symp. 171 (1992) 184-198.

[25] S.H. Zeisel, Metabolic crosstalk between choline/1-carbon metabolism and energy homeostasis. Clin. Chem. Lab. Med. 51 (2013) 467-475.

[26] U. Doshi, and A.P. Li, A comparison of aldehyde oxidase activity in cryopreserved hepatocytes from multiple animal species: CD-1 mouse, SD rat, Wistar rat, Beagle dog, Cynomolgus monkey, Rhesus monkey and man. Drug Metab. Rev. 45 (2014) 200-200.

[27] X. Zhang, Z. Zhu, Z. Huang, P. Tan, and R.Z. Ma, Microsatellite genotyping for four expected inbred mouse strains from KM mice. J. Genet. Genomics. 34 (2007) 214-222.

[28] R.A. Gibbs, G.M. Weinstock, M.L. Metzker, D.M. Muzny, E.J. Sodergren, S. Scherer, G. Scott, D. Steffen, K.C. Worley, P.E. Burch, G. Okwuonu, S. Hines, L. Lewis, C. DeRamo, O. Delgado, S. Dugan-Rocha, G. Miner, M. Morgan, A. Hawes, R. Gill, R.A. Holt, M.D. Adams, P.G. Amanatides, H. Baden-Tillson, M. Barnstead, S. Chin, C.A. Evans, S. Ferriera, C. Fosler, A. Glodek, Z.P. Gu, D. Jennings, C.L. Kraft, T. Nguyen, C.M. Pfannkoch, C. Sitter, G.G. Sutton, J.C. Venter, T. Woodage, D. Smith, H.M. Lee, E. Gustafson, P. Cahill, A. Kana, L. Doucette-Stamm, K. Weinstock, K. Fechtel, R.B. Weiss, D.M. Dunn, E.D. Green, R.W. Blakesley, G.G. Bouffard, J. de Jong, K. Osoegawa, B.L. Zhu, M. Marra, J. Schein, I. Bosdet, C. Fjell, S. Jones, M. Krzywinski, C. Mathewson, A. Siddiqui, N. Wye, J. McPherson, S.Y. Zhao, C.M. Fraser, J. Shetty, S. Shatsman, K. Geer, Y.X. Chen, S. Abramzon, W.C. Nierman, R.A. Gibbs, G.M. Weinstock, P.H. Havlak, R. Chen, K.J. Durbin, A. Egan, Y.R. Ren, X.Z. Song, B.S. Li, Y. Liu, X. Qin, S. Cawley, G.M. Weinstock, K.C. Worley, A.J. Cooney, R.A. Gibbs, L.M.

D'Souza, K. Martin, J.Q. Wu, M.L. Gonzalez-Garay, A.R. Jackson, K.J. Kalafus, M.P. McLeod, A. Milosavljevic, D. Virk, A. Volkov, D.A. Wheeler, et al., Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature 428 (2004) 493-521.

[29] V.M. Katoch, K. Katoch, V.P. Bharadwaj, A.K. Datta, V.D. Sharma, C.T. Shivannavar, and K.B. Kannan, Metabolic studies on mycobacteria. V. A preliminary report on the ATP synthesis by mycobacteria including M. leprae by using different substrates. Indian J. Leprosy. 59 (1987) 163-70.

[30] S. Morris, and A.M. Adamczewska, Utilisation of glycogen, ATP and arginine phosphate in exercise and recovery in terrestrial red crabs, Gecarcoidea natalis. Comp. Biochem. Physiol. A Mol. Integr. Physiol. 133 (2002) 813-825.

[31] T.S. Kang, D.R. Korber, and T. Tanaka, Contributions of citrate in redox potential maintenance and ATP production: metabolic pathways and their regulation in Lactobacillus panis PM1. Appl. Microbiol. Biot. 97 (2013) 8693-8703.

[32] M. Zeng, Y. Liang, H. Li, B. Wang, and X. Chen, A metabolic profiling strategy for biomarker screening by GC-MS combined with multivariate resolution method and Monte Carlo PLS-DA. Anal. Methods-UK. 3 (2011) 438-445.

Figure Legends

Fig.1. 600 MHz ^{1}H NMR spectra of serum **(A)** and urine **(B)** obtained from different laboratory animals and human. The aliphatic region (in the dashed box) was respectively magnified 20 (for serum) or 8 times (for urine) compared with corresponding aromatic region for the purpose of clarity. Keys: 1-MH: 1-Methylhistidine; 2-HB: 2-Hydroxybutyrate; 3-HB: 3-Hydroxybutyrate; AA: Acetoacetate; AC: Acetylcholine; Ace: Acetate; Aco: cis-Aconitate; Act: Acetone; AD: Acetamide; AH: Aminohippurate; Ala: Alanine; All: Allantoin; Alt: Allantoate; BA: Bile acid; Ben: Benzoate; Bet: Betaine; Bu: Butyrate; Cit: Citrate; Cn: Creatinine; Cr: Creatine; DG: Deoxyguanosine; DHM: 3,4-Dihydroxymandelate; DHT: Dihydrotyymine; DMA: Dimethylamine; DMG: N, N-Dimethylglycine; EA: Ethanolamine; Eth: Ethanol; For: Formate; G: Glycerol; GA: Guanidoacetate; GABA: gama-Aminobutyrate; Gln: Glutamine; Glu: Glutamate; Gly: Glycine; GPC: Glycerolphosphocholine; HG: Homogentisate; HIB: 2-Hydroxyisobutyrate; Hip: Hippurate; His: Histidine; IB: Isobutyrate; Ile: Isoleucine; IP: Isopropanol; IV: Isovalerate; Kg: α-Ketoglutarate; KV: 2-Ketoisovalerate; L1: LDL, C$\underline{H}_3$-$(CH_2)_n$-; L2: VLDL, C$\underline{H}_3$-$(CH_2)_n$-; L3: LDL, CH_3-$(C\underline{H}_2)_n$-; L4: VLDL, CH_3-$(C\underline{H}_2)_n$-; L5: Lipid, -C$\underline{H}_2$-CH_2-C=O; L6: Lipid, -C$\underline{H}_2$-CH=CH-; L7: Lipid, -CH_2-C=O; L8: Lipid, =CH-C$\underline{H}_2$-CH=; L9: Lipid, -CH=CH-; Lac: Lactate; Leu: Leucine; Lys: Lysine; M: Malonate; *m*-HPA: *meta*-Hydroxyphenylacetate; *m*-I: *myo*-Inositol; MA: Methylamine; Met: Methoinine; MG: Methylguanidine; MM: Methylmalonate; Mth: Methanol; NAG: N-acetylglutamate; NAS: N-acetyl glycoprotein signals; NMN: N-Methylnicotinamide; *o*-HPA: *ortho*-Hydroxyphenylacetate; OA: 3-Oxoadipate; OAc: Oxaloacetae; OAG: O-acetyl glycoprotein signals; *p*-HPA: *para*-Hydroxyphenylacetate; PAG: Phenylacetylglycine; PC: Phosphocholine: Phe: Phenylalanine; Py: Pyruvate; SB: Sebacate; Suc: Succinate; Tau: Taurine; Thr: Threonine; TMA: Trimethylamine; Tri: Trigonelline; Trp: Troptophan; Tyr: Tyrosine; U: Unknown; Uc: Urocanate; Val: Valine; Xan: Xanthine; α-Glc: α-Glucose; β-Glc: β-Glucose

Fig.2. 3D score plots to compare the metabolic profile of human and different-specie/strain laboratory animals based on the NMR data of serum **(A)** and urine **(B)** metabonomes. Abbreviation: D: dairy cow, H: human, M: mouse, P: piggy, R: rat, G: guinea pig, M-B: Balb/c mouse, M-K: KM mouse, Ra: rabbit, R-S: SD rat, R-W: Wistar rat

Fig.3. OPLS-DA score plots (left panel) and the corresponding coefficient loading plots (middle and right panels) derived from NMR data of serum **(A)** and urine **(B)** obtained from different pair-wise groups. H: human, M: mouse, R: rat, P: piggy, G: guinea pig, M-B: Balb/c mouse, M-K: KM mouse, Ra: rabbit, R-S: SD rat, R-W: Wistar rat. Peaks in the positive direction indicate metabolites that are more abundant in the group in the positive direction of first principal component, and vice versa. Keys for the assignments are shown in the Fig. 1

Fig.4. The comparison of metabolic variations among different pairwise species. **(A)** The comparison of correlation coefficient of each metabolite in serum between human-mouse and human-rat; **(B)** Overall coefficient of the metabolites in serum from different pairwise species; **(C)** Overall coefficient of the metabolites in urine from different pairwise species
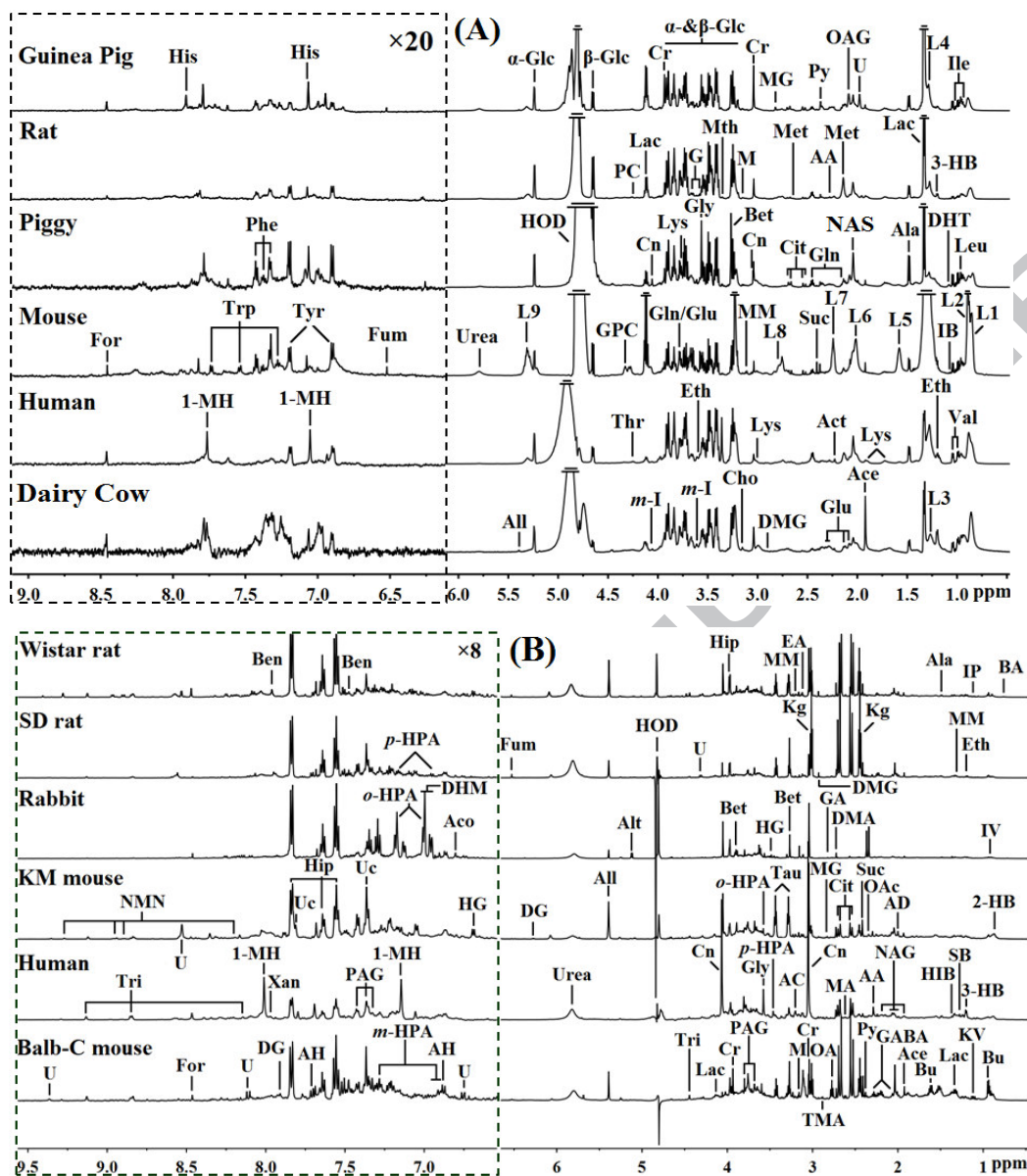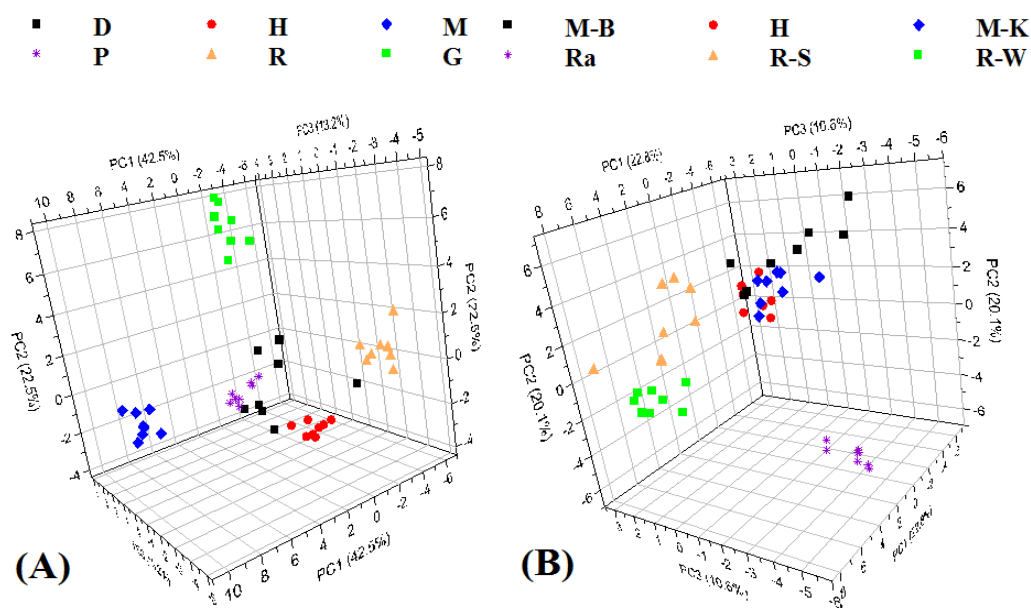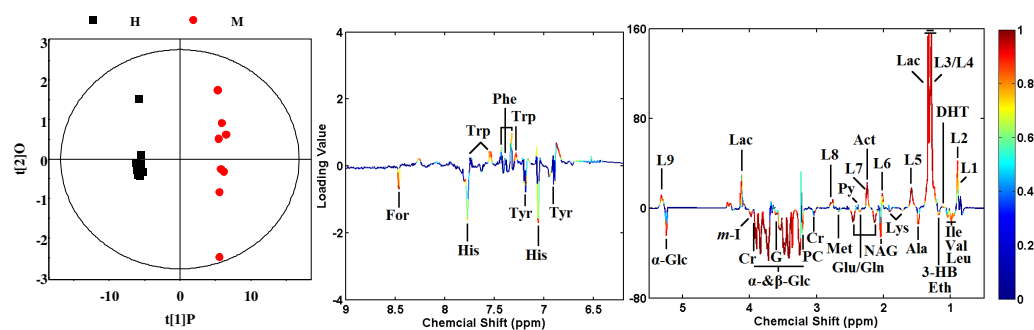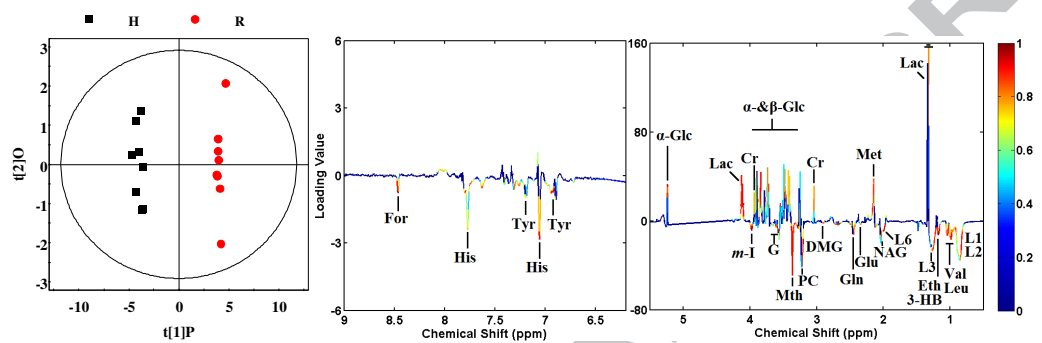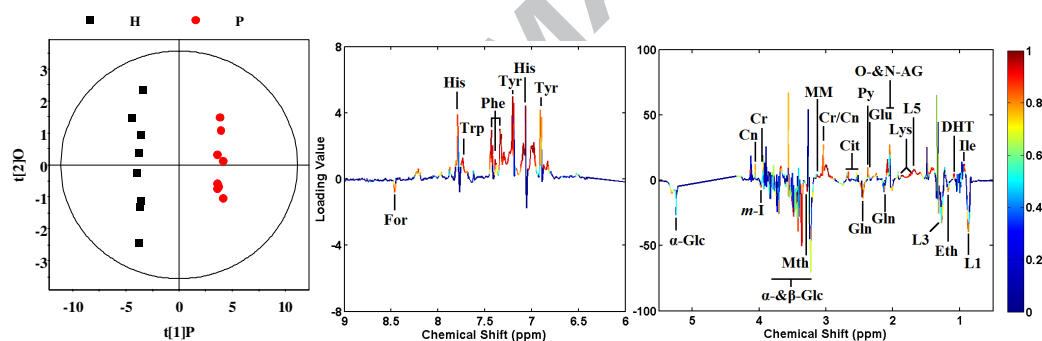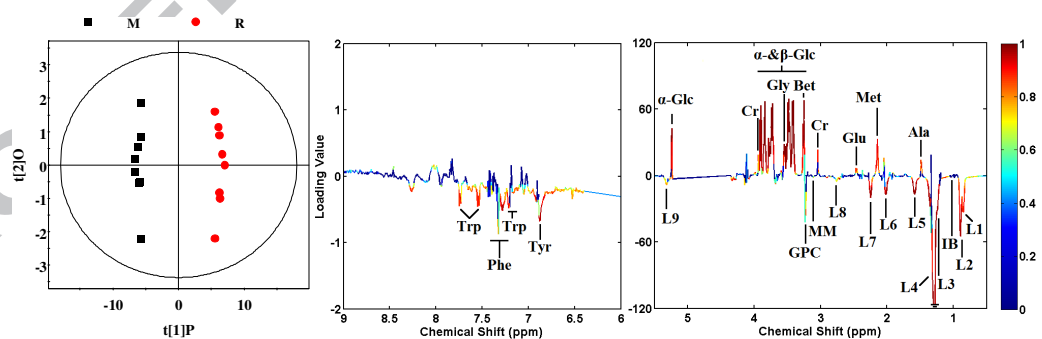
Fig.1.

Fig.2.
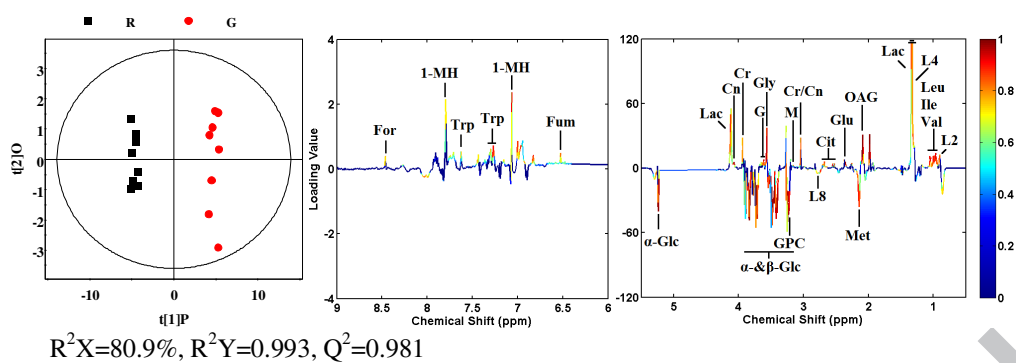
25

$R^2X=87.7\%$, $R^2Y=0.996$, $Q^2=0.988$
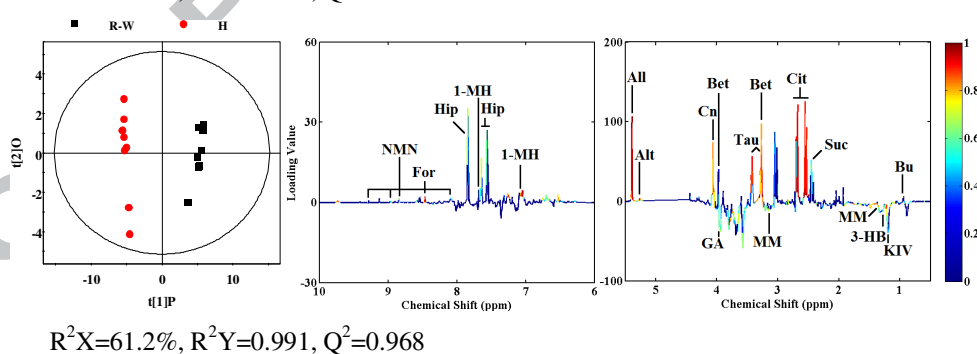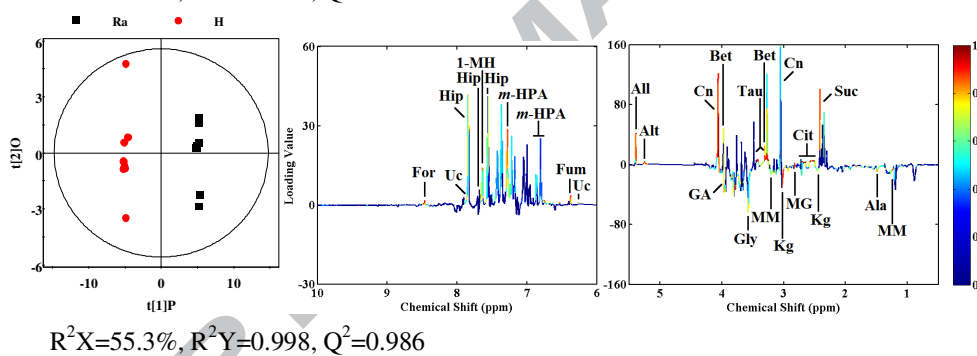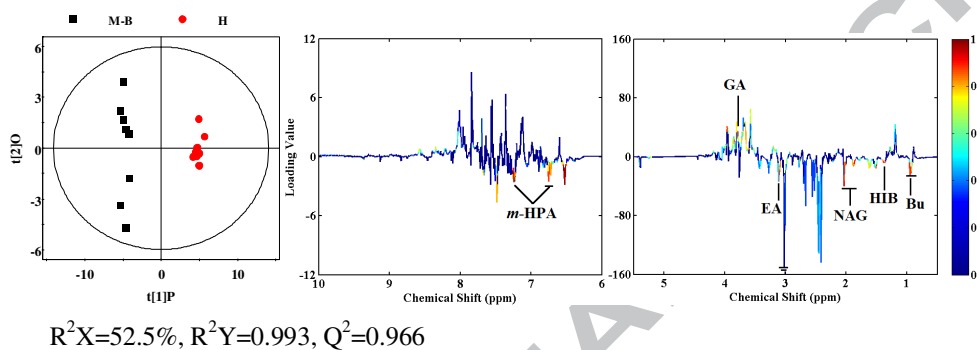
$R^2X=72.2\%$, $R^2Y=0.994$, $Q^2=0.971$

$R^2X=71.7\%$, $R^2Y=0.996$, $Q^2=0.987$

$R^2X=90.0\%$, $R^2Y=0.995$, $Q^2=0.989$

26

R^{2}X=80.9%, R^{2}Y=0.993, Q^2=0.981

(A)

R^{2}X=52.5%, R^{2}Y=0.993, Q^2=0.966

R^{2}X=55.3%, R^{2}Y=0.998, Q^2=0.986

R^{2}X=61.2%, R^{2}Y=0.991, Q^2=0.968

27

R²X=41.9%, R²Y=0.986, Q²=0.917

R²X=56.2%, R²Y=0.993, Q²=0.967

R²X=56.7%, R²Y=0.972, Q²=0.920

(B)

Fig.3.

A The comparsion of metabolic variations in serum between human-mouse and human-rat

B The metabolic variations in serum between different pairwise species

C The metabolic variations in urine between different pairwise species

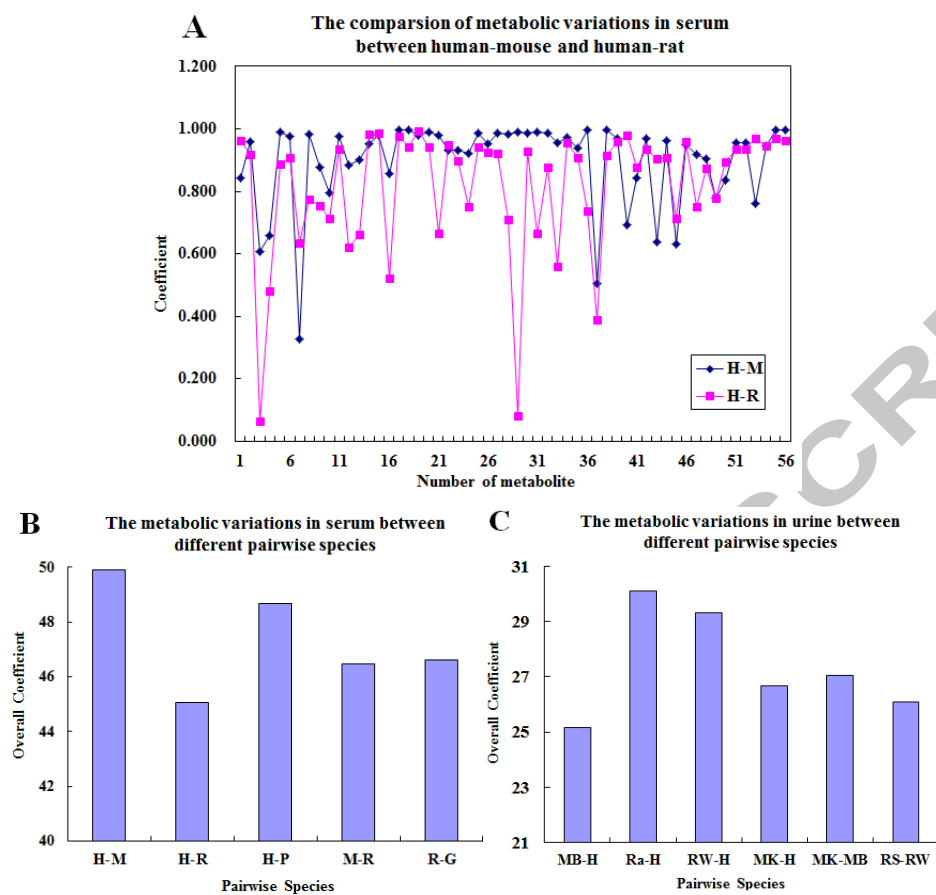**Fig.4.**

Table 1. The involved pathways corresponding to metabolic differences in serum (A) and urine (B) between pair-wise species

(A) Serum

Pair-wise species	Metabolic pathway
Human-Mouse	Microbial metabolism in diverse environments (9/40[a])
	Biosynthesis of biological substance including secondary metabolites (7/40), amino acid (6/40) and aminoacyl-tRNA (5/40)
	ABC transporters (7/40)
	Protein digestion and absorption (6/40)
	Carbon metabolism (5/40)
	Other specific amino acid metabolisms (4/40) and degradation (4/40)
Human-Rat	Microbial metabolism in diverse environments (7/30)
	Biosynthesis of biological substance including secondary metabolites (6/30), amino acids (5/30) and aminoacyl-tRNA (5/30)
	ABC transporters (6/30)
	Protein digestion and absorption (5/30)
	Carbon metabolism (4/30)
	Other specific amino acid metabolisms (6/30)
Human-Piggy	Microbial metabolism in diverse environments (10/35)
	Biosynthesis of biological substance including secondary metabolites (8/35), amino acids (7/35) and aminoacyl-tRNA (5/35)
	ABC transporters (5/35)
	Protein digestion and absorption (5/35)
	Carbon metabolism (7/35)
	Other specific carboxylic acids metabolisms (5/35)
Mouse-Rat	Microbial metabolism in diverse environments (6/32)
	Biosynthesis of biological substance including secondary metabolites (5/32), amino acids (5/32) and aminoacyl-tRNA (4/32)
	ABC transporters (4/32)
	Protein digestion and absorption (5/32)
	Other specific amino acid metabolisms (4/32)
Rat-Guinea pig	Microbial metabolism in diverse environments (8/27)
	Biosynthesis of biological substance including secondary metabolites (6/27), amino acids (4/27) and alkaloids (4/27)
	Carbon metabolism (6/27) and carbon fixation (5/27)
	Other specific amino acid (9/27) and carboxylic acids (5/27) metabolisms

(B) Urine

Pair-wise species	Metabolic pathway
Balb/c mouse-Human	Microbial metabolism in diverse environments (2/5)
	Other specific amino acids metabolism (4/5)
Rabbit-Human	Microbial metabolism in diverse environments (9/20)
	Biosynthesis of secondary metabolites (4/20)
	ABC transporters (4/20)
	Carbon metabolism (6/20) and carbon fixation (4/20)

30

	Other specific amino acids (9/20) and carboxylic acids (4/20) metabolisms and degradation (4/20)
Wistar rat-Human	Microbial metabolism in diverse environments (4/17)
	Biosynthesis of secondary metabolites (2/17) and alkaloids (8/17)
	Carbon metabolism (3/17) and carbon fixation (3/17)
	Other specific amino acids (2/17) and carboxylic acids (3/17) metabolisms
KM mouse-Human	Microbial metabolism in diverse environments (4/9)
	Protein digestion and absorption (1/9)
	Carbon fixation (1/17)
	Other specific amino acids metabolism (6/9) and degradation (3/9)
KM mouse-Balb/c mouse	Microbial metabolism in diverse environments (6/15)
	Biosynthesis of secondary metabolites (3/15) and alkaloids (5/15)
	Carbon fixation (2/15)
	Other specific amino acids metabolism (13/15) and degradation (6/15)
SD rat-Wistar rat	ABC transporters (2/7)
	Other specific amino acids metabolism (2/7) and degradation (1/7)

[a] The ratio of metabolites involved in any a metabolic pathway to all the significant metabolites corresponding to the species difference.