Metabolomics of biofluids : from analytical tools to data interpretation

Nevedomskaya, E.

General Introduction

Variability of genomes in populations is a necessary prerequisite for evolution. Owing to it, populations adapt to the changing environment, conquer new territories and new species evolve. However, this genetic variability can only be seen at the population level; at the level of the individual its genome is constant and static. The genome defines the possibilities of a given organism to adapt to the environment, but does not reflect its actual state. As in modern medicine there is a conceptual shift to personalized health, the attention is being redirected from populations to individuals. Though the importance of genetics is hard to overestimate, new ways of assessing human individuality, or, in other words, the phenotype, are sought. The biochemical representation of the phenotype is believed to be most closely approximated by the metabolome – a collection of low-molecular-weight (<1 kDa) compounds (metabolites) present in an organism.(1) Metabolites are the products of all the biochemical processes in the organism, which makes them a more appropriate target for phenotype-based research than transcripts and proteins, which are information messengers and executors of the biochemical reactions, respectively.

The key words that characterize metabolomics are diversity and variation. These characteristics can be both virtue and vice for the metabolomics workflow (Figure 1). At different steps variation has either to be explored and used or reduced to the possible minimum.

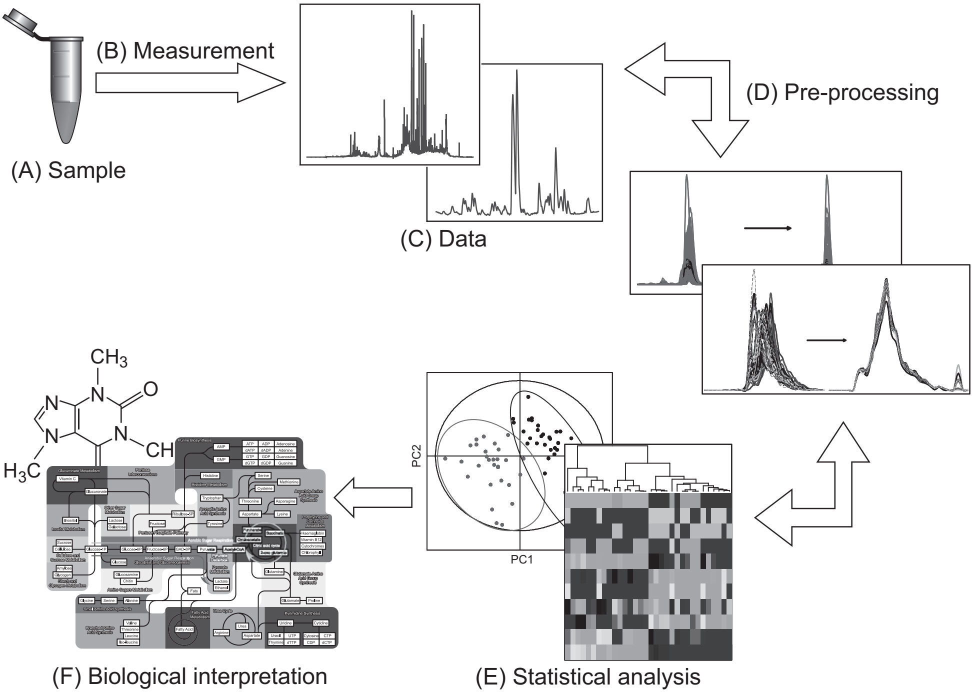

Figure 1. Typical metabolomics workflow.

The variability that occurs in metabolomics, as well as in other ‘omics’-experiments, has three sources – biological, pre-analytical and analytical.(2) Unlike in animal experiments, in which it is relatively easy to standardize both the conditions under which the animals are kept and the handling of the samples, clinical experiments are much more susceptible to bias and interference. This variability affects the outcome of the existing clinical tests(3), and without question influences metabolomics results as well.

Already at the stage of planning the experiments and the study design it is important to reduce the part of biological variation which is not related to the question addressed by the study. When, for example, two groups of patients or patients and healthy controls are to be compared, it is important that the selected groups have minimum differences not related to the research aim. These can be associated with gender, age and diet. These factors have been shown to have a large effect on metabolic profiles.(4;5) It also has been shown that the differences due to the fact that samples are collected in different countries and cities can be easily detected in urine metabolic profiles.(6;7) These extremes should be avoided of course, but it is not difficult to imagine that in a typical multi-center clinical study samples come from different hospitals, than the results obtained from the data should be considered with great caution.

When samples are collected (Figure 1(A)) it is very important that the collection is highly standardized. First of all, the time of collection is an important factor to consider. Diurnal variation is not that obvious in case of blood plasma, but has a strong impact on metabolic profiles of urine.(8) The subjects should also either follow a certain diet or fast before the sample collection. Sample handling obviously also plays a role and such factors as the collection tubes, time on ice before freezing, temperature and time of storage, the number of thaw-freeze cycles can introduce a considerable bias.(9)

Assuming that the study had been properly designed and that the samples had been collected, the aim of a metabolomics experiment would be to generate a comprehensive view on the low-molecular-weight components in the samples (Figure 1(B)). At this stage another issue of variability in metabolomics is faced: which is the diversity of metabolites themselves. And this represents a major difference between metabolomics and the other ‘omics’ technologies. In genomics, transcriptomics and proteomics the molecules measured belong to the same chemical classes; metabolites, however, are extremely diverse in their chemical and physical properties. It is impossible to cover all of the metabolites by a single analytical technique, which is the reason why a number of analytical platforms are being used in the field.(10) Those platforms are either Mass-Spectrometry (MS) or Nuclear Magnetic Resonance (NMR) based. Both approaches have their specific advantages and disadvantages.(11) For example, NMR has a lower concentration sensitivity than MS, but requires less sample pretreatment and is non-destructive. MS-based technologies are very sensitive, but the range of the compounds detected is defined by the separation method used, thus making it more targeted compared to ^{1}H NMR, which is universal for all molecules containing hydrogen, which in case of biological samples means very few exceptions. The combination of analytical methods can increase the coverage of the metabolome and thus lead to better understanding of the biological process under study as well as improve identification of the compounds.(12) However, due to many factors, such as, for instance, the high costs of the different instruments, the whole range of the machines suitable for metabolomics experiments is rarely present in one lab. Though the analytical capabilities and drawbacks of each of the platforms are known, it is less known which impact they have on the recovery of biological information. This knowledge is essential for making a decision about the suitability of a given analytical method for a certain biological/clinical application.

The reproducibility of analytical methods is not perfect and this obviously can result in variability in the data. MS-based methods are more prone to this variability compared to NMR, not only due to the nature of the instrumentation, but also because the samples need more extensive pre-treatment and preparation. The reproducibility of every technique should be carefully assessed in order to understand the possible drawbacks and the extent of post-processing needed. Also each of the methods has to be optimized to give minimum variability. A way to control the instrument performance while running long sequences of measurements is to use a set of analytical standards and quality control (QC) samples.(13)

Despite all the efforts to get rid of the unwanted analytical variability in the final data, this can only be minimized. A common loss in repeatability is the drift in peak position in the spectra. In NMR this is related to the difference in pH and ion strength between the samples which is not completely abolished by the addition of buffer into the samples before the measurement. In chromatography-MS techniques misalignment of chromatographic peaks is caused by the sample matrix, pH, column ageing; the stability of the MS can also be compromised and cause run-to-run differences.(14) Peak shift is also considered a serious drawback of capillary electrophoresis.(15) A trivial method that allows overcoming the drift problem to a certain extent is binning – dividing the available data axis in short segments (bins) and integrating the signal intensity of each of them. More sophisticated extensions of the method are also available, such as, for example, adaptive binning (16) and Gaussian binning.(17) Binning is often used in case of NMR, although there are reasons to consider this method not optimal.(14) For chromatography-MS techniques it is even less advantageous due to the more complex nature of the data, the large number of variables generated and the loss of peak information. For this reason, in chromatography-MS a combination of alignment and peak picking is most regularly used with a variety of available algorithms and software.(18-20)

After all the manipulations mentioned above the only variability that should be left in a metabolomics dataset is of biological origin. The unwanted part of it is related to the different dilution of the samples, which is most prominent in the case of urine due to different water uptake by the subjects. Normalization is the step that removes this variation. For this, a few methods are available: normalization to the total sum, to some "housekeeping" molecules (*e.g.* to creatinine) and more sophisticated ones (*e.g.* Probabilistic Quotient Normalization(21)), but none of them are optimal for all cases and the choice should be made depending on the biological context.(22)

The rest of the variation in the data has to be explored as this represents the biological variability (Figure 1(E)). The first step of the analysis is to investigate data structure in order to find patterns, natural grouping of the samples and possible outliers. This is done by means of unsupervised methods, which do not use any *a priori* information about the data and thus give an unbiased view. The most often used procedures are Principal Component Analysis (PCA) and its variations and clustering.

PCA is a projection-based method that summarizes the variation present in the data in a lower-dimension space. When applied to data containing both biological and QC samples it is possible to estimate the analytical variation in the data in comparison to the biological variability: QCs must have orders of magnitude less variation than the real samples, thus clustering tightly together. PCA is also extremely useful for identifying abnormal samples (outliers), which may have to be removed prior to any further analysis, and for detecting any grouping of samples. The latter might be related to the question of interest, addressed by the study, or might be related to other phenomena. The first case is very encouraging for continuation of the analysis. The second does not mean that subsequent analysis cannot be done, because often the studied differences are subtle and masked by other sources of biological diversity; it however implies a more careful selection of strategies for discrimination and especially for validation, as natural clustering of the samples might influence the results of cross-validation.

Clustering is a collection of unsupervised methods to assess inter-sample relations and identify natural groups of samples. There are many variants of clustering that use different measures for the distance between the samples (Euclidean, Mahalanobis, Manhattan *etc.*), different algorithms (hierarchical, partitioning *etc.*) and various initial assumptions (*e.g.* whether samples belong to only one or to multiple clusters). Clustering not only allows discovering grouping of samples, but also assessing the quality of the data.

As has already been mentioned above, in unsupervised methods the variation of interest is not necessarily reflected in the natural grouping of samples. The questions often posed in clinical metabolomics research are finding the differences between two and more groups of samples (from patients and controls, from subjects under various conditions and/or interventions) and predicting to which of the studied groups new samples belong. Statistically speaking the tasks are discrimination and classification, which are often carried out together in one method.

The abundance of discrimination/classification methods may appear confusing: projection-based methods (Partial Least Square Discriminant analysis (PLS-DA) and Orthogonal PLS-DA (OPLS-DA), Soft Independent Modelling of Class Analogies (SIMCA)), k-nearest neighbor algorithm (k-NN), artificial neural networks (ANN), support vector machine and others. The choice of a particular method for a certain application might to a large extent be based on the expertise of the user; however there are some factors that should be taken into consideration when selecting the procedure to be applied. One of them is the assumption about the distribution of the data: if there is the information available parametric methods can be used (for example, projection-based methods); without such information non-parametric methods are the preferred choice (k-NN, ANN). SIMCA, ANN, k-NN cope better with a large number of classes than discrimination methods. Discrimination methods show the best performance when the classes are tight, homogeneous in terms of dispersion and covariance structure; otherwise, classes should be modeled separately by means of SIMCA, for example. If samples belong to a number of definite classes, discrimination techniques are used; if not, class modeling techniques should be chosen.(23)

As mentioned above, most often in metabolomics and in particular in clinical metabolomics, the aim is to find the differences in profiles of two or more groups of samples. These differences might not be uncovered in a simple analysis. Thus larger groups have to be investigated in search of systematic variation and more sophisticated statistical methods are used. The latter can result in substantial overfitting and are more difficult to validate and interpret.

Modeling groups against each other averages the effects between the samples from one group and reduces the individual-specific variability. It already has been recognized by the clinical and especially the pharmaceutical community that averaging health and medication intervention effects between people is a dead end street for the development of future medicine and that more personalized approaches have to be found.(24)

The pharmaceutical industry was the first to respond to this need due to certain stagnation in the field, a decrease in development of new drug and the withdrawals of drugs due to unforeseen side effects. With the genomics boom after the completion of the Human Genome project, the answers for personalized medical care and treatment were sought with the help of genomics and resulted in the emergence of a new discipline – pharmacogenomics. Despite all the expectations of this new research field, the number of genomics-driven drug discoveries is low.(25) A possible explanation is that, despite the importance of genes for defining the phenotype, they are not the only factors responsible for determining the actual physiological and/or pathological state of the organism, the response to treatment and the clinical outcome. The metabolome, on the other hand, is much closer to the phenotype in comparison to other 'omes'. The switch from genetic to metabolic "individuality" thus is logical and the need for such a switch is already recognized by the community.(26)

The idea of the connection of metabolism and human individuality and integrity is not new. It was first proposed and documented in the classical work of Sir Archibald E. Garrod "The incidence of alkaptonuria: a study in chemical individuality", the title of which speaks for itself. "...No two individuals of a species are absolutely identical in bodily structure," Garrod wrote as long ago as in 1902, "neither are their chemical processes carried out on exactly the same lines" .(27) The next progress in the field was made almost 50 years later by Roger J. Williams who demonstrated "evidence indicating that each individual possesses what may be called a "metabolic personality"-that is, a distinctive pattern of metabolic traits" and that these traits are maintained over a period of several months.(28)

As many other fundamental ideas, the idea of "metabolic personality", although maybe not enough appreciated at the time it appeared, came back in the 21st century. In the recent publications of Assfalg *et al.* and Bernini *et al.* the two collaborating groups elaborated and experimentally supported exactly the same basic thoughts – that "metabolic phenotypes" (the name changed slightly 60 years after R.J. Williams) do indeed exist and that they are stable over time.(29;30) As the analytical technologies have advanced enormously in the last decades, the analytical basis of the latest research is different from that of R.J. Williams – the use of NMR makes it possible to measure a large number of molecules in one run and to obtain precise quantitative information on these molecules. However, as shown by Assfalg *et al.*, the full variety of metabolites assessed by NMR is not necessary to define the individual metabolic patterns – equally good results can be obtained using only a limited set of 12 compounds. The latter are even to a certain extent overlapping with those measured by Williams. The importance of metabolic phenotypes in relation to disease, nutrition and response to various stimuli has been outlined in both studies.

All the environment and nutrition influences in, for instance, urine are superimposed on the invariant profiles represented by metabolic phenotypes. The existence and stability of such distinct profiles are related to homeostasis. Homeostasis is defined as “a state of dynamic balance with the variables fluctuating between tolerance limits”. The moment when “tolerance limits” are crossed and homeostasis lost can be considered as the moment of the onset of disease.(31) Detecting this occasion might enable early diagnosis, prognosis and possibilities for more successful intervention.

Both health and disease are dynamic entities, understanding of which would be possible only by monitoring them in time. As has been mentioned above, the key feature of the genome is that it is static. The metabolome, on the contrary, is highly dynamic, reflecting the changes happening over time and the reaction of the organism to the altering environment. As such, time-correlated changes of the metabolome will likely have more diagnostic and prognostic power than a single time point measurement.

There is an increasing awareness that metabolomics is of great importance for the medicine of the future due to its “personalized” and dynamic nature. However, to make this possible, the present strategies in metabolomics experiment design and data treatment should be changed. Currently metabolomics literature is dominated by “case-control” studies, which both average the effects between individuals and neglect the beneficial dynamic essence of metabolic profiles. Clinical metabolomics is not an exception, though the concept of “dynamic disease” has been around for a considerable time.(32) The advantages and the gain in information recovery obtained by dynamic profiling are starting to be recognized, but are not universally applied. In order to generalize its use it is important to change study design, sample collection and data analysis strategies.

Even when dynamic metabolomics data is being collected, the advantages of it are not always exploited as it is sometimes analyzed with the use of “classical” statistical methods, such as PCA and PLS-DA.(33) However, it has been clearly shown that those methods are not optimal for such data and optimized or new strategies should be applied.(34)

Metabolomics data generated by NMR and MS-based technologies is multivariate by its nature due to the large number of molecular species measured in one run. With the addition of the time dimension the data becomes also multilevel as different levels of variation, for instance, between- and within-individual, can be assessed. A collection of powerful methods for dealing with such data and for separating the levels of variability is thus called multilevel.(35;36) Exploring the between-individual variation allows neglecting the intra-individual changes, which may be non-systematic day-to-day differences that do not relate to the question of interest. On the other hand, the within-individual block comprises the time-related information and using this is a more personalized approach for data analysis as each person acts as its own control.

To sum up, metabolomics is an attractive methodology for clinical research as it is the most accurate reflection of the actual physiological and biochemical state of the organism. The dynamic and highly "individualized" nature of the metabolome is a strong indication that it could provide the means to make personalized medicine go all the way from an "elusive dream"(37), via "proof-of-principle", to real application. The current thesis does not offer a recipe how to do it, but describes a number of essential components for the development of this new type of medicine, such as robust and reproducible analytical methods, pre-processing routines, various data analysis methods and also metabolomics applications to both animal and clinical experiments that use the longitudinal study design.

REFERENCES

1. Lindon,J.C., Holmes,E., Bollard,M.E., Stanley,E.G., and Nicholson,J.K. 2004. Metabonomics technologies and their applications in physiological monitoring, drug safety assessment and disease diagnosis. Biomarkers 9:1-31.

2. De Noo,M.E., Tollenaar,R.A., Deelder,A.M., and Bouwman,L.H. 2006. Current status and prospects of clinical proteomics studies on detection of colorectal cancer: hopes and fears. World J. Gastroenterol. 12:6594-6601.

3. Lippi,G., Guidi,G.C., Mattiuzzi,C., and Plebani,M. 2006. Preanalytical variability: the dark side of the moon in laboratory testing. Clin. Chem Lab Med. 44:358-365.

4. Holmes,E., Loo,R.L., Stamler,J., Bictash,M., Yap,I.K., Chan,Q., Ebbels,T., de,I.M., Brown,I.J., Veselkov,K.A. *et al* 2008. Human metabolic phenotype diversity and its association with diet and blood pressure. Nature 453:396-400.

5. Slupsky,C.M., Rankin,K.N., Wagner,J., Fu,H., Chang,D., Weljie,A.M., Saude,E.J., Lix,B., Adamko,D.J., Shah,S. *et al* 2007. Investigations of the effects of gender, diurnal variation, and age in human urinary metabolomic profiles. Anal Chem 79:6995-7004.

6. Lenz,E.M., Bright,J., Wilson,I.D., Hughes,A., Morrisson,J., Lindberg,H., and Lockton,A. 2004. Metabonomics, dietary influences and cultural differences: a ^{1}H NMR-based study of urine samples obtained from healthy British and Swedish subjects. J. Pharm. Biomed. Anal 36:841-849.

7. Holmes,E., Loo,R.L., Stamler,J., Bictash,M., Yap,I.K., Chan,Q., Ebbels,T., de,I.M., Brown,I.J., Veselkov,K.A. *et al* 2008. Human metabolic phenotype diversity and its association with diet and blood pressure. Nature 453:396-400.

8. Lenz,E.M., Bright,J., Wilson,I.D., Morgan,S.R., and Nash,A.F. 2003. A ^{1}H NMR-based metabonomic study of urine and plasma samples obtained from healthy human subjects. J. Pharm. Biomed. Anal 33:1103-1115.

9. Lindon,J.C., Nicholson,J.K., Holmes,E., Keun,H.C., Craig,A., Pearce,J.T., Bruce,S.J., Hardy,N., Sansone,S.A., Antti,H. *et al* 2005. Summary recommendations for standardization and reporting of metabolic analyses. Nat. Biotechnol. 23:833-838.

10. Lindon,J.C., and Nicholson,J.K. 2008. Analytical technologies for metabonomics and metabolomics, and multi-omic information recovery. Trac-Trends in Analytical Chemistry 27:194-204.

11. Lindon,J.C., and Nicholson,J.K. 2008. Spectroscopic and Statistical Techniques for Information Recovery in Metabonomics and Metabolomics. Annual Review of Analytical Chemistry 1:45-69.

12. Crockford,D.J., Holmes,E., Lindon,J.C., Plumb,R.S., Zirah,S., Bruce,S.J., Rainville,P., Stumpf,C.L., and Nicholson,J.K. 2006. Statistical heterospectroscopy, an approach to the integrated analysis of NMR and UPLC-MS data sets: application in metabonomic toxicology studies. Anal Chem 78:363-371.

13. van der Greef,J., Martin,S., Juhasz,P., Adourian,A., Plasterer,T., Verheij,E.R., and McBurney,R.N. 2007. The art and practice of systems biology in medicine: mapping patterns of relationships. J. Proteome Res. 6:1540-1559.

14. Aberg,K.M., Alm,E., and Torgrip,R.J. 2009. The correspondence problem for metabonomics datasets. Anal Bioanal Chem.

15. Garcia-Perez,I., Vallejo,M., Garcia,A., Legido-Quigley,C., and Barbas,C. 2008. Metabolic fingerprinting with capillary electrophoresis. J. Chromatogr. A 1204:130-139.

16. De Meyer,T., Sinnaeve,D., Van,G.B., Tsiporkova,E., Rietzschel,E.R., De Buyzere,M.L., Gillebert,T.C., Bekaert,S., Martins,J.C., and Van,C.W. 2008. NMR-based characterization of metabolic alterations in hypertension using an adaptive, intelligent binning algorithm. Anal Chem 80:3783-3790.

17. Anderson,P.E., Reo,N.V., DelRaso,N.J., Doom,T.E., and Raymer,M.L. 2008. Gaussian binning: a new kernel-based method for processing NMR spectroscopic data for metabolomics. Metabolomics 4:261-272.

18. Katajamaa,M., and Oresic,M. 2007. Data processing for mass spectrometry-based metabolomics. J. Chromatogr. A 1158:318-328.

19. van Nederkassel,A.M., Daszykowski,M., Eilers,P.H., and Heyden,Y.V. 2006. A comparison of three algorithms for chromatograms alignment. J. Chromatogr. A 1118:199-210.

20. Lange,E., Tautenhahn,R., Neumann,S., and Gropl,C. 2008. Critical assessment of alignment procedures for LC-MS proteomics and metabolomics measurements. BMC Bioinformatics 9:375.

21. Dieterle,F., Ross,A., Schlotterbeck,G., and Senn,H. 2006. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in ^{1}H NMR metabonomics. Anal Chem 78:4281-4290.

22. Craig,A., Cloarec,O., Holmes,E., Nicholson,J.K., and Lindon,J.C. 2006. Scaling and normalization effects in NMR spectroscopic metabonomic data sets. Anal Chem 78:2262-2267.

23. Gonzalez,A.G. 2007. Use and misuse of supervised pattern recognition methods for interpreting compositional data. Journal of Chromatography A 1158:215-225.

24. Woodcock,J. 2007. The prospects for "personalized medicine" in drug development and drug therapy. Clinical Pharmacology & Therapeutics 81:164-169.

25. Goldsmith,P., Fenton,H., Morris-Stiff,G., Ahmad,N., Fisher,J., and Prasad,K.R. 2010. Metabonomics: a useful tool for the future surgeon. J. Surg. Res. 160:122-132.

26. Nicholson,J.K., Wilson,I.D., and Lindon,J.C. 2010. Pharmacometabonomics as an effector for personalized medicine. Pharmacogenomics 12:103-111.

27. Garrod,A.E. 1902. THE INCIDENCE OF ALKAPTONURIA : A STUDY IN CHEMICAL INDIVIDUALITY. The Lancet 160:1616-1620.

28. Williams,R.J., Berry,L.J., and Beerstecher,E. 1949. Individual Metabolic Patterns, Alcoholism, Genetotrophic Diseases. Proceedings of the National Academy of Sciences of the United States of America 35:265-271.

29. Assfalg,M., Bertini,I., Colangiuli,D., Luchinat,C., Schafer,H., Schutz,B., and Spraul,M. 2008. Evidence of different metabolic phenotypes in humans. Proceedings of the National Academy of Sciences of the United States of America 105:1420-1424.

30. Bernini,P., Bertini,I., Luchinat,C., Nepi,S., Saccenti,E., Schafer,H., Schutz,B., Spraul,M., and Tenori,L. 2009. Individual Human Phenotypes in Metabolic Space and Time. J. Proteome Res.

31. van der Greef,J., and Smilde,A. 2005. Symbiosis of chemometrics and metabolomics: past, present, and future. JOURNAL OF CHEMOMETRICS 19:376-386.

32. Glass,L., and Mackey,M.C. 1988. From clocks to chaos : the rhythms of life. Princeton University Press. Princeton, NJ.

33. Li,J.A., Wijffels,G., Yu,Y.H., Nielsen,L.K., Niemeyer,D.O., Fisher,A.D., Ferguson,D.M., and Schirra,H.J. 2011. Altered Fatty Acid Metabolism in Long Duration Road Transport: An NMR-based Metabonomics Study in Sheep. Journal of Proteome Research 10:1073-1087.

34. Smilde,A.K., Westerhuis,J.A., Hoefsloot,H.C.J., Bijlsma,S., Rubingh,C.M., Vis,D.J., Jellema,R.H., Pijl,H., Roelfsema,F., and van der Greef,J. 2010. Dynamic metabolomic data analysis: a tutorial review. Metabolomics 6:3-17.

35. Jansen,J.J., Hoefsloot,H.C.J., van der Greef,J., Timmerman,M.E., and Smilde,A.K. 2005. Multilevel component analysis of time-resolved metabolic fingerprinting data. Analytica Chimica Acta 530:173-183.

36. Westerhuis,J.A., van Velzen,E.J.J., Hoefsloot,H.C.J., and Smilde,A.K. 2010. Multivariate paired data analysis: multilevel PLSDA versus OPLSDA. Metabolomics 6:119-128.

37. Lesko,L.J. 2007. Personalized medicine: Elusive dream or imminent reality? Clinical Pharmacology & Therapeutics 81:807-816.