Biomarker Discovery in Life Sciences

Proefschrift
ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus Dr. D.D. Breimer,
hoogleraar in de faculteit der Wiskunde en
Natuurwetenschappen en die der Geneeskunde,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 13 april 2005
te klokke 14.15 uur

door

Robert-Jan Antonius Nicolaas Lamers
geboren te Arnhem
in 1971

PROMOTIECOMMISSIE

Promotor:	Prof.dr. J. van der Greef
Co-promotor:	Dr. J.H.J. van Nesselrooij (TNO)
Referent:	Prof.dr. A.K. Smilde (Universiteit van Amsterdam)
Overige leden:	Prof.dr. G.J. Mulder Prof.dr. M. Danhof Prof.dr. T. Hankemeier

ISBN 90 74538 002

Publication of this thesis was financially supported by TNO Quality of Life, Zeist, The Netherlands.

Contents

1

General introduction

Biomarkers

Biomedical research is continuously challenged to discover and elucidate relationships between health and disease on the one hand, and environmental factors such as nutrition and the use of pharmaceuticals on the other hand. Biomarkers can facilitate this quest [1-3]. A biomarker is defined as a parameter that is objectively measured and evaluated as an indicator of normal biological or pathological processes, or pharmacological responses to a therapeutic intervention [4].

A biological system must keep internal conditions within tolerable limits to continue a healthy functioning (see Figure 1). When a change occurs in a biological system, the system usually responds to reverse the change to maintain equilibrium. This control phenomenon is known as homeostasis.

For instance, the body requires glucose to meet the demand for energy. The amount of glucose that is needed will depend on physical exertion. Glucose is mainly obtained from the diet and when the body is provided with a surplus of glucose, this is converted to glycogen which is the principal storage form of glucose in the body. Under certain conditions, between meals for instance, glycogen can be converted to glucose if needed. The pancreas monitors the level of glucose in the blood and it controls the glucose concentration by releasing the hormone insulin or glucagon. Insulin is released as a result of an increase in the blood glucose level and stimulates the uptake of glucose into cells where it is stored as glycogen. Glucagon is released when the blood glucose level decreases, and thus promotes the conversion of glycogen into glucose. In this way, the glucose level in blood is maintained within tolerable limits. In case of diabetes type 1, the body is not able to produce sufficient insulin to stimulate the uptake of glucose into cells. Without this corrective feedback process the blood glucose level can raise above the tolerable

limits, which will result in biological derailment and finally in symptoms of disease. When the body is out of balance, drugs are in many cases the only remedy to reduce the unbalance but often the situation has reached an irreversible state. Nutrition, however, is often suitable for interventions between, but close to, the tolerable limits of equilibrium.

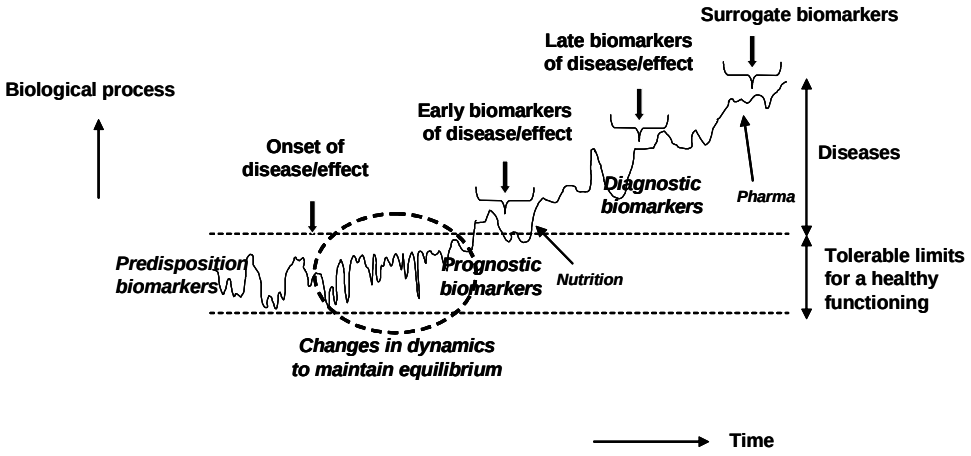

Figure 1. *A biological system must maintain internal conditions within tolerable limits to continue a healthy functioning. This phenomenon is known as homeostasis. Biomarkers are indicators for a biological process that is, or is likely to be, out of equilibrium and may reflect prognosis, diagnosis or progression of such a process.*

Biomarkers may reflect different stages of a biological process that is, or is likely to get, out of equilibrium. There are three types of biomarkers: predisposition, prognostic and diagnostic biomarkers [5]. A *predisposition* biomarker reflects the sensitivity of a subject to a disease, like breast cancer predisposition genes for example. A *prognostic* biomarker has the ability to predict whether a subject will be susceptible to a disorder. For instance, a well-known prognostic biomarker is cholesterol, which is used to identify the risk of a heart disease. A *diagnostic* biomarker measures the incidence and progression of a disease process. Haemoglobin A_{1C} is an example of such a biomarker. When blood glucose levels increase in subjects with diabetes, the levels of haemoglobin A_{1C} in the blood increase accordingly, providing a diagnostic marker of the progression of the

disease. Diagnostic biomarkers can be classified as early, late and surrogate biomarkers and have in common that they are indicators for a disorder that has already developed. Early biomarkers measure the appearance of a disease in an early stage, when curing may still be possible. Late and surrogate biomarkers reflect a state of disease in which relieving symptoms is often the most optimal remedy. A surrogate biomarker is intended to substitute for a clinical endpoint of a disease, a characteristic or variable that reflects how a patient feels, functions or survives.

Research on the identification of biomarkers can greatly benefit from a systems biology approach [6]. Systems biology looks at all the elements of a biological system and reveals their interconnection and interdependence when the system functions in response to biological or environmental perturbations. This will lead to the discovery of new biomarkers, which can be represented by genes, proteins, metabolites, or a combination of these.

Metabolites are intermediate or end products of biological processes [7]. As a result, metabolites play an important role as biomarkers. Metabolites can be endogenous or exogenous, primary or secondary. Endogenous metabolites are produced within a biological system. Exogenous metabolites are obtained from external factors, such as food or drugs that are converted in the system. Primary metabolites have a direct relation to an exogenous or endogenous biological process. Secondary metabolites have an indirect relation to a biological process. Biomedical research targets mainly endogenous primary and secondary metabolites [8]. In intervention studies and toxicological research, endogenous and exogenous as well as primary and secondary metabolites are subject of investigations [9].

Analytical techniques and MVDA for biomarker profile selection

Metabolites provide valuable information on the activity of a biological system, and thus about its state. Biological fluids, such as urine and blood, contain thousands of metabolites that are potential biomarkers. Analytical techniques like proton nuclear magnetic resonance spectroscopy (NMR), gas chromatography-mass spectrometry (GC-MS) and liquid chromatography-mass spectrometry (LC-MS), enable quantification and identification of metabolites in biological fluid samples [10].

However, data obtained with these techniques is very intricate, due to the mixture of numerous metabolites present in body fluids. Biological fluid samples from, for instance, a group of healthy versus a group of diseased subjects that are analysed with NMR, often show variations in metabolite levels between the groups that are too small and too complex to be recognized by visual inspection [11]. To find these differences, multivariate data analysis (MVDA) is needed to explore recurrent patterns in the data, as depicted in Figure 2. It is a powerful tool for the analysis of data sets with a large number of variables. In MVDA, samples are classified according to fine distinctions in the original data, such as NMR spectra, which are caused by small differences between metabolite levels of samples. Metabolites that discriminate between groups of samples and that can be correlated to clinical endpoints, are potential biomarkers.

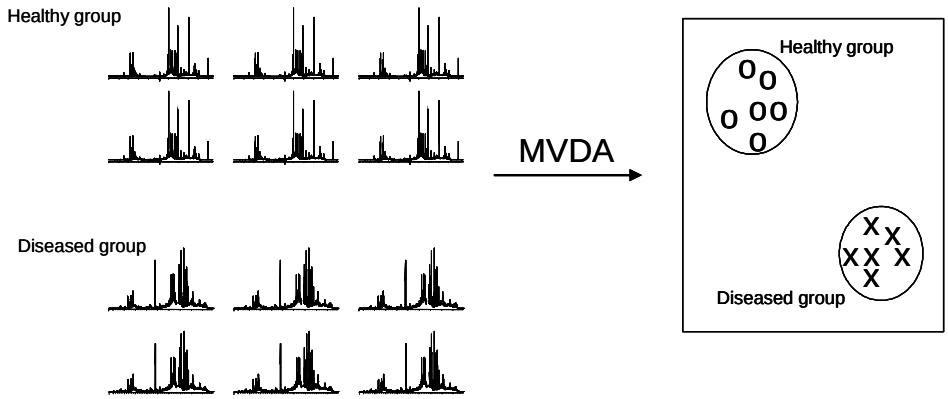

Figure 2. *NMR spectra of biological fluid samples from a group of healthy versus a group of diseased subjects. The data often shows variations in metabolite levels between the groups that are too small and complex to be recognized by eye. To find these differences, multivariate data analysis (MVDA) is needed to explore recurrent patterns in the data.*

The combination of NMR and/or GC-MS, LC-MS with subsequent MVDA is also referred to as metabolic fingerprinting, biomarker profiling, metabolomics, metabonomics or metanomics [7]. The technology has emerged from the biological fluid profiling approaches that were developed many decades ago for the study of

inborn errors of metabolism and effects of nutrition. The early work in this area was mainly driven by GC-MS, which allows low- concentration components to be measured in single profiles. In the eighties, mass spectrometric profiling has become a powerful fingerprinting methodology, especially when combined with MVDA [6].

Both GC-MS and LC-MS are highly suitable for identification of all metabolites in a sample, thus covering the complete metabolome. However, for global biomarker profiling, NMR is an attractive approach, as a wide range of metabolites can be quantified simultaneously without extensive sample preparation. More in-depth studies, using techniques like LC-MS-MS or 2-dimensional NMR, may be used subsequently to identify unknown metabolites and thus to elucidate metabolic pathways involved, especially when metabolite information is integrated with genome, gene expression (transcriptome) and proteome data in systems biology strategies [6].

In this thesis, the term *metabolic fingerprinting* or *biomarker profiling* will be used, because the focus will be on the recognition of global profiles, especially from NMR data, and less on the coverage of the complete metabolome.

Challenges in biomarker profiling

Biomarker profiling has already proven to be a valuable tool in several areas of life science research, such as toxicology and biomedicine. However, there are still frontiers of knowledge. For instance, differences in disease severity or in diet complicate the identification of a biomarker profile. Therefore often well-defined animal studies are used as a starting point. Once a biomarker profile is identified in an animal study, it is supposed that this profile, or part thereof, may also be applicable to the human situation. This issue of translational science is a major challenge for biomarker research.

Other intriguing topics are early and prognostic biomarkers. Diseases as well as environmental factors disturb biorhythms. Such perturbations affect the biological system's metabolism and are supposed to show up in time-course biological data. The metabolites that represent the first perturbation in time are early biomarkers. Early biomarkers may bring prevention of diseases nearer [5].

Most biomarker profiling research is based on in vivo studies. Such studies are usually labour-intensive and expensive. In vitro studies are not hampered by these disadvantages and form therefore an interesting alternative. In vitro experiments could help in generating hypotheses about effects to be expected in vivo. Besides, in vitro research can help in the differentiation between primary and secondary metabolites. This may especially be helpful in the study of complex mixtures, like nutraceuticals and traditional Chinese medicine (TCM) and early biomarkers.
Finally, systems biology should be employed for a better understanding of biomarker profiles and of the underlying biological processes, from genes through proteins to the observed metabolites.

Aim of this thesis

This thesis describes the exploration of metabolic fingerprinting in life sciences and the identification of several novel biomarker profiles, using NMR with subsequent MVDA. In a systems biology approach, metabolic profiling is also used to further understand biological processes.

Outline of this thesis

The individual chapters of this thesis each address a different aspect. In Chapters 2 and 3, the theory behind the analytical techniques NMR and LC-MS and multivariate data analysis is described. In Chapter 4, an application of NMR and MVDA is illustrated with a study in which a diagnostic biomarker profile is searched for osteoarthritis in guinea pigs. Osteoarthritis is further elaborated on in Chapter 5 with the presentation of results of a study in human patients. The combination of NMR and MVDA as a method to find an early biomarker profile is depicted in Chapter 6, with acute rejection after kidney transplantation as an example. In Chapter 7, an in vitro study is presented to show that in vitro research is a useful approach to generate hypotheses about affected metabolic pathways. In Chapter 8, a step is made towards systems biology, by combining metabolite profiles with transcriptomics in a hepatotoxic study with bromobenzene. In Chapter 9, conclusions are drawn and perspectives discussed.

References

1. Colburn, W.C. (2003) Biomarkers in drug discovery and development: from target identification through drug marketing. J. Clin. Pharmacol. 43: 329-341.
2. Frank, R., Hargreaves, R. (2003) Clinical biomarkers in drug discovery and development. Nat. Rev. Drug Disc. 2: 566-580.
3. Schulte, P.A. (1995) Opportunities for the development and use of biomarkers. Toxicol. Lett. 77: 25-29.
4. Biomarkers definitions working group (2001) Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin. Pharmacol. Ther. 69: 89-95.
5. Branca, F., Hanley, A.B., Pool-Zobel, B., Verhagen, H. (2001) Biomarkers in disease and health. Brit. J. Nutr. 85, Suppl.1: S55-S92.
6. Van der Greef J., Davidov E., Verheij E., Vogels J.T.W.E., van der Heijden R., Adourian A.S., Oresic M., Marple E.W., Naylor S. (2002) The Role of Metabolomics in Drug Discovery: A New Vision for Drug Discovery and Development. In: *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis* (Harrigan, G.G. & Goodacre R., eds.). Kluwer Academic Publishing. Boston MA: 170-198.
7. Fiehn O. (2002) Metabolomics: the link between genotypes and phenotypes. Plant Molecular Biology 48: 155-171.
8. Naylor, S. (2003) Biomarkers: current perspectives and future prospects. Expert Rev. Mol. Diagn. 3: 525-529.
9. Timbrell, J.A. (1998) Biomarkers in toxicology. Toxicology 129: 1-12.
10. Nicholson J.K., Lindon J.C., Holmes E. (1999) ''Metabonomics'': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica 29: 1181-1189.
11. Lindon J.C., Nicholson J.K., Holmes E., Everett J.R. (2000) Metabonomics: metabolic processes studied by NMR spectroscopy of biofluids combined with pattern recognition. Concepts Magn. Reson. 12: 289-320.

2

Analytical techniques

Proton NMR spectroscopy

Nuclear magnetic resonance spectroscopy is a method for structure research of compounds which have magnetic nuclei [1]. An example of such a nucleus is the proton, ^{1}H. A proton has a magnetic dipole moment, which tends to align itself with respect to the direction of a sufficiently strong external magnetic field. If a proton is placed in such a field, its axis of rotation (spin) can take only two spatial orientations: aligned with the field or against it. Each of these orientations corresponds to a discrete energy level of the proton. Alignment with the field is the more stable one, and energy must be absorbed to flip the spin from this state to the less stable alignment against the field. Transitions between the two energy levels take place through absorption or emission of electromagnetic radiation:

$$\Delta E = h\nu \qquad (1)$$

with ΔE the energy difference between the two states, h the Planck constant and υ the frequency of the absorbed or emitted radiation.

The energy needed to flip the proton spin depends on the strength of the external magnetic field. The stronger the magnetic field, the greater the tendency of the proton to remain aligned with the field. As a consequence, the required frequency of the radiation will be higher and is given by the relation:

$$\nu = \frac{\gamma_p B}{2\pi} \qquad (2)$$

In this formula is B the strength of the magnetic field and γ_p the gyromagnetic ratio of the proton.

The frequency at which a proton absorbs energy depends on the total magnetic field strength, which is felt by the proton. This so-called effective field strength is not exactly the same as the applied field strength but also includes magnetic contributions from electrons and nuclei around the proton. As a consequence, the absorption frequency of a proton depends on the surroundings in which it is situated, such as the electron density at its location and the presence of neighboring protons. Each set of equivalent protons will have a different environment from other sets, and will thus require a different applied field strength to produce the same effective field strength. At a given frequency, all protons absorb at the same effective field but at a different applied field. This applied field strength is varied and the absorption of radiation, with a maximum for each set of protons, is plotted. The result is an NMR spectrum, an example of which is shown in Figure 1. The distance (δ) of an absorption peak to a reference position is called its chemical shift. Chemical shifts, measured in frequency units, are expressed in relation to a reference compound, which is often sodium 3-trimethylsilyl-propionate-2, 2, 3, 3-$^{2}H_4$ (TMSP). The shift is expressed in parts per million (ppm):

$$\delta = (\nu - \nu_{reference}) x 10^6 / \nu_{reference} \qquad (3)$$

The peak area is proportional to the amount of equal protons, that is protons with the same chemical shift. Spin-spin interactions between protons in the same molecule cause splitting of signals. An NMR spectrum provides valuable information about (i) how many different kinds of nuclei there are in a molecule, (ii) the environment of each kind of nucleus, (iii) how many nuclei of each kind there are, and (iv) the environment of a nucleus with respect of other nearby nuclei. Mixtures like biological fluids show very complex NMR spectra. Due to the enormous amount of metabolites that each have many different kinds of nuclei, such mixtures show many NMR signals. Consequently, the identification of compounds in such NMR spectra is hampered.

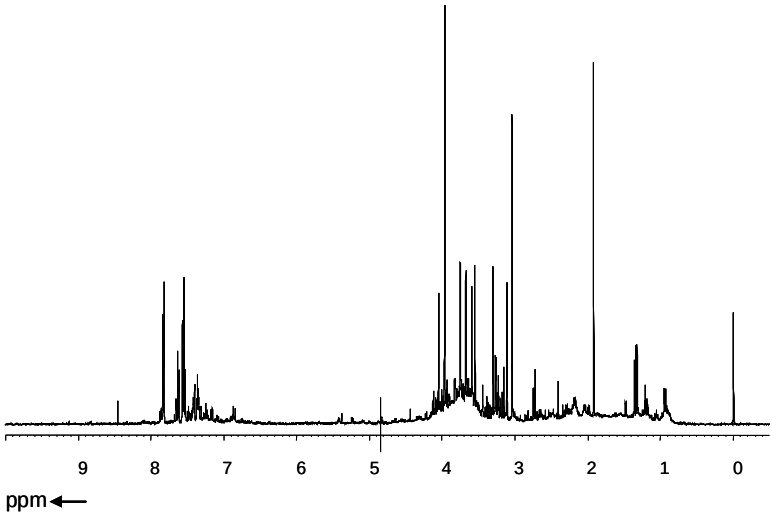

Figure 1. *An example of an NMR spectrum.*

Chromatography and Mass Spectrometry

Liquid chromatography (LC) is a separation method in which the sample components to be separated are distributed between two phases, a stationary phase and a mobile phase which is a liquid [2]. The chromatographic process occurs as a result of repeated sorption and desorption steps during the movement of sample components along the stationary phase. Separation of components takes place due to differences in interaction of the individual sample components with the stationary phase. Gas chromatography (GC) is based on the same principle, with the difference that in this case the mobile phase is a gas. It has the capability to separate very well, and in a short time, and is therefore preferable to LC. However, LC is very useful when compounds are not volatile, or not stable enough at the higher temperatures used in GC.

Both GC and LC can be combined with mass spectrometry (MS), which is a powerful tool for analysing components of thermal lability and low volatility [3]. It is a quantitative method of analysis, in which molecules of a component are ionised and (partly) broken up into fragments. The fragments are selected according to their masses and their intensities are measured. In a mass spectrum, intensities

are plotted as function of the mass fragments upon which the chemical structure of a component can be determined.

Many ionisation methods are available, such as chemical ionisation and electrospray ionisation (ESI). The latter is used in combination with LC for many biochemical analyses. It generates ions directly from the solution by creating a spray of highly charged droplets in the presence of a strong electric field. When a droplet decreases in size, the electric charge density on its surface will increase. Ions then leave the droplet when the mutual repulsion between like charges on the surface becomes large enough, upon which the escaping ions are directed into the mass analyzer.

Comparison of the techniques

Both GC-MS and LC-MS are highly suitable for identification of a broad range of metabolites in a sample, thus covering the complete metabolome. Both techniques are sensitive but require extensive sample preparation. The main advantage of NMR, compared to GC-MS and LC-MS, is the fact that a wide range of compounds in a sample can be quantified simultaneously without extensive sample preparation. The disadvantage is that NMR is less sensitive and that the identification of metabolites is more complicated. However, for global metabolic profiling NMR is an attractive approach, producing a good general impression of the contents of a sample. When interesting results show up, techniques like GC-MS, LC-MS or 2-dimensional NMR, may still be used afterwards for a more detailed study of the sample and to identify metabolites of interest.

References

1. Morrison, R.T., Boyd, R.N. (1987) Spectroscopy and Structure. In: *Organic Chemistry*. Allyn and Bacon Inc. Newton MA: 569-631.
2. Poole, C.F., Scheutte, S.A. (1984) Contemporary practice of chromatography. Elsevier, Amsterdam.
3. Niessen, W.M.A. (1998) Liquid chromatography-mass spectrometry, second edition, revised and expanded, Chromatographic Science Series. Marcel Dekker Inc., New York.

3

Multivariate data analysis

General introduction

In biomedical studies, enormous amounts of data are produced. Besides, data obtained from biological fluids with GC-MS, LC-MS as well as NMR is highly complex. The large number of metabolites present in biological fluids produces an overwhelming amount of signals. To find differences and similarities in such data, MVDA is indispensable. It visualizes the correlation between variables in complex or large data sets (e.g. hundreds of signals in NMR or LC-MS spectra) in relation to a target variable such as disease status.

In MVDA, unsupervised and supervised techniques can be used. Unsupervised methods such as principal component analysis (PCA) determine patterns within data sets, without prior knowledge, and visualize the data in such a way as to emphasize similarities and differences. With such methods, a direct comparison of for instance NMR spectra is made and samples are clustered, solely on the basis of NMR spectral similarities. Supervised methods, such as principal component discriminant analysis (PCDA) and partial least-squares (PLS), are more powerful tools. They use additional information in the analysis of the data set, such as biochemical, histopathological or clinical data, to identify differences between pre-defined groups.

In metabolic profiling, techniques like PCA, PCDA and PLS are often applied. However, new methods are being developed continuously, such as PLS batch processing and multilevel simultaneous component analysis (MSCA). These routines are extremely suitable to handle time-course data with a multilevel structure.

Component methods

In component models such as PCA, a large set of related variables is transformed to a smaller set of non-correlated variables that express the variation in the original variables to a maximum. This principle is shown in Figure 1. Two groups of samples are measured at two variables, v1 and v2. Most of the variation in the measurements is explained in the direction of the line PC1, and secondly in the direction of the line PC2, which is orthogonal to PC1. The new variables PC1 and PC2 are called components and each of them depicts an axis in multidimensional space. PC1 and PC2 are independent linear combinations of the original variables v1 and v2.

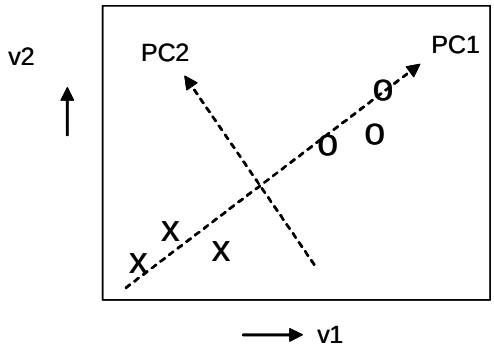

Figure 1. *Representation of principal component analysis (PCA). Six samples are measured on the original variables, v1 and v2. New non-correlated variables, so-called principal components (PC1 and PC2), describe most of the variation in the measurements.*

More in general, PCA compares objects (e.g. NMR spectra) and forms subsets of these objects on the basis of variable similarities. For this, a data matrix X, containing *I* samples and *J* variables, is transformed from a large set of related variables into a smaller set of non-correlated variables, called principal components (PC), which are chosen to express the maximum variation in the original variables [1; 2]. A mathematical description of the PCA model is given in equation (1)

$$X = TP' + E \qquad (1)$$

where X is the original *IxJ* set of data, P' is a transpose matrix (*FxJ)* of variable coefficients (loadings), T is an *IxF* matrix of object scores and E is an *IxJ* matrix containing the residuals not explained by the model using *F* principal components. Equation (1) can also be written as

$$x_{ij} = \sum_{r=1}^{R} t_{ir} p_{jr} + e_{ij} \qquad (2)$$

with x_{ij}, t_{ir}, p_{jr}, e_{ij} the typical elements of X, T, P and E and *R* the number of components.

The score on a PC is the distance of the projection of an original object on the PC to the zero point. Scores are plotted in a score plot, with the PCs as axes. For example, NMR spectra are recorded from biological fluid samples. Scores are situated close to each other in a score plot when the NMR spectra of the samples

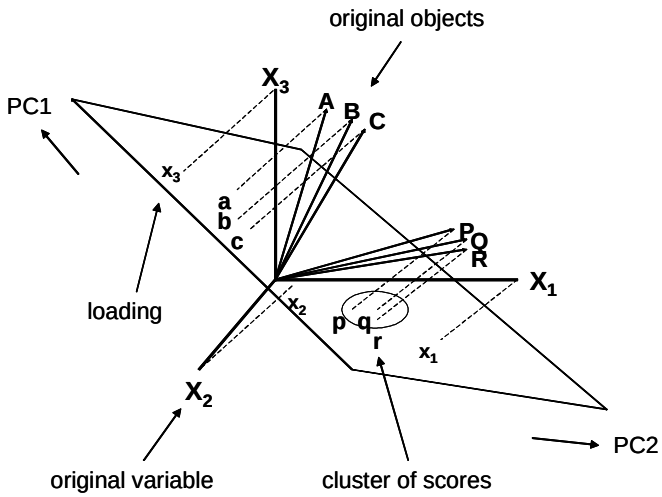

Figure 2. *A combined score and loading plot. A graphical representation of the projection of objects (e.g. NMR spectra) A, B, C, P, Q and R and the original variables (e.g. NMR signals) X_1, X_2 and X_3 onto a plane through the first and second principal component (PC1 and PC2). The lower case characters denote the scores and loads. The distance between the clusters of scores is largest in the direction of PC1.*

are similar (see Figure 2). When the clustering of scores (NMR spectra) matches the controls, treated or diseased subjects in the original study set-up, a connection can be linked up between affected NMR signals, and treatment or disease.
The contribution of a variable to a principal component is named a loading. The loading on a PC is the distance of the projection of an original variable on the PC to the zero point. A high loading indicates a strong contribution of the original variable to the investigated PC (see Figure 2). In a so-called factor spectrum, henceforth named metabolic fingerprint or biomarker profile, loadings are presented as lines (see Figure 3). Biomarker profiles are usually constructed in the direction in which the distance between clusters of scores is largest. The position of a line in a biomarker profile corresponds to the position of a variable in the original data, e.g. the position of a signal in an NMR spectrum in the mentioned example. The length of a line denotes the contribution of this variable, such as an NMR signal, to the clustering of scores in the investigated direction.

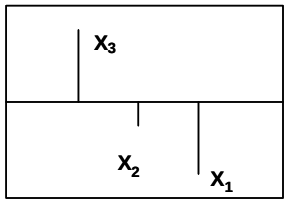

Figure 3. *An example of a factor spectrum, obtained from Figure 2 in the direction of PC1. Loadings are presented as lines. The location of the lines corresponds to the location of the variables in the original data, for instance NMR signals. The length of a line denotes the contribution of a variable to a principal component. X_1 is highly increasing for P, Q and R, whereas X_3 is highly increasing for A, B and C.*

Discriminant analysis

Discriminant analysis is often applied when the interest is centered on differences between groups, whereas component methods such as PCA are essentially dimension-reduction techniques. Discriminant analysis is based on the assumption that samples from a given group are more similar to each other than to samples from other groups. The technique aims at finding and identifying structures in the original data that show large differences in their group means. A priori knowledge

as to which samples are similar is needed; hence discriminant analysis is a supervised technique, in contrast to PCA.

Discriminant analysis combines variables in such a way that differences between predefined groups are maximized. This is illustrated in Figure 4, where two groups of samples are measured on two variables.

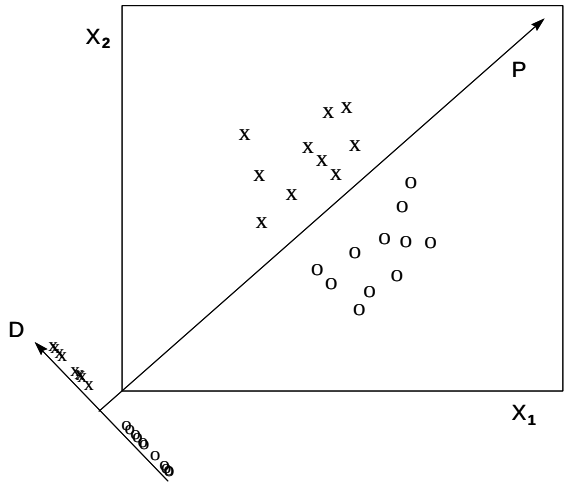

Figure 4. *Principle of discriminant analysis. D is the discriminant axis, P is a projection line, X1 and X2 are two original variables and x and o represent samples from two different groups. Projection of samples on X1 and X2 shows no separation between the two clusters, whereas projection on line D shows a complete separation.*

According to the principal component maximum variance criterion, these samples should be projected on line P. For discriminating between groups of samples, however, this is not the optimal solution. Projection of samples on line D shows a complete separation between the two clusters and is therefore a better solution [3]. Analogous to the PC's in PCA, the new calculated variable in discriminant analysis is named D, discriminant axis. Discriminant axes are expressed as:

$$D = \sum_{j=1}^{J} w_j x_j \qquad (3)$$

with *D* the discriminant score, while w_i and x_j are weighting coefficient and score, respectively, for the j^{th} out of J variables. The new variables D describe differences between means of groups. Because these new variables must provide the most efficient representation of the differences between the means of the groups, the optimization criterion is the ratio of the between-group variation over the within-group variation. In a technique like PCDA, the scores from PCA are used as a starting point for linear discriminant analysis, as depicted in Figure 5.

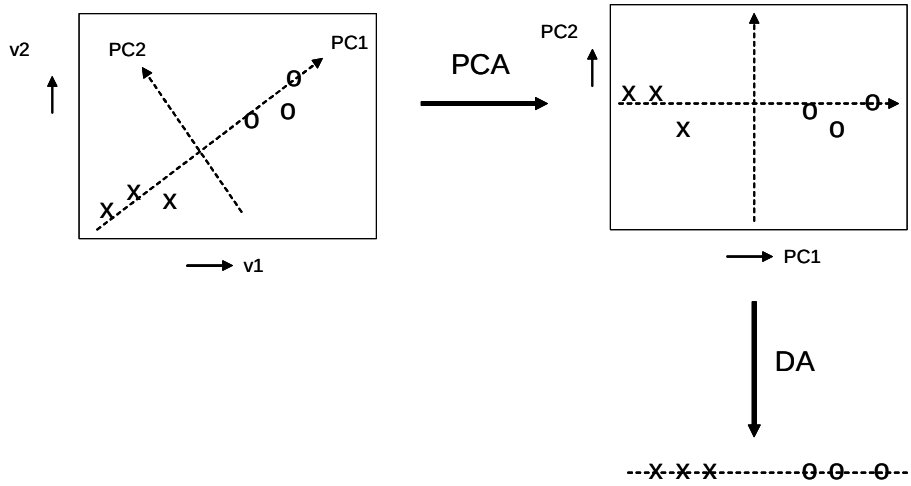

Figure 5. *Principal component discriminant analysis. The scores from PCA are used as a starting point for linear discriminant analysis (DA).*

Regression analysis

A partial least-squares model has a predictive nature, in contrast to a PCA model that describes data. In PLS, pairs of scores and loadings, so-called latent variables, are not only calculated to maximize the explained variance in the predicting data set, X, but also to maximize the correlation with Y, the data to be predicted [4]. This can be written in the equations:

$$X = TP^T + E \qquad (4)$$

and

$$Y = TQ^T + F \quad (5)$$

where X represents an I x J matrix of independent variables and Y an I x K matrix containing the dependent variables. The matrices P^T and Q^T are transpose S x J and S x K matrices, containing the dependent and the independent variable loadings, respectively. The matrix T is an I x S matrix of S latent scores, whereas E and F are I x J and I x K matrices containing the residuals of the independent and the dependent variables, respectively.

Data preprocessing

Normalisation is a procedure that is often applied to analytical data, for instance when spectra contain no reference peaks. The assumption behind this is that each sample has the same amount of information present in the spectra. Normalising the spectra makes their absolute amount of information mutually comparable.

To avoid that variables with relatively small values are hard to detect amidst variables with relatively large values, it is often necessary to center or to scale the data. In case no centering or scaling techniques are used, it may occur that after application of MVDA only one or a few variables describe the full variance present in the dataset. If variables with different units are present in the dataset it may also be necessary to center or to scale the data before the application of MVDA.

Orthogonal signal correction (OSC) is a preprocessing method that may be applied to spectral data prior to developing a PLS model [7]. In situations where a PLS model captures a very large amount of predictor block (X) variance in the first factor but gets very little of the predicted variable (Y), it can be useful to remove extraneous variance from X that is unrelated to Y. OSC does this by finding directions in X that describe large amounts of variance while being orthogonal to Y.

Reliability of MVDA models

Data for MVDA is often randomly divided into a training data set and a test data set. Models are built upon the training data set. Subsequently the test data set is used to test the reliability of the training model. This is done by passing the test

data through the training model. In this way the model's prediction of classification of the test data is obtained. Predictions should be in agreement with the actual known results.

To obtain an objective assessment of the prediction error of a model obtained by MVDA, cross validation can also be carried out. For this, the measurements of one or more samples are left out of a data set. MVDA is applied then on the remaining data. The obtained model is used to predict the scores of the left out sample(s). Repeating this procedure until each of the samples is left out once and then comparing the predicted scores with the actual known scores gives an estimate of the reliability of the MVDA model. Leaving out one sample at a time is generally referred to as leave-one-out cross validation [8].

References

1. Joliffe, I.T. (1986) Principal Component Analysis. Springer Verlag, New York.
2. Smilde, A.K., Kiers, H.A.L. (1998) Internal report: Three-way analysis in chemistry. University of Amsterdam, The Netherlands.
3. Hoogerbrugge, R., Willig, S.J., Kistemaker, P.G. (1983) Discriminant Analysis by Double Stage Principal Component Analysis. Anal. Chem. 55: 1710-1712.
4. Geladi, P., Kowalski, B.R. (1986) Partial Least Squares Regression: A Tutorial. Anal. Chim. Acta 185: 1-17.
5. Wold, S., Kettaneh, N., Friden, H., Holmberg, A. (1998) Modelling and diagnostics of batch processes and analogous kinetic experiments. Chemometrics Intell. Lab. Syst. 44: 331-340.
6. Antti, H., Bollard, M.E., Ebbels, T., Keun, H., Lindon, J.C., Nicholson, J.K., Holmes, E. (2002) Batch statistical processing of 1H NMR-derived urinary spectral data. J. Chemometrics 16: 461-468.
7. Wold, S., Antti, F., Lindgren, F., Ohman, J. (1998) Orthogonal signal correction of near-infrared spectra. Chemom. Intell. Lab. Syst. 44: 175-185.
8. Burghard, D.R., Kuznicki, J.T. (1990) Chemometrics: chemical and sensory data. CRC Press, Boca Raton.

4

Identification of disease-related metabolic fingerprints in osteoarthritic guinea pigs

Abstract

Osteoarthritis, one of the most common diseases among the elderly, is characterized by the progressive destruction of joint tissues. The etiology of osteoarthritis (OA) is largely unclear and no effective disease-modifying treatment is currently available. Metabolic fingerprinting provides a novel tool for the identification of biomarkers. A metabolic fingerprint consists of a typical combination of metabolites in a biological fluid and in this study is identified by a combination of ^{1}H Nuclear Magnetic Resonance spectroscopy (NMR) and multivariate data analysis (MVDA). The current feasibility study was aimed at identifying a metabolic fingerprint for OA. Urine samples were collected from osteoarthritic male Hartley guinea pigs (n=15) at 10 and 12 months of age, treated with a medium vitamin C dose (30 mg/d) and from healthy male Strain 13 guinea pigs (n=8) at 12 months of age, treated with 30 mg/d vitamin C. NMR measurements were performed on all urine samples. Subsequently, MVDA was carried out on the obtained NMR data. An NMR fingerprint was identified that reflected the osteoarthritic changes in guinea pigs. The metabolites that comprised the fingerprint indicate that energy metabolism is of major importance in OA. This study demonstrates the feasibility of metabolic fingerprinting to identify disease-specific profiles of urinary metabolites. NMR fingerprinting is a promising means of identifying new disease markers and of gaining fresh insights into the pathophysiology of disease.

Based on: Lamers, R.A.N., DeGroot, J., Spies-Faber, E.J., Jellema, R.H., Kraus, V.B., Verzijl,N.,TeKoppele, J.M., Spijksma, G., Vogels, J.T.W.E., van der Greef, J., van Nesselrooij, J.H.J. (2003) Identification of disease and nutrient related metabolic fingerprints in osteoarthritic guinea pigs. J. Nutr., 133, 1776-1780.

Introduction

Biomedical research is continuously facing the challenge of elucidating the relationship between health, disease and metabolism on one hand and effects of nutrition or pharmaceuticals on the other hand. Genomics will contribute to clarifying the etiology of most common genetic diseases and provide approaches for therapeutic intervention. However, knowledge of genomics is not the single universal tool for predictive medicine and nutritional strategies [1-2]. A person's phenotype results from the interaction of the genotype with the environment, in which nutrition plays a major role. Metabolites are the quantifiable molecules that best reflect phenotype [3] and are attractive candidates for biomarker fingerprints in nutritional intervention studies.

Biological fluids, such as urine and blood, contain a large number of metabolites which may provide valuable information on the metabolism of an organism, and thus about its health status. Metabolic fingerprinting, also referred to as metabolomics, metabonomics [4], metanomics [1], or related terms, is a method that enables quantification and identification of metabolites in biological fluids. The methodology has emerged from the profiling of body fluid approaches that were developed many decades ago for the study of inborn errors of metabolism and effects of nutrition. The early work in this area was mainly driven by mass spectrometric techniques (GC-MS), which allow low concentration components to be measured in single profiles. In the eighties, especially with the combination of multivariate data analysis (MVDA), mass spectrometric profiling had become a powerful fingerprinting methodology [5].

For a full coverage of a complex mixture of metabolites, a combination of analytical techniques is desirable. However, for global screening, ^{1}H Nuclear Magnetic Resonance spectroscopy (NMR) is an attractive approach, as a wide range of metabolites can be quantified at the same time without extensive sample preparation. More in-depth studies can subsequently elucidate the metabolic pathways involved, especially when metabolite information is integrated with gene expression and proteomic data in systems biology strategies [6].

NMR spectra of biological fluids are very complex due to the mixture of numerous metabolites present in these fluids. Therefore, variations between samples are often too small to be recognized by eye. In order to increase the comparability of

NMR spectra and thereby maximize the power of the subsequent data analysis, we have developed a Partial Linear Fit algorithm [7] in the past. This algorithm adjusts minor shifts in the spectra while maintaining the resolution. To find significant differences, multivariate data analysis (MVDA) is needed to explore recurrent patterns in a number of NMR spectra [4-5]. MVDA is a powerful tool for the analysis of data sets with a large number of variables. It visualizes the correlation between variables in complex or large data sets (e.g. thousands of signals in NMR spectra) in relation to a target variable such as disease status. MVDA falls into two general classes: unsupervised and supervised techniques. Unsupervised methods such as principal component analysis (PCA) determine patterns within data sets, without prior knowledge, and visualize the data in such a way as to emphasize their similarities and differences. With such methods, a direct comparison of NMR spectra is made and subsets of data are formed, solely on the basis of NMR spectral similarities.

In PCA, data is transformed from a large set of related variables (e.g. NMR signals) to a smaller set of uncorrelated variables. The new created variables are called principal components (PCs) and aim at expression of maximum variation in the original variables. Each PC forms an axis in multidimensional space and the calculated distance of an object (e.g. a complete NMR spectrum of a guinea pig urine sample) to this axis is a so-called *score*. The contribution of each variable (e.g. a single NMR signal) to a PC can also be calculated, giving a so-called *loading*. A high loading indicates a strong contribution of the original NMR signal to the investigated PC. Loadings can be displayed in a so-called factor spectrum. Loading vectors are described as lines then, with a position equal to the position of the variables in the original spectra. The height of the lines indicates the contribution of the variables to the investigated direction.

Supervised methods such as partial least squares (PLS) and principal component discriminants analysis (PCDA) are more powerful tools, which use additional information on the data set such as biochemical, histopathological or clinical data to identify differences between pre-defined groups (8). In PCDA, the scores from PCA are used as a starting point for linear discriminant analysis. Discriminant analysis works by combining the PCs in such a way that differences between pre-defined groups are maximised.

Osteoarthritis (OA), the most common form of arthritis, is a multi-factorial, chronic joint disease that is characterized by the progressive destruction of articular cartilage, resulting in impaired movement, pain and ultimately disability [9]. A variety of systemic and local risk factors have been identified that predispose to the development of OA, including -but not limited to- age, gender, bone density, obesity, joint injury and nutritional factors [10]. Despite the growing body of information on the pathogenesis of OA, its etiology is far from clear and effective disease-modifying treatment is lacking. Diagnosis of OA is currently based on clinical symptoms in combination with imaging techniques such as radiology or MRI, to visualize the degenerative changes in the joint. These changes can only be observed in an advanced stage of the disease, in which joint tissue damage is considered irreversible. Alternative methods are therefore needed in order to detect osteoarthritic changes in the joints in an early stage of the disease in a quantitative, reliable, and sensitive manner [11]. By measuring a combination of relevant metabolites in biological fluids, metabolic fingerprinting potentially meets these criteria.

The Hartley outbred strain guinea-pig develops spontaneous progressive knee OA, with features similar to the human disease [12-13] and was chosen to investigate the potential of metabolic fingerprinting as a tool for diagnosis.

Materials and Methods

Animal handling procedure

Fifteen male Hartley guinea pigs that develop OA during aging were purchased at two months of age from Charles River Laboratories (Wilmington, MA, USA) and maintained on standard guinea pig feed. The guinea pigs were maintained at a medium dose of vitamin C (30 mg) provided with feed daily (supplemented with standard Purina Lab Diet 5025 (Purina Mills, LLC, St. Louis, MO, USA) without vitamin C *ad libitum*). Furthermore, eight Strain 13 guinea pigs (obtained from Crest Caviary, Prundale, CA, USA) which develop OA to a much lesser extent than the Hartley strain [14] were housed individually in solid bottom cages and fed 30 mg/d of vitamin C. Metabolic cages (PLAS-LABS, Lansing MI) suitable for guinea pigs were used to collect 24-hour urine samples at 10 and 12 months of age for the Hartley guinea pigs and at 12 months of age for the Strain 13 guinea pigs. The

collected urines were centrifuged at 3000 rpm for 10 minutes to remove debris, and stored at –80°C until analyses. For all experiments “Principles of laboratory animal care” were followed and American guidelines and laws were applicable.

NMR analysis of urine samples

Prior to NMR spectroscopic analysis, 200 µL urine samples were lyophilized and reconstituted in 1 mL sodium phosphate buffer (0.1 mmol/L, pH 6.0, made up with D_2O), to minimize spectral variance arising from differences in urinary pH. Sodium trimethylsilyl-[2,2,3,3,-$^2H_4$]-1-propionate (TMSP; 0.025 mmol/L) was added as an internal standard. NMR spectra were recorded in random order and in triplicate in a fully automated manner on a Varian UNITY 400 MHz spectrometer using a ^1H NMR set-up operating at a temperature of 293 K.

Free induction decays (FIDs) were collected as 64K data points with a spectral width of 8.000 Hz; 45 degree pulses were used with an acquisition time of 4.10 s and a relaxation delay of 2 s. The spectra were acquired by accumulation of 128 FIDs. The signal of the residual water was removed by a pre-saturation technique in which the water peak was irradiated with a constant frequency during 2 s prior to the acquisition pulse. The spectra were processed using the standard Varian software. An exponential window function with a line broadening of 0.5 Hz and a manual baseline correction was applied to all spectra. After referring to the internal NMR reference (TMSP δ= 0.0), line listings were prepared using the standard Varian NMR software. To obtain these listings all lines in the spectra above a threshold corresponding to about three times the signal-to-noise ratio were collected and converted to a data file suitable for multivariate data analysis applications.

NMR data processing and multivariate data analysis

The NMR data reduction file was imported into Winlin (V1.10, TNO, The Netherlands). Minor variations from comparable signals in different NMR spectra were adjusted and lines were fitted without loss of resolution. To correct for urinary dilution, the data were auto-scaled so that small and large peaks contribute similarly to the final study result. Where needed endogenous and exogenous metabolites of vitamin C were eliminated from the NMR spectra leading to more

universal OA related changes and principal component discriminant analysis (PCDA) was performed [15].

Age, strain or dose was used as a priori information respectively. The NMR data set was randomly divided into a training data set and a test data set. PCDA models were built upon the training data set. Subsequently the test data set was used to test the reliability of the training model. Predictions were in agreement with the actual groupings.

The resulting discriminants were quantified for each of the urinary NMR spectra and the scores were plotted to visualize clustering. Factor spectra were used then to correlate the scores to the original NMR features in the spectra. These metabolic fingerprints provided insight into the type of metabolites responsible for differences between categories.

Results

Metabolic fingerprinting in disease diagnosis

The Hartley outbred stock albino guinea pig spontaneously develops an osteoarthritic condition that closely resembles its human counterpart. The earliest histological signs of the disease appear at 3 months of age in the medial tibial plateau which gradually progress to extensive cartilage degeneration in guinea pigs aged 12 months or older [16].

The underlying hypothesis of the present study is that OA will disturb metabolism, which will be reflected in an aberrant urinary metabolic composition. Using metabolic fingerprinting such OA-induced abnormal urinary composition may be quantified. However, also aging may cause disturbances in metabolism that are independent of a pathological change [17]. Therefore, a suitable control group was essential for the construction of a representative metabolic fingerprint for OA. To exclude the possibility that metabolic differences caused by aging interfered with those caused by OA, NMR analysis was performed on urine samples of Hartley guinea pigs that were followed longitudinally. Samples were collected of the same guinea pigs at 10 and 12 months of age, treated with 30 mg/d vitamin C. This approach minimized age effects on the metabolite profile: at 10 months of age guinea pigs are fully grown, while the OA severity is expected to increase in the guinea pigs from 10 months onward [14]. Comparison of two urinary spectra of a

single guinea pig at 10 and 12 months of age (Figure 1) showed that, on first sight, the differences between the respective NMR spectra were small.

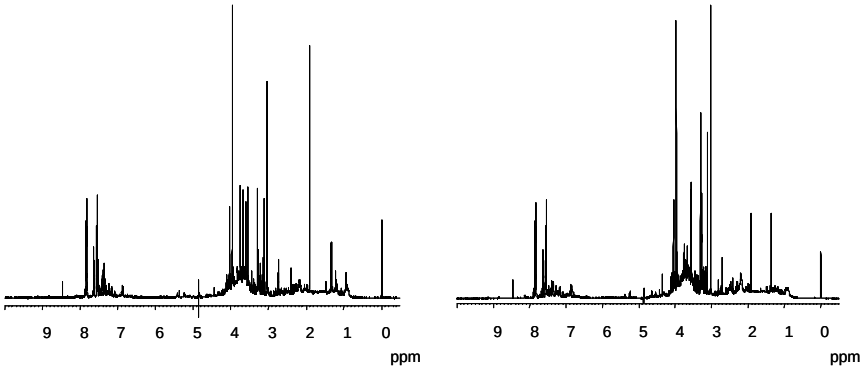

Figure 1. *NMR spectra of urine taken at 10 months (left) and 12 months (right) from one Hartley guinea pig. On first sight, no differences exist between the two spectra.*

Since the NMR spectra contain a range of different signals, combinations of which represent the different metabolites, principal component discriminant analysis (PCDA) was used to visualize disease related differences in these spectra [15], using age (10 or 12 months) as additional input information. The resulting plot (Figure 2) clearly showed that scores of samples were different per age group (and thus per OA severity). The scores of the 10 months samples are below zero, whereas the scores of the 12 months samples are above zero ($P<0.001$). The distinction between scores is based on the concentrations of metabolites which are characteristic for each group. A metabolic fingerprint was derived from the plot of the scores, which is likely to reflect OA (Figure 3). In this fingerprint, NMR signals of urinary guinea pig metabolites are visualized which increased or decreased in association with age and thus OA.

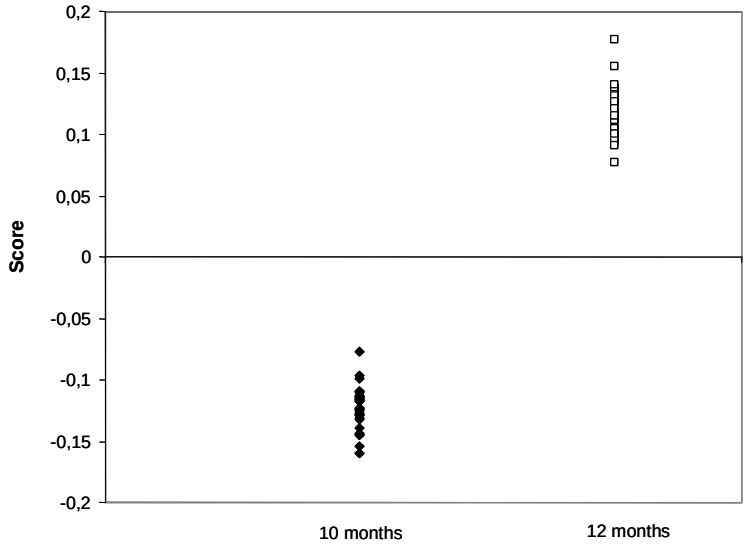

Figure 2. *PCDA score plot of urinary NMR spectra of Hartley guinea pigs at 10 versus 12 months of age (the points represent the complete urinary NMR spectra of the guinea pigs). At 10 months of age guinea pigs are fully grown, while at 12 months of age their OA severity has substantially increased (14). This increase in OA severity is reflected by the urinary composition, as there is a clear difference between the positions of the two groups (P<0.001).*

Using the approach described above, in principle the small age difference (2 months) rather than a different severity of OA could be responsible for the clustering of the data and thus the metabolic fingerprint. Therefore, to ascertain that the obtained metabolic fingerprint reflected OA differences and not age differences, a complementary approach was used. NMR spectra were obtained from urine samples of 12-months old Strain 13 guinea pigs, treated with 30 mg/d vitamin C, that show only very limited OA changes [14]. These NMR spectra were subjected to PCDA together with the data obtained from 12-months old Hartley guinea pigs, treated with 30 mg/d vitamin C. Using this design, only age-matched samples were compared, thereby completely eliminating age-related changes as confounders of the differences in the metabolite profile. For PCDA an ideal situation was assumed in which, at 12 months of age, Hartley guinea pigs have OA, whereas Strain 13 guinea pigs did not have

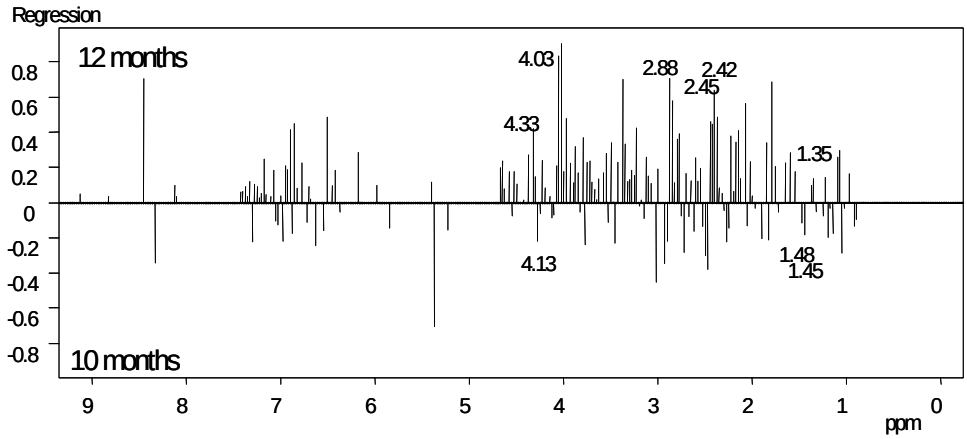

Figure 3. *Factor spectrum, or metabolic fingerprint, of urinary NMR spectra of 10 versus 12 month old Hartley guinea pigs that is typical for OA. Peaks (representing NMR signals) in the positive direction indicate metabolites that are more abundant in urine of guinea pigs with severe OA (at 12 months) than in urine of guinea pigs with milder disease (at 10 months). Consequently, metabolites that are more abundant in urine of these healthier guinea pigs are presented as peaks in the negative direction.*

OA [14]. The score plot resulting from this analysis, in which "strain" (i.e. Hartley *versus* Strain 13) was used as additional input data for PCDA, again showed a clear difference between scores of the two groups (Figure 4; $P<0.001$). The scores of Hartley guinea pigs are above zero, whereas the scores of the Strain 13 guinea pigs are below zero.

A metabolic fingerprint was used to assign the NMR signals that explain differences between the two groups, and thus health status (Figure 5). In this analysis, both OA as well as strain differences could in principle be responsible for the separation in groups.

Combining the results from both analyses allowed elimination of the confounding factor from each of the approaches, namely age and strain respectively. The metabolic fingerprint obtained from the comparison of urinary NMR spectra of 10 and 12 months old Hartley guinea pigs was almost similar to the metabolic fingerprint obtained from the age-matched strain comparison (compare Figure 2B and 3B). Thus, metabolites that were present in both fingerprints are likely to be OA specific. Using this approach, NMR regions in the metabolic fingerprint that emerged as strongly associated with OA are 2.0-2.9, 3.0-4.7 and 6.3-7.5 ppm. Signals at 1.35, 2.42, 2.45, 2.88, 4.03 and 4.33 ppm in a positive direction and

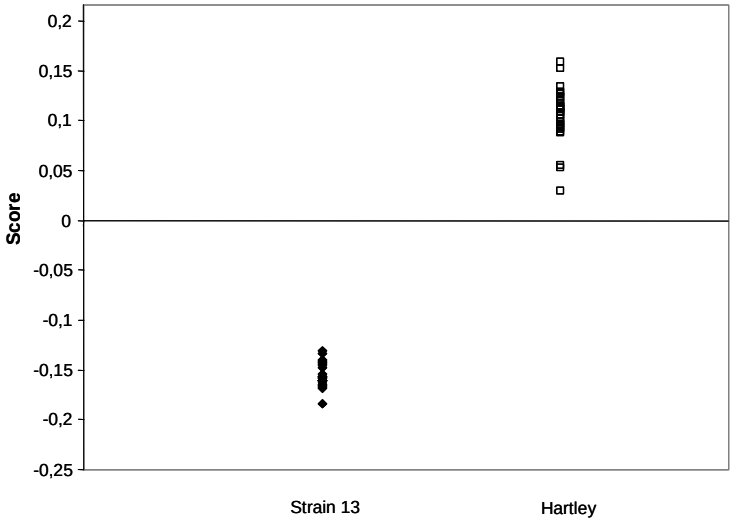

Figure 4. *PCDA score plot of urinary NMR spectra of 12 months old Hartley versus Strain 13 guinea pigs. Strain 13 guinea pigs are healthy, while Hartley guinea pigs suffer from OA (14). The difference in urinary composition is reflected by the clear separation into two groups in the graph (P<0.001).*

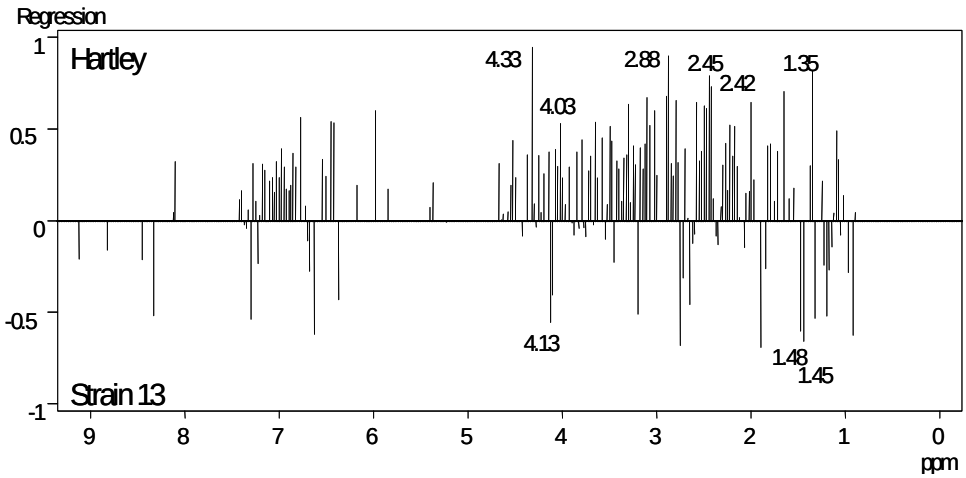

Figure 5. *OA specific metabolic fingerprint of urinary NMR spectra of 12 months old Hartley versus Strain 13 guinea pigs. Peaks in the positive direction indicate metabolites that are more abundant in urine of Hartley guinea pigs than in urine of Strain 13 guinea pigs. Consequently, metabolites that are more abundant in urine of Strain 13 guinea pigs are presented as peaks in the negative direction.*

1.45, 1.48 and 4.13 in a negative direction were especially abundant in the OA state. Signals in these regions represent, amongst others, hydroxybutyrate, creatine/creatinine, pyruvate, and lactic acid. These metabolites are consistent with the hypothesis suggesting altered energy utilization with OA [18].

Discussion

In the present study, a novel method was employed to identify a metabolic fingerprint for OA, using NMR in combination with MVDA on urine samples of outbred male Hartley guinea pigs that spontaneously develop knee OA. An NMR fingerprint for OA was identified that was independent of age or strain effects and therefore can be used as a diagnostic tool for OA in guinea pigs.

In this study it was essential to use samples from a well-defined animal study, rather than using human material, since differences in OA severity, medication, diet, and habits etc create additional variability which would greatly hinder *de novo* identification of an OA specific urinary fingerprint. However, once identified amidst the thousands of other metabolites, the urinary fingerprint for OA may also be quantified in humans. A feasibility study to this purpose will be initiated shortly.

This study demonstrates the feasibility of metabolic fingerprinting to identify metabolite profiles in (pre)clinical research. As shown here, this technique has the ability to distinguish a disease from a non-disease state. In addition to their contribution to the fingerprint, the individual metabolites may provide additional insight into the pathogenesis of OA. Lactic acid, malic acid, hypoxanthine and alanine contributed heavily to the fingerprint, suggesting their involvement in the osteoarthritic process. Altered energy demand may thus play an important role in OA. Further identification of the metabolites involved and combining the current metabolic data with proteomics and genomics approaches in order to form a holistic, integrated picture of the metabolic pathways implicated in OA, may provide new insights into OA pathogenesis and thereby identify new disease targets. This approach also has the potential to catalyze development of new biomarkers for OA. Metabolic fingerprinting has the ability to distinguish disease specific metabolites. Metabolic fingerprints, as demonstrated in this study of OA, can also provide a sensitive outcome measurement tool that can be used to evaluate the effects of a nutrient or drug intervention on the incidence and progression of disease.

Therefore, this powerful technique has broad applicability in the field of clinical nutritional and pharmaceutical research.

Acknowledgements

This work was partially funded by a grant from the Arthritis Foundation (to Virginia B. Kraus). The authors gratefully acknowledge Janet L. Huebner at Duke University Medical Center, Durham, NC, for technical expertise and histopathological scoring.

References

1. Watkins S.M.: Comprehensive lipid analysis: a powerful metanomic tool for predictive and diagnostic medicine (2000) Isr Med Assoc J 2:722-724.
2. van Ommen B., Stierum R.: Nutrigenomics: exploiting systems biology in the nutrition and health area (2002) Curr Opin Biotechnol13:517-521.
3. Watkins S.M., Hammock B.D., Newman J.W., German J.B.: Individual metabolism should guide agriculture toward foods for improved health and nutrition (2001) Am J Clin Nutr 74:283-286.
4. Nicholson J.K., Lindon J.C., Holmes E.: ''Metabonomics'': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data (1999) Xenobiotica 29:1181-1189.
5. van der Greef J.: The role of metabolomics in Systems Biology, a new vision for drug discovery and development, Metabolic profiling: its role in biomarker discovery and gene function analysis (2002) Edited by GG Harrigan, R Goodacre. Boston, Kluwer Academic Publishing.
6. Ideker T., Galitski T., Hood L.: A new approach to decoding life: systems biology (2001) Annu Rev Genomics Hum Genet 2:343-372.
7. Vogels J.T.W.E., Tas A.C., Venekamp J., van der Greef J.: Partial linear fit: a new NMR spectroscopy preprocessing tool for pattern recognition applications (1996) J Chemometrics 10:425-438.
8. Lindon J.C., Nicholson J.K., Holmes E., Everett J.R.: Metabonomics: metabolic processes studies by NMR spectroscopy of biofluids combined with pattern recognition (2000) Concepts Magn Reson 12:289-320.
9. Creamer P., Hochberg M.C.: Osteoarthritis (1997) Lancet 350:503-508.
10. Felson D.T.: Osteoarthritis: new insights. Part 1: The disease and its risk factors (2000) Ann Intern Med 133:635-646.
11. DeGroot J., Bank R.A., Tchetverikov I., Verzijl N., TeKoppele J.M.: Molecular markers for osteoarthritis: the road ahead (2002) Curr Opin Rheumatol 14:585-589.
12. Huebner J.L., Otterness I.G., Freund E.M., Caterson B., Kraus V.B.: Collagenase 1 and collagenase 3 expression in a guinea pig model of osteoarthritis (1998) Arthritis Rheum 41:877-890.
13. Bendele A.M.: Animal models of osteoarthritis (2001) J Musculoskel Neuron Interact 1:363-376.
14. Huebner J.L., Hanes M., Beekman B., Tekoppele J.M., Kraus V.B.: A comparative analysis of bone and cartilage metabolism in two strains of guinea-pig with varying degrees of naturally occurring osteoarthritis (2002) Osteoarthritis Cartilage 10:758.
15. Hoogerbrugge R., Willig S.J., Kistemaker P.G.: Discriminant analysis by double stage principal component analysis (1983) Anal Chem 55:1710-1712.
16. Bendele A.M., White S.L., Hulman J.F.: Osteoarthrosis in guinea pigs: histopathologic and scanning electron microscopic features (1989) Lab Anim Sci 39:115-121.
17. Timbrell J.: Principles of biochemical toxicology (2000) London, Taylor and Francis: 117-171.

18. Damyanovich A.Z., Staples J.R., Marshall K.W.: 1H NMR investigation of changes in the metabolic profile of synovial fluid in bilateral canine osteoarthritis with unilateral joint denervation (1999) Osteoarthritis Cartilage 7:165-172.

5

Identification of a urinary metabolite profile associated with osteoarthritis

Abstract

Objective: Osteoarthritis (OA) is one of the most common diseases among the elderly. The main characteristic is the progressive destruction of articular cartilage. We lack quantitative and sensitive biomarkers for OA to detect changes in the joints in an early stage of the disease. In this study, we investigated whether a urinary metabolite profile could be found that could serve as a diagnostic biomarker for osteoarthritis in humans. We also compared the profile we obtained previously in the guinea pig spontaneous OA model.

Methods: Urine samples of 92 participants (47 non-OA controls and 45 individuals with radiographic OA of the knees or hips) were selected from the Johnston County Osteoarthritis Project (North Carolina, USA). Participants ranged in age from 60 to 84 years. Samples were measured by ^{1}H Nuclear Magnetic Resonance spectroscopy (NMR) with subsequent principal component discriminant analysis (PCDA).

Results: Differences were observed between urine NMR spectra of OA cases and controls ($P<0.001$ for both male and female subjects). A metabolite profile could be determined which was strongly associated with osteoarthritis. This profile largely resembled the profile previously identified for guinea pigs with OA. A correlation was found between the metabolite profile and radiographic OA severity (R^2= 0.82 (male); R^2= 0.93 (female)).

Conclusion: This study showed that a urine metabolite profile may serve as a novel discriminating biomarker of osteoarthritis.

Submitted for publication.

Introduction

Osteoarthritis (OA), or cartilage degeneration, is the most common form of arthritis [1]. An important feature of the disease is the progressive destruction of articular tissues, resulting in impaired movement, pain and ultimately disability. A variety of systemic and local risk factors have been identified that predispose to the development of OA, including age, gender, bone density, obesity, joint injury and nutritional factors [2]. Despite the growing knowledge on the pathogenesis of OA, its etiology is still not clear and effective disease-modifying treatment is lacking. Diagnosis of OA is currently based on clinical symptoms [3-5] in combination with radiology. Radiological evaluation of joints mainly images bone and is relatively insensitive: a follow-up period of two years is often needed to assess disease progression [6]. Magnetic Resonance Imaging (MRI) has the ability to simultaneously visualize all joint tissues. The technique is currently being optimized but has not yet reached its full potential.

Alternative methods are therefore needed in order to detect osteoarthritic changes in the joints in an early stage of the disease in a quantitative, reliable, and sensitive manner. Biomarkers that monitor molecular events taking place during disease are well suited for this purpose. A good biomarker is disease-specific, reflects actual disease progression, is sensitive to changes due to therapeutic intervention and can predict disease outcome. Currently no single biomarker exists that meets these requirements [6]. Combining several biomarkers has been shown to improve the discriminatory capability considerably [7]. Recent developments in the field of metabolomics now provide the tools to go one step further: identify profiles of metabolites that together serve as a biomarker [8-9]. Biological fluids, such as urine and blood, contain a large number of metabolites that may provide valuable information on the metabolism of an organism, and thus about its health status. Metabolic profiling, also referred to as metabolomics, metabonomics [8], or related terms, is a technique that enables quantification and identification of metabolites in biological fluids. The technology has emerged from approaches to the profiling of body fluid that were developed many decades ago for the study of inborn errors of metabolism and the effects of nutrition. Our previous research has shown that such a metabolomics approach is also feasible for identifying a biomarker profile for OA. We discovered differences between urine samples of 10- and 12-month old Hartley

guinea pigs that spontaneously develop OA using nuclear magnetic resonance spectroscopy (NMR) and multivariate data analysis (MVDA). A metabolite profile was detected which was strongly associated with OA [10]. For initial metabolic profiling it was essential to use samples from a well-defined animal study, rather than using human material, since differences in OA severity, medication, diet and habits create additional variability that would greatly hinder the identification of an OA specific urinary metabolic profile in humans. However, once identified amidst the numerous other metabolites, the urinary metabolic profile for OA in the guinea pig model may also be quantified in humans. The present study was initiated from this assumption and designed to identify a biomarker profile that could distinguish unaffected from OA affected individuals.

Materials and Methods

Study population and sample selection

Urine samples of 92 subjects were obtained from participants in the ongoing Johnston County Osteoarthritis Project (North Carolina, USA), that is described in detail elsewhere [11]. Participants were selected who were not using any medicine for joint complaints (NSAIDS and COX-2 inhibitors), varied in age between 60 and 84 years, and had a body mass index (BMI) between 21 and 34. Radiographic knee osteoarthritis was defined from weight bearing bilateral anteroposterior radiographs of the knee, according to the Kellgren-Lawrence (K-L) grading scheme [12]. An OA case was defined as K-L grade $\geq$2 of at least two joints out of the four joints considered (knees and hips); controls were defined as K-L grade 0 in both knees and 0 or 1 in both hips. The selected group of participants consisted of 47 controls (20 male and 27 female) and 45 patients with radiographic knee and/or hip OA (21 male and 24 females).

Second morning voided urines were collected and centrifuged at 3000 rpm for 10 minutes to remove debris, and stored at –80°C until analyses. The study was approved by the Institutional Review Board of the University of North Carolina School of Medicine and the Centers for Disease Control and Prevention. Written informed consent was obtained from all participants.

NMR analysis of urine samples

Prior to NMR spectroscopic analysis, 1 mL urine samples were lyophilized and reconstituted in 1 mL sodium phosphate buffer (0.1 mmol/L, pH 7.4, made up with D_2O), to minimize spectral variance arising from differences in urinary pH. Sodium trimethylsilyl-[2,2,3,3,-$^2H_4$]-1-propionate (TMSP; 0.1 mmol/L) was added as an internal standard. NMR measurements were carried out in random order and in triplicate in a fully automated manner on a 600 MHz spectrometer (Avance, Bruker BioSpin GmbH, Rheinstetten, Germany), using a proton NMR set-up operating at a temperature of 300K. For each sample, 128 free induction decays (FID) were collected. Each FID was induced using a 45-degree pulse, an acquisition time of 2.73 s and a relaxation delay of 2 s. The FIDs were collected as 64K data points with a spectral width of 12.000 Hz. The spectra were processed using the standard Bruker software. An exponential window function with a line broadening of 0.3 Hz and a manual baseline correction were applied to all spectra. After referring to the internal NMR reference (TMSP δ= 0.0), line listings were prepared with the standard Bruker NMR software. To obtain these listings all lines in the spectra above a threshold corresponding to about three times the signal-to-noise ratio were collected and converted to a data file suitable for multivariate data analysis applications. The NMR data file was imported into Winlin (V2.1, TNO, The Netherlands). Minor variations from comparable signals in different NMR spectra were adjusted and lines were fitted without loss of resolution, after which MVDA was carried out.

Multivariate data analysis

To correct for urinary dilution NMR data were centered and scaled to unit variance so that small and large peaks contributed similarly to the final study result. Subsequently Winlin was used to perform principal component discriminant analysis (PCDA) on the data. Principal component discriminant analysis is a powerful tool to identify and maximize differences between pre-defined groups in data sets with a large number of variables [13]. In this study, health status (controls versus OA cases) was used as a priori knowledge for PCDA discrimination. The NMR data set was randomly divided into a training data set and a test data set. The PCDA models were built upon the training data set. Subsequently the test data

set was used to test the reliability of the training model. Predictions were in agreement with the actual groupings (controls versus cases).

To visualize differences between NMR spectra of controls and OA cases, the PCDA scores were plotted and the unpaired T-test was performed to evaluate the statistical significance of the difference between the PCDA scores of the two groups (Excel Office 2003, Microsoft Corporation, USA). Subsequently, the original NMR spectra were used to calculate which metabolites contributed to the PCDA scores that distinguished controls from OA cases (for male and female case separately). The combination of these metabolites (visualized via the position in ppm within the NMR-spectrum) formed (gender-specific) metabolic profiles for OA. These metabolic profiles provided insight into the type of metabolites responsible for the difference in PCDA scores of the two groups.

Partial least square (PLS) regression analysis was carried out in Matlab (Version 6.5, The MathWorks Inc., Natick, MA, USA) using the PLS toolbox (Version 3.0, Eigenvector Research Inc., Manson, WA, USA) to correlate urine NMR spectra with the sum of knee and hip K-L grades, a measure for OA status. Leave-one-out was used as a method for cross-validation to obtain a goodness of fit (R^2) for the PLS model in its prediction of K-L grades from the urine NMR spectra. The PLS regression vectors, showing NMR signals that were correlated to the K-L grade, were plotted using Excel.

Results

Description of sample

The characteristics of the OA and control participants are given in Table 1. To minimize variation in urinary metabolites (and thus increase the chance of finding an OA-specific metabolic fingerprint) male and female subjects were analysed separately and for both genders, cases and controls were matched for age, height and weight. No consistent statistically significant differences were observed except for the presence of OA. The OA patients had a summed K-L grade (knee and hips) ranging from 4 and 10 while the control subjects had a summed K-L grade ranging from 0 – 2.

	Female		
	Control		Case
	mean (sd) range	p value	mean (sd) range
Number	27		24
Age [years]	69.5 (6.0) [60 - 84]	0.253	67.7 (4.8) [61 - 77]
Height [inches]	63.2 (2.0) [57.5 - 66.5]	0.183	62.4 (2.3) [58.3 - 65.5]
Weight [lbs]	161 (26) [109 - 198]	0.411	156 (25) [106 - 204]
Body mass index	28.4 (4.3) [20.9 - 34.0]	0.715	28.0 (3.5) [21.9 - 33.9]
BMD left hip	0.9 (0.1) [0.55 - 1.16]	**0.049**	0.9 (0.1) [0.65 - 1.18]
BMD spine	63.2 (0.2) [0.49 - 1.31]	0.112	1.1 (0.2) [0.73 - 1.46]
Affected joint [#]	0.0 (0.0) [0 - 0]	**< 0.001**	2.4 (0.7) [2 - 4]
K/L grade right knee	0.0 (0.0) [0 - 0]	**< 0.001**	1.3 (1.1) [0 - 3]
K/L grade left knee	0.0 (0.0) [0 - 0]	**< 0.001**	1.5 (1.2) [0 - 4]
K/L grade right hip	0.9 (0.3) [0 - 1]	**< 0.001**	1.8 (0.7) [0 - 3]
K/L grade left hip	1.0 (0.2) [0 - 1]	**< 0.001**	1.7 (0.7) [0 - 3]
Summed K/L grade	1.9 (0.5) [0 - 2]	**< 0.001**	6.2 (1.6) [4 - 9]
Current NSAID use	none		none
Current Cox-2 inhibitor use	none		none

	Male		
	Control		Case
	mean (sd) range	p value	mean (sd) range
Number	20		21
Age [years]	66.9 (4.0) [61 - 75]	0.512	67.8 (4.8) [61 - 76]
Height [inches]	68.4 (3.0) [61.0 - 73.5]	0.774	68.1 (2.9) [63.3 - 73.5]
Weight [lbs]	177 (17) [144 - 212]	0.090	188 (23) [139 - 237]
Body mass index	26.6 (2.2) [24.0 - 31.4]	**0.032**	28.4 (3.0) [23.6 - 34.0]
BMD left hip	1.0 (0.1) [0.74 - 1.31]	0.453	1.0 (0.2) [0.76 - 1.52]
BMD spine	1.2 (0.2) [0.89 - 1.59]	0.566	1.2 (0.1) [0.96 - 1.54]
Affected joint [#]	0.0 (0.0) [0 - 0]	**< 0.001**	2.2 (0.6) [2 - 4]
K/L grade right knee	0.0 (0.0) [0 - 0]	**< 0.001**	1.5 (1.0) [0 - 3]
K/L grade left knee	0.0 (0.0) [0 - 0]	**< 0.001**	1.3 (1.0) [0 - 3]
K/L grade right hip	0.8 (0.4) [0 - 1]	**< 0.001**	1.7 (0.7) [0 - 3]
K/L grade left hip	0.9 (0.3) [0 - 1]	**< 0.001**	1.6 (0.5) [1 - 2]
Summed K/L grade	1.7 (0.6) [0 - 2]	**< 0.001**	6.1 (1.4) [4 - 10]
Current NSAID use	none		none
Current Cox-2 inhibitor use	none		none

Table 1. *The characteristics of the OA and control participants.*

Human urinary metabolite profile

The underlying hypothesis of the present study is that OA leads to, or is accompanied by, metabolic disturbances that are reflected in an aberrant urinary metabolite composition. NMR with subsequent multivariate data analysis revealed such OA-related alterations in the urinary metabolite composition, resulting in a metabolic biomarker fingerprint that distinguished healthy individuals without OA from individuals with OA.

When applying PCDA on the male and female NMR data sets, clear differences for both genders were observed between the NMR spectra of OA cases and controls (P<0.001 for both male and female subjects). Plots of the PCDA scores clearly showed this (Figure 1A and 1C). Thus, PCDA resulted in a distinct separation

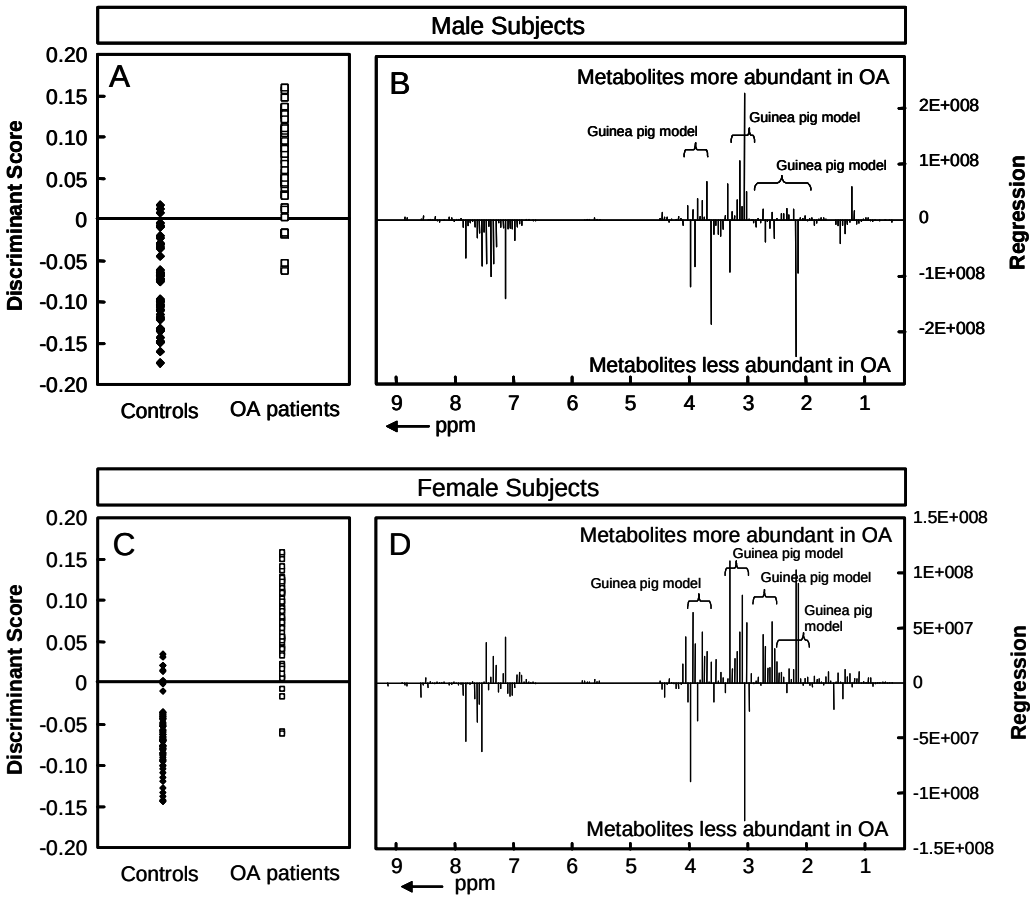

Figure 1. *Plot of the scores of urinary NMR spectra of male **(1A)** and female **(1C)** participants without OA versus participants with OA (the points represent the complete urinary NMR spectra of the subjects). The prevalence of OA is reflected by the urinary composition, as there is a clear difference between the scores of the two groups (P<0.001 for both male and female subjects). In addition, metabolite profiles of male **(1B)** or female **(1D)** participants without OA versus participants with OA. Peaks (representing NMR signals, expressed in ppm) in the positive direction indicate metabolites that are more abundant in urine of participants with OA than in urine of participants without OA. Consequently, metabolites that are less abundant in urine of OA subjects as compared to the controls are presented as peaks in the negative direction. Signals that were also found in the guinea pig OA study are indicated.*

between groups based on the metabolite composition and metabolite concentration of urine which is characteristic for each group. In other words, a specific combination of metabolites can distinguish OA cases from controls.
The profiles of these metabolites for male and female subjects are shown in Figure 1B and 1D. In these 'fingerprints', NMR signals of urinary metabolites were depicted according to their relative abundance in OA versus non-OA subjects. NMR signals of the metabolic profile that displayed an association with OA for both males as well as females and that showed up in the same direction were δ 1.18, 2.02, 2.22, 2.38, 2.58, 2.74, 3.02, 3.14, 3.18, 3.22, 3.26, 3.70, 3.74, 3.78, 3.94 ppm in the positive direction and δ 1.38, 3.58, 3.98, 7.02, 7.06, 7.10, 7.54, 7.58, 7.62, 7.66, 7.78, 7.82 and 7.86 ppm in the negative direction. Signals at δ 3.06 ppm in a positive direction and δ 2.14, 2.18, 3.30, 3.62, 7.14, 7.38 ppm in the negative direction varied strongly between males and females.

Metabolite profile in relation to summed KL-grade

Partial least squares regression was performed to correlate urine NMR spectra with the summed knee and hip K-L grades of subjects. A model was obtained that could predict the K-L grade from the urine NMR spectra for male participants (R^2=0.82) and female participants (R^2=0.93) (Figure 2 left and right panel respectively), thus showing the sensitivity of the metabolite profile for OA. The NMR patterns (thus metabolite profiles) that were shown by PLS regression to be correlated to the K-L grade, were largely similar to the metabolite profiles obtained by PCDA on controls *versus* OA cases, respectively (Figure 3). NMR signals that correlated to OA and that showed similar patterns in men and women were δ 1.18, 2.38, 2.58, 2.74, 3.10, 3.14, 3.18, 3.70, 3.74, 3.78, 3.94 ppm in a positive direction and δ 1.38, 3.58, 3.98, 7.02, 7.10, 7.54, 7.58, 7.62, 7.66, 7.82 and 7.86 ppm in a negative direction.

Identities of NMR signals

Although the NMR signals that make up the fingerprints are not yet characterized (which would require extensive mass spectroscopy), rough identification of the metabolites can be done based on the comparison of their NMR pattern with databases containing NMR signatures of known molecules. According to these databases, the NMR signals at δ 1.18, 2.38, 2.58, 2.74, 3.10, 3.14, 3.18, 3.70,

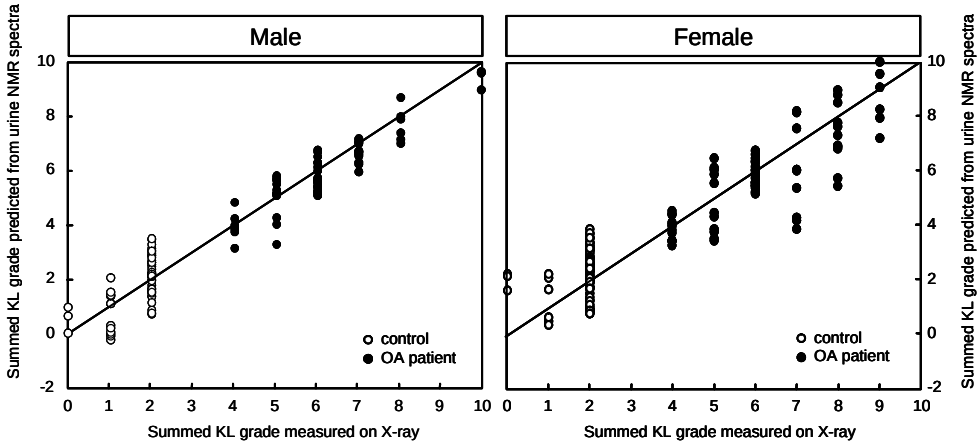

Figure 2. *Plot of measured summed K-L grades versus predicted summed K-L grades from urine NMR profiles for male **(2A)** and female **(2B)** participants, obtained with a PLS model. The goodness of fit (R^2) of the model is 0.93 and 0.82 for male and female subjects respectively, showing that the correlation between K-L grades and the urine metabolite profile is high.*

3.74, 3.78, 3.94 ppm in the positive direction (levels increased with OA) represent, among others, metabolites like hydroxybutyrate, pyruvate, creatine/creatinine and glycerol. Signals at δ 1.38, 3.58, 3.98, 7.02, 7.06, 7.10, 7.54, 7.58, 7.62, 7.66, 7.78, 7.82 and 7.86 ppm in a negative direction (levels decreased with OA) represent, among others, compounds like histidine and methylhistidine.

Discussion

A biomarker for OA that is useful for early disease detection, that reflects the course of joint destruction, and that predicts long-term outcome is currently lacking [14]. Such a biomarker could be used to assess disease progression and the effects of therapy and thereby serve as an outcome measure in clinical trials. Ultimately, an ideal OA biomarker would potentially facilitate the development of effective individualized treatment plans and approaches.

Classical biomarker development, based on the detection of known tissue synthesis and resorption markers, has thus far not yielded biomarkers sufficiently specific and/or sensitive enough for the above mentioned applications [6]. The

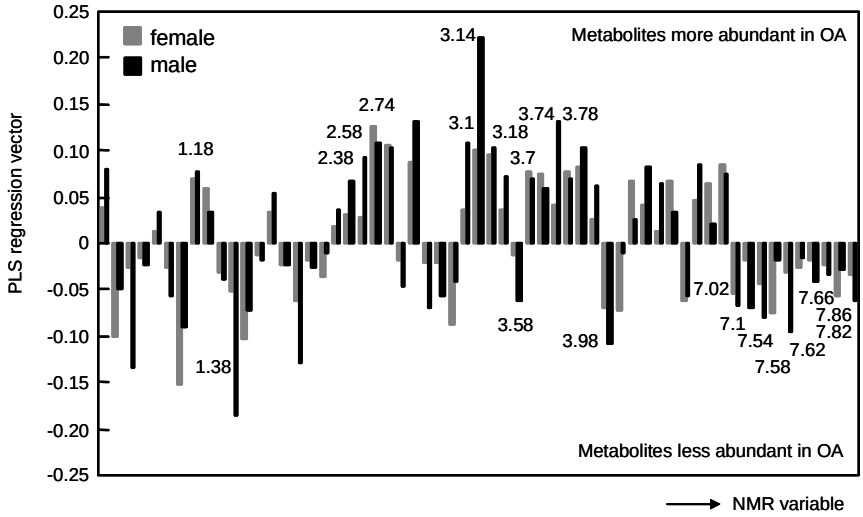

Figure 3. *Plot showing the NMR signals that correlate to the K-L grades according to PLS for both male and female participants. Peaks (representing NMR signals) in the positive direction indicate metabolites that are more abundant in urine of patients with OA than in urine of participants without OA. Consequently, metabolites that are less abundant in urine of participants with OA are presented as peaks in the negative direction. There is a large similarity with the metabolite profiles obtained with PCDA.*

recent and rapid progression in the field of NMR with subsequent MVDA has heralded the advent of a new approach to OA biomarker development. From the overall mixture of metabolites in a biological fluid like urine, a combination of molecules can be identified that together best reflect a disease process [8].

In the present study, we were able to discriminate between subjects who did not have OA and subjects with radiological OA, based on small differences in urinary metabolite composition and metabolite levels as detected by NMR and subsequent MVDA. We identified a urine metabolite profile that was strongly associated with OA and which appeared in male as well as in female subjects. A regression model showed that this metabolite profile correlated with the summed Kellgren Lawrence scale of radiological OA, for male as well as female subjects. Hence, from the metabolite profile it was possible to discriminate between OA cases and controls and, moreover, to predict the OA state and severity in a sensitive manner. The metabolite profile that we identified in human samples strongly resembled the one

we identified previously in samples from the guinea pig model for OA [10]. This further supports the supposition that the metabolite profile could serve as a biomarker for OA.

In addition to its use as a biomarker for osteoarthritis, the OA-specific metabolic fingerprint also provides information on the cellular processes that occur during the disease and as such, the metabolic profile contributes to our understanding of the pathophysiology of the disease. The presence of hydroxybutyrate, pyruvate, creatine/creatinine and glycerol in the metabolite profile could point at an enhanced use of fat, and hence an altered energy utilization. This is consistent with studies described in the literature in which the involvement of altered energy metabolism in OA has been proposed [15]. It is possible that some of the metabolites associated with OA are intermediaries in these metabolic pathways.

The metabolite profile for human OA also indicated an alteration in histidine metabolism. Our results demonstrated lower levels of histidine and methylhistidine in association with OA. This finding is consistent with the hypothesis that histidine is consumed by metabolism to histamine, itself responsible for stimulating the proliferation of articular chondrocytes into clusters, a characteristic of OA cartilage [16-17]. The synthesis of histamine from histidine is catalysed by histidine decarboxylase [18-19] and both histamine and histidine decarboxylase have been demonstrated in chondrocytes of OA cartilage [16-17]. These results would suggest that lowered levels of histidine may be caused by over-expression of histidine decarboxylase. However, more studies are needed to confirm this hypothesis.

In conclusion, our study provides evidence of a diagnostic metabolite profile associated with OA that correlates with K-L grades. Our findings are consistent with other studies reporting effects on altered energy and histidine metabolism in association with OA. The metabolite profile may provide a sensitive outcome measurement tool that can be used to evaluate the effects of nutrients and drugs on the incidence and progression of the disease. Results are promising but further research will be necessary to validate this hypothesis. Moreover, this metabolite profile may provide a tool to allow physicians to better quantify the extent of disease. Mass-spectroscopy-based identification of the unknown metabolites will be an important next step in promoting an understanding of the disease

Acknowledgements

This study was partially supported by a grant from the Dutch Arthritis Association and the Association of Schools of Public Health/Centers for Disease Control and Prevention, S1734 and the National Institute of Arthritis, Musculoskeletal, and Skin Diseases Multipurpose Arthritis and Musculoskeletal Disease Center 5-P60-AR30701 (JMJ, JBR, ADD, GL). Elly J. Spies-Faber and Gerwin K. Spijksma are gratefully acknowledged for carrying out NMR experiments.

References

1. Lawrence, R.C., Helmick C.G., Arnett F.C. (1998) Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. Arthritis Rheum. 43: 778-799.
2. Felson, D.T. (2000) Osteoarthritis: new insights. Part 1: The disease and its risk factors. Ann. Intern. Med. 133: 635-646.
3. Altman R., Alarcon G., Appelroth D. (1986) The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the knee. Arthritis Rheum. 29: 1039-1049.
4. Altman R., Alarcon G., Appelrouth D., Bloch D., Borenstein K., Brandt K. (1990) The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hand. Arthritis Rheum. 33: 1601-1610.
5. Altman R., Alarcon G., Appelrouth D., Bloch D., Borenstein D., Brandt K. (1991) The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Arthritis Rheum. 34: 505-514.
6. DeGroot J., Bank R.A., Tchetverikov I., Verzijl N., TeKoppele J.M. (2002) Molecular markers for osteoarthritis: the road ahead. Curr. Opin. Rheumatol. 14: 585-589.
7. Garnero P., Rousseau J.C., Delmas P.D. (2000) Molecular basis and clinical use of biochemical markers of bone, cartilage, and synovium joint diseases. Arthritis Rheum. 43: 953-968.
8. Nicholson J.K., Lindon J.C., Holmes E.: ''Metabonomics'': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data (1999) Xenobiotica 29:1181-1189.
9. Van der Greef J, Davidov E, Verheij ER, Vogels J, van der Heijden R, Adourian AS, Oresic M, Marple EW, Naylor S. The role of metabolomics in Systems Biology (Chapter 10) In: *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis.* Edited by Harrigan GG, Goodacre R. Boston/Dordrecht/London: Kluwer Academic Publishers; 2003: 170-198.
10. Lamers, R.A.N., DeGroot, J., Spies-Faber, E.J., Jellema, R.H., Kraus, V.B., Verzijl, N., TeKoppele, J.M., Spijksma, G., Vogels, J.T.W.E., van der Greef, J., van Nesselrooij, J.H.J. (2003) Identification of disease and nutrient related metabolic fingerprints in osteoarthritic guinea pigs. J. Nutr.133: 1776-1780.
11. Jordan J.M., Linder G.F., Renner J.B., Fryer J.G. (1995) The impact of arthritis in rural populations. Arthritis Care Res. 8: 242-250.
12. Kellgren, J.H., Lawrence, J.S. (1957) Radiological assessment of osteoarthritis. Ann. Rheum. Dis. 16: 494-502.
13. Hoogerbrugge R., Willig S.J., Kistemaker P.G.: Discriminant analysis by double stage principal component analysis (1983) Anal. Chem. 55:1710-1712.
14. Otterness, I.G., Weiner, E., Swindell, A.C., Zimmerer, R.O., Ionescu, M., Poole, A.R. (2000) An analysis of 14 molecular markers for monitoring osteoarthritis: segregation of the markers into clusters and distinguishing osteoarthritis at baseline. Osteoarthritis Cartilage 8: 180-185.

15. Damyanovich, A.Z., Staples, J.R., Chan, A.D., Marshall, K.W. (1999) Comparative study of normal and osteoarthritic canine synovial fluid using 500 MH 1H magnetic resonance spectroscopy. J. Orthop. Res. 17: 223-231.
16. Tetlow, L.C., Woolley, D.E. (2002) Histamine stimulates matrix metalloproteinase-3 and-13 production by human articular chondrocytes in vitro. Ann. Rheum. Dis. 61: 737-740.
17. Tetlow, L.C., Woolley, D.E. (2003) Histamine stimulates the proliferation of human articular chondrocytes in vitro and is expressed by chondrocytes in osteoarthritic cartilage. Ann. Rheum. Dis. 62: 991-994.
18. Salway, J.G. (2004) Metabolism at a glance. Blackwell Publishing, Oxford.
19. Michal, G. (1999) Biochemical pathways: an atlas of biochemistry and molecular biology. Spektrum Akademischer Verlag, Heidelberg.

6

Urinary metabolite profile as early biomarker for acute rejection after kidney transplantation

Abstract

Objective: Renal transplantation is the preferred treatment option for patients with (or approaching) end-stage renal failure. A frequent problem after such an operation is acute rejection of the transplanted organ due to an attack of the immune system against the allograft. This may lead to functional deterioration of the kidney and to a negative effect on long-term graft survival. Early diagnosis of acute rejection may prevent structural damage to the kidney and is supposed to be beneficial for long-term allograft survival. This study was set up to investigate whether a urinary metabolite profile can be identified that is associated with an acute rejection episode. Such a profile could also serve as an early biomarker for acute allograft rejection.

Methods: Urine samples of nineteen male participants who underwent a kidney transplantation were selected from a prospectively collected consecutive cohort transplanted at the Leiden University Medical Center. Nine patients with biopsy-confirmed acute rejection in the early posttransplantation period were identified, as well as ten control subjects who maintained proper function of their kidney transplant. Samples were measured by ^{1}H Nuclear Magnetic Resonance spectroscopy (NMR) with subsequent multivariate data analysis (MVDA).

Results: Urine composition was different for patients showing acute rejection as compared to control subjects at time of biopsy ($P<0.001$) and also five days prior to the biopsy ($P<0.001$). A urinary metabolite profile was determined which was associated with the occurrence of an acute rejection episode. Five days prior to acute rejection, patients already showed an aberrant metabolite profile.

Conclusion: This study shows that a urinary metabolite profile may serve as a novel early biomarker that is able to predict the occurrence of an acute rejection episode in renal transplant recipients.

In preparation for publication.

Introduction

The kidneys allow the excretion of waste products from the body, regulate the volume of extracellular fluid and the balance of electrolytes [1]. *Acute* renal failure is characterized by a sudden decrease in renal function due to injury, disease or toxins. When remedies are used in time, kidney functions may restore. In case of *chronic* renal failure, the functions of the kidney are slowly getting worse. With good medication however, it is possible to control the process of deterioration [1-2]. In *end-stage* renal disease, chronic renal failure progresses to a point at which the kidneys work less than 10% of their capacity [3]. At this point the kidneys are no longer fulfilling their role of removing waste and excess fluids from the body. Toxins start to build up in the blood causing progressive complaints and potentially life-threatening complications such as hyperkalaemia. Currently there is no cure for end-stage renal disease. This condition is fatal to the patient unless dialysis on a regular basis or a kidney transplant is performed [2].

Renal transplantation is the preferred treatment of patients with, or those approaching end-stage rnal failure [4-5]. Allograft rejection remains an important problem after renal transplantation. In case of an acute rejection episode, a transplant recipient's immune system attacks the transplanted organ or tissue. Therefore, kidney transplant patients require life-long treatment with immunosuppressive drugs that suppress the alloimmune response [4]. With the current regimen, the overall incidence of early acute rejection still about 20-30% of kidney transplant recipients [5]. Acute rejection occurs most often within three months after transplantation but it can occur at any time after transplantation. Acute rejection is predominantly mediated by T-cells that infiltrate the graft and cause tissue destruction [5]. Treatment with high-dose steroids, T cell antibodies or a change in immunosuppressant drug can in the large majority of cases effectively control the rejection process. The functional response of an acute rejection episode to therapy has shown to be a strong impact factor for long-term graft survival after kidney transplantation [6-7]. Rejection episodes that do not affect renal function had no impact on graft survival whereas acute rejections in which baseline function was not restored were shown to have a profound impact on renal survival [7].

Currently, serum creatinine is used to monitor allograft function. Creatinine is a compound that is completely filtered from the blood by the glomerulus and also

secreted by the renal tubular epithelial cells. Hence its clearance is a good estimate of the glomerular filtration rate [1-2]. Although this is an easy and relatively inexpensive assay, it is not very sensitive and the blood creatinine level does not rise until significant injury to the kidney has occurred. Currently no biomarkers exist that can detect an acute rejection episode in a reliable and sensitive manner before the creatinine levels rise. Episodes of acute rejection are usually diagnosed by a kidney biopsy, which is performed at the time of graft dysfunction according to sequential serum creatinine levels. Since biopsy cannot be done frequently, it is not a very satisfactory method for early detection of acute rejection, i.e. before loss of function occurs. A prognostic biomarker for acute rejection will facilitate the appropriate and timely treatment of patients in order to prevent irreversible loss of renal allograft function. This is likely to have a beneficial effect on long-term graft survival. Next to that, the discovery and development of new therapies would greatly benefit from an early biomarker for acute rejection.

Metabolomics is the untargeted profiling of metabolites in biological samples [8-10]. This technology, together with approaches such as transcriptomics and proteomics [11], provides more insight regarding the pathogenesis of diseases. Developments in the field of metabolomics now provide the tools to investigate pathways and molecules involved in allograft rejection and to identify metabolite profiles that contain potential biomarkers. The present study was initiated from the assumption that a biomarker profile, which reflects acute kidney rejection, may be detected in urine. Using ^{1}H Nuclear Magnetic Resonance spectroscopy (NMR) and subsequent multivariate data analysis (MVDA) we investigated whether acute rejection after kidney transplantation is preceded by an aberrant urine metabolite profile that can be used to predict this occurrence in time.

Materials and Methods

Study population and sample selection

Early morning urine samples of nineteen male participants who underwent a kidney transplant operation were selected from a prospective study at the Leiden University Medical Centre (Leiden, The Netherlands). Subjects received a maintenance immunosuppressive regimen including prednisone, ciclosporin micro emulsion and mycophenolate mofetil. Patients with delayed graft function after

transplantation were excluded from the current study. Characteristics of patients their donor and the transplanted kidney were known (degree of mismatches for HLA antigens, ischemia time, sex, age). Daily serum creatinine values were collected to monitor kidney function. When creatinine values on two subsequent days were raised more than 10% over their baseline values, a biopsy was carried out after exclusion of potential pre- and post renal causes for graft dysfunction. Nine patients with a biopsy-confirmed acute rejection episode in the early posttransplantation period, as well as ten subjects with stable renal allograft function were selected. Urine samples collected at the time of rejection and at five days prior to the event were identified and samples collected within at comparable timeframe were chosen from the control subjects. The collected urine samples were centrifuged at 3000 rpm for 10 minutes to remove debris, and stored at –80°C until analyses.

NMR analysis of urine samples

Prior to NMR spectroscopic analysis, 1 mL urine samples were lyophilized and reconstituted in 1 mL sodium phosphate buffer (0.1 mmol/L, pH 7.4, made up with D_2O), to minimize spectral variance arising from differences in urinary pH. Sodium trimethylsilyl-[2,2,3,3,-$^2H_4$]-1-propionate (TMSP; 0.1 mmol/L) was added as an internal standard. NMR measurements were carried out in random order and in triplicate in a fully automated manner on a 600 MHz spectrometer (Avance, Bruker BioSpin GmbH, Rheinstetten, Germany), using a proton NMR set-up operating at a temperature of 300K. For each sample, 128 free induction decays (FID) were collected. Each FID was induced using a 45-degree pulse, an acquisition time of 2.73 s and a relaxation delay of 2 s. The FIDs were collected as 64K data points with a spectral width of 12.000 Hz. The spectra were processed using the standard Bruker software. An exponential window function with a line broadening of 0.3 Hz and a manual baseline correction were applied to all spectra. After referring to the internal NMR reference (TMSP δ= 0.0), line listings were prepared with the standard Bruker NMR software. To obtain these listings all lines in the spectra above a threshold corresponding to about three times the signal-to-noise ratio were collected and converted to a data file suitable for multivariate data analysis applications. The NMR data file was imported into Winlin (V2.1, TNO, The

Netherlands). Minor variations from comparable signals in different NMR spectra were adjusted and lines were fitted without loss of resolution, after which MVDA was carried out.

Multivariate data analysis

Data was centered and scaled to unit variance, upon which principal component discriminant analysis (PCDA) was performed using the Winlin software package. Health status (controls versus patients) was used as a priori knowledge for discrimination in PCDA. The NMR data set was randomly divided into a training data set and a test data set. The PCDA models were built upon the training data set. Subsequently the test data set was used to test the reliability of the training model. Predictions were in agreement with the actual groupings (patients versus controls).

Scores were plotted against the discriminant axis. The unpaired T-test was performed to evaluate the statistical significance of the difference between the PCDA scores of the two groups (Excel Office 2003, Microsoft Corporation, USA). The difference between scores was correlated to the original NMR features in the spectra. The resulting metabolite profiles provided insight into the type of metabolites responsible for the disparity.

Results

Description of sample

The mean characteristics of the group with patients showing acute rejection and the group with control participants are given in Table 1. No consistent statistically significant differences were observed except for the cold ischemia time ($P<0.05$) and creatinine clearance at three months ($P<0.05$).

Urinary metabolite profile of patients versus controls at time of rejection

The underlying hypothesis of the present study is that a metabolite profile could be hidden in urine of patients who show acute kidney rejection after transplantation. This profile should be specific for this process. Using NMR with subsequent MVDA, such abnormal urinary composition may be quantified.

When PCDA was applied on the NMR spectra of urine samples taken at the time of

biopsy, a difference showed up between scores of controls and patients (Figure 1A; P<0.001). Roughly, the scores of the controls are above zero, whereas the scores of the patients are below zero. The scores of two patients show overlap with the scores of the controls.

	patients with rejection	patients without rejection	P value
Female acceptor (%)	25	56	0.35
Age years (median)	53	39	0.15
First transplant (%)	62.5	77.8	0.60
Mismatches			
HLA-A zero	12.5	55.6	0.15
HLA-B zero	12.5	44.4	0.30
HLA-DR zero	37.5	55.6	0.65
CMV status recipient negative (%)	25	33.3	1.00
CMV status donor negative (%)	50	55.6	1.00
Female donor (%)	37.5	33.3	1.00
Panel Reactive Abs (highest %)	17.86	34.67	0.35
Warm ischemia time (min)	32.5	30	0.75
Cold ischemia time (h)	27.22	17.38	0.05
creatinin clearance at 3 month (median; mL/min)	47.1	79,6	0.02
creatinin clearance at 6 month (median; mL/min)	60.2	82,1	0.2

Table 1. *Mean characteristics of the group with patients showing acute rejection and the group with control participants. No consistent statistically significant differences were observed except for the cold ischemia time (P<0.05) and creatinine clearance at three months (P<0.05).*

A metabolite profile could be constructed that marks the difference between the groups, and which is likely to reflect acute kidney rejection (Figure 1B). In this fingerprint, NMR signals of urinary metabolites are visualized which increased or decreased in association with acute graft rejection at time of biopsy. NMR signals that were strongly associated with the event, were δ 2.18, 3.22, 7.46 ppm in positive direction. These signals represent trimethylamine N-oxide (TMAO), amongst other not identified signals. Signals at δ 3.18, 3.46, 3.58, 3.66, 3.78, 3.82, 3.90 and 4.06 ppm in negative direction represent, amongst others, metabolites like glucose and creatinine.

Metabolite profile five days before rejection

The question is now whether the observed differences in the metabolite profile at the time of acute rejection already can be detected at a time point when creatinine levels have not raised yet. When PCDA was carried out on the NMR data of urine samples taken five days prior to biopsy, a difference between scores of controls and patients showing acute rejection later on showed up (Figure 2A; $P<0.001$). Scores of the controls are above zero, whereas scores of the patients are below zero.

Thus, five days before acute rejection is assessed with biopsy, MVDA showed a visible separation between patients and controls. This difference is based on the concentrations of metabolites which were characteristic for each group. The metabolite profile for kidney rejection at five days prior to biospy showed similarities but was in part different from the one obtained at the day of rejection (Figure 2B). NMR signals that were strongly associated with rejection, were δ 2.26, 2.78, 3.10, 3.22 and 3.66 ppm in positive direction. Signals in these regions represent, amongst other not identified signals, metabolites like TMAO and dimethylglycine. Signals at δ 3.02, 3.26, 3.42, 3.50, 3.78, 3.82, 3.9 and 4.66 ppm in negative direction represent, amongst others, glucose that lowers in patients that showed rejection according to biopsy five days later on.

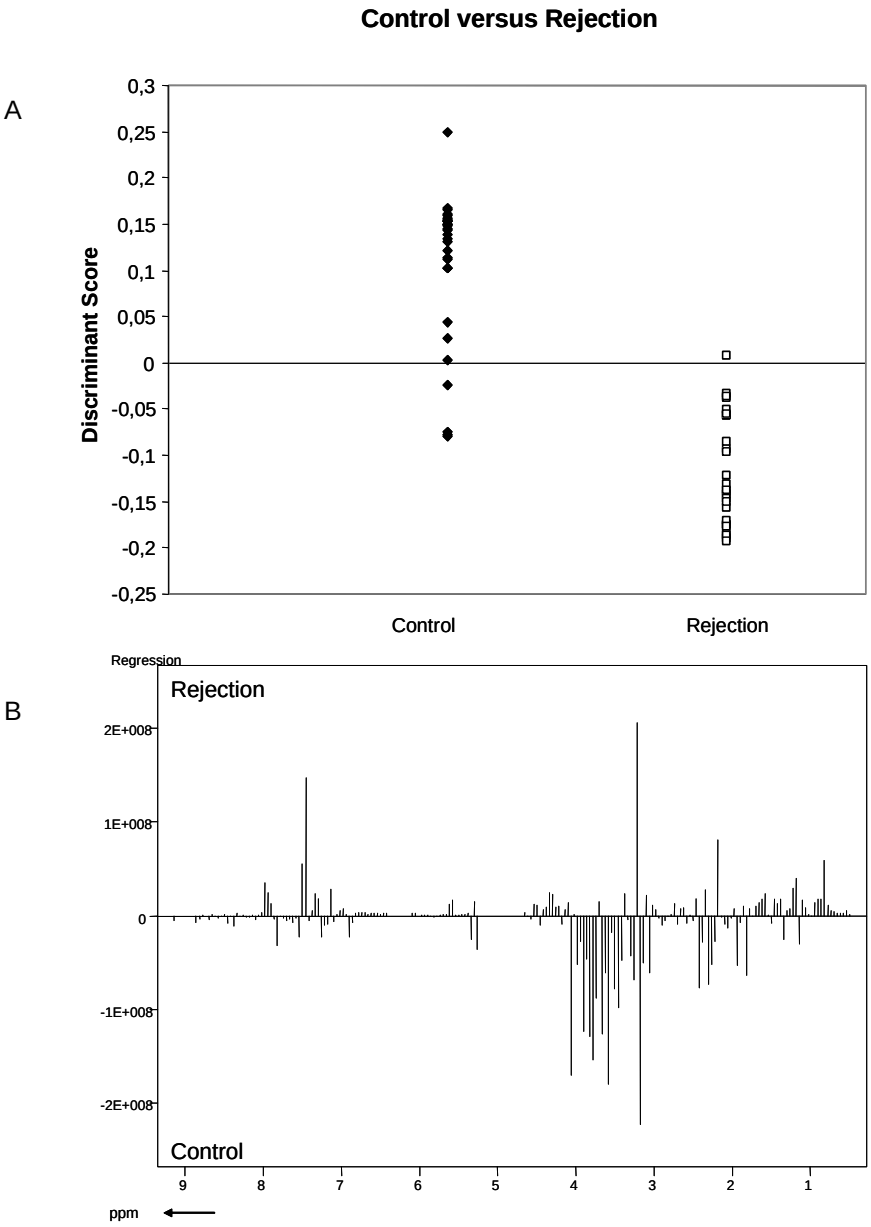

Figure 1. (A) *A PCDA score plot of urinary NMR spectra of patients showing acute rejection upon kidney transplantation versus control subjects that presented no rejection (the points represent the complete urinary NMR spectra of the subjects). Urine was taken at time of biopsy. The prevalence of rejection is reflected in the urinary composition, as there is a clear difference between the scores of the two groups on the discriminant axis (P<0.001).* **(B)** *Metabolite profile showing differences between urinary NMR spectra of controls versus patients showing acute rejection. Peaks (representing NMR signals) in the positive direction indicate metabolites that are more abundant in urine of patients showing acute rejection than in urine of the controls. Consequently, metabolites that are more abundant in urine of these controls are presented as peaks in the negative direction.*

Discussion

A biomarker that can early diagnose acute kidney rejection is important as it will clear the way for preventive therapy. This will allow for less functional damage to the kidney and may eventually result in prolonged long-term renal allograft survival. Currently no biomarkers exist that can detect an acute rejection episode in a reliable and sensitive manner and in an early stage. Using a metabolomics approach, we investigated whether a metabolite profile could be identified that is predictive for an acute rejection episode in renal allograft recipients. With NMR and subsequent MVDA we analyzed urine samples of nineteen male participants who underwent a kidney transplant operation. Nine of these patients were faced with acute rejection in the early posttransplantation months according to biopsy, whereas ten subjects showed no complications and served as controls.

We were able to discriminate patients who showed acute graft rejection from control subjects. The distinction between these groups was based on small differences between urinary metabolite levels of the respective groups as measured by NMR. We found a urinary metabolite profile that was strongly associated with acute rejection. We also showed that, already five days before rejection was assessed according to biopsy, discrimination between urine contents of patients and controls was feasible. A metabolite profile was presented that may predict the subsequent acute rejection episode at least five days prior to the event. According to our results, the creatinine level in urine at biopsy was decreased in patients showing acute rejection. This finding was in line with clinical chemistry, in which patients with acute rejection showed a significant lower creatinine clearance as compared to the control subjects. An elevated blood creatinine level and reduced urinary creatinine excretion is a well-known characteristic in patients showing acute rejection [1-2].

However, the metabolite profile was not dominated by creatinine solely. TMAO was identified as a metabolite highly associated with acute graft rejection at biopsy as well as five days prior to the event. The metabolite TMAO has been associated with the occurrence of delayed graft function, and in particular to the cold ischemia time, the time a kidney is preserved on ice before transplantation [12-15]. Damage to the renal medulla due to cold preservation may cause the release in urine of TMAO and dimethylglycine [16-17]. In our study, the cold ischemia time was

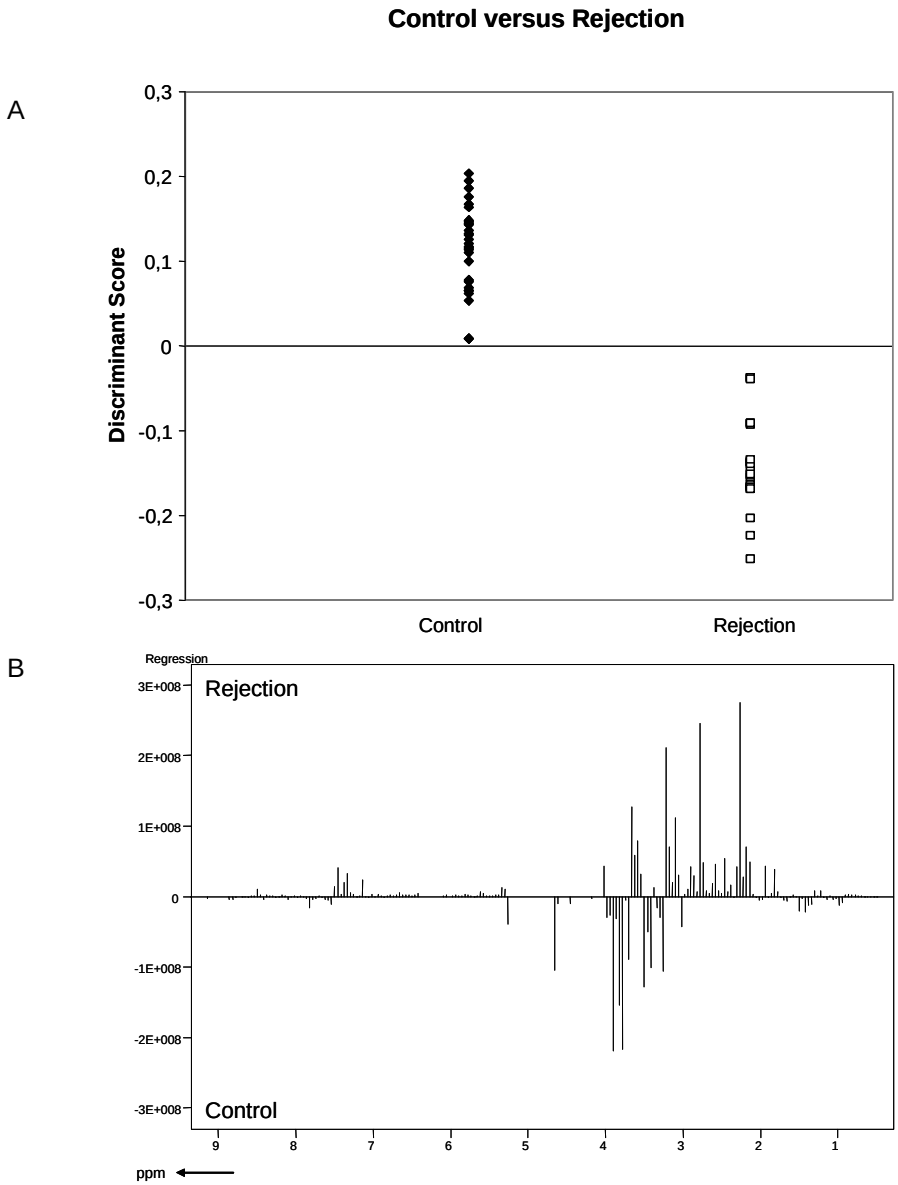

Figure 2. (A) *A PCDA score plot of urinary NMR spectra of patients developing acute rejection versus controls that presented no rejection (the points represent the complete urinary NMR spectra of the subjects). Urine was taken five days before the actual rejection found place according to biopsy. The prevalence of acute rejection is already reflected in the urinary composition five days before biopsy, as there is a clear difference between the scores of the two groups on the discriminant axis at this time point (P<0.001).* **(B)** *Metabolic profile showing differences between urinary NMR spectra of controls versus patients showing acute rejection. Peaks (representing NMR signals) in the positive direction indicate metabolites that are more abundant in urine of patients showing acute rejection than in urine of the controls. Consequently, metabolites that are more abundant in urine of these controls are presented as peaks in the negative direction. This biomarker profile predicts acute rejection five days prior to biopsy.*

significantly higher for the patients showing acute rejection, which is in agreement with an elevated amount of TMAO in urine of these patients, but none of the patient had delayed graft function.

The metabolite profile also pointed at lowered amounts of glucose in urine, which may be explained by the fact that renal perfusion is hampered by an acute rejection episode and consequently the glomeruli function less properly.

This study shows that use of NMR with subsequent MVDA enables to measure biological fluids without pretreatment in a quick manner. It provides an unbiased and broad overview of many important metabolites present in biological samples and can thus be beneficial in the search for new biomarkers. However, many of the NMR signals in the metabolic profiles that were measured in our study could not be assigned to metabolites. Nominating compounds in NMR spectra of biological fluids is a common problem [10]. The choice for assigning an NMR signal to a metabolite may be influenced by existing knowledge about metabolic pathways that is already known. However, there are still many metabolites and pathways in the body we do not know. The interpretation of signals in an NMR spectrum of a biological fluid is hampered by an incomplete knowledge of metabolism [18]. More efforts should therefore be made to identify and validate metabolites. Seen in this light, for example liquid chromatography-mass spectrometry (LC-MS) and gas chromatography-mass spectrometry (GC-MS) are valuable complementary analytical tools.

In conclusion, the results of this study give evidence of a metabolite profile associated with acute kidney rejection at time of biopsy. We also identified an aberrant metabolite profile five days prior to the clinical assessment of an acute rejection episode. The obtained metabolite profiles, especially the metabolites upon which the identified profile is based, could play an important role as potential early biomarker for acute graft rejection. Findings about involved metabolites are in line with clinical chemistry and other studies that reported effects of metabolites related to the ischemia/reperfusion injury. The metabolites in the metabolic fingerprints could not all be nominated, but our results justify further research in this area. The identified early metabolite profile may contain potential early biomarkers that will provide a tool for physicians to diagnose acute rejection in time such that

preventive and custom-made therapy can be applied. Eventually, this should lead to less functional renal damage and an increase the longevity of renal allografts.

References

1. Thabrew, I., Ayling, R.M. (2001) Biochemistry for clinical medicine. Greenwich Medical Media Ltd., London.
2. Mayne, P.D. (2001) Clinical chemistry in diagnosis and treatment. Arnold, London.
3. Cotran, R.S., Kumar, V., Robbins, S.L. (1989) Pathologic basis of disease. W.B. Saunders Company, Philadelphia.
4. Morris, P.J. (1997) Renal transplantation: a quarter century of achievement. Semin. Nephrol. 17: 188-195.
5. Salmela, K., Isoniemi, H. (1997) Kidney transplantation. Ann. Chir. Gynaecol. 86: 94-100.
6. De Fijter, J.W., Mallat, M.J.K., Doxiadis, I.I.N., Ringers, J., Rosendaal, F.R., Claas, F.H.J., Paul, L.C. (2001) Increased immunogenicity and cause of graft loss of old donor kidneys. J. Am. Soc. Nephrol. 12: 1538-1546.
7. Meier-Kriesche, H.U., Schold, J.D., Srinivas, T.R., Kaplan, B. (2004) Lack of improvement in renal allograft survival despite a marked decrease in acute rejection rates over the most recent era. Am. J. Transplant. 4: 378-383.
8. Fiehn O. (2002) Metabolomics, the link between genotypes and phenotypes. Plant Molecular Biology 48:155-171.
9. Nicholson J.K., Lindon J.C., Holmes E. (1999) ''Metabonomics'': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica 29: 1181-1189.
10. Lamers, R.A.N., DeGroot, J., Spies-Faber, E.J., Jellema, R.H., Kraus, V.B., Verzijl, N., TeKoppele, J.M., Spijksma, G., Vogels, J.T.W.E., van der Greef, J., van Nesselrooij, J.H.J. (2003) Identification of disease and nutrient related metabolic fingerprints in osteoarthritic guinea pigs. J. Nutr. 133: 1776-1780.
11. Schaub, S., Rush, D., Wilkins, J., Gibson, I.W., Weiler, T., Sangster, K., Nicolle, L., Karpinski, M., Jeffery, J., Nickerson, P. (2004) Proteomic-based detection of urine proteins associated with acute renal allograft rejection. J. Am. Soc. Nephrol. 15: 219-227.
12. Le Moyec, L., Pruna, A., Eugene, M., Bedrossian, J., Idatte, J.M., Huneau, J.F., Tome, D. (1993) Proton nuclear magnetic resonance spectroscopy of urine and plasma in renal transplantation follow-up. Nephron 65: 433-439.
13. Foxall, P.J.D., Mellotte, G.J., Bending, M.R., Lindon, J.C., Nicholson, J.K. (1993) NMR spectroscopy as a novel approach to the monitoring of renal transplant function. Kidney Int. 43: 234-245.
14. Neild, G.H., Foxall, P.J.D., Lindon, J.C., Holmes, E.C., Nicholson, J.K. (1997) Uroscopy in the 21st century: high-field NMR spectroscopy. Nephrol. Dial. Transplant 12: 404-417.
15. Rush, D., Somorjai, R., Deslauriers, R., Shaw, A., Jeffery, J., Nickerson, P. (2000) Subclinical rejection – a potential surrogate marker for chronic rejection – may be diagnosed by protocol biopsy or urine spectroscopy. Ann. Transpl. 5: 44-49.
16. Hauet, T., Gibelin, H., Godart, C., Eugene, M. Carretier, M. (2000) Kidney retrieval conditions influence damage to renal medulla: evaluation by proton nuclear magnetic resonance (NMR) pectroscopy. Clin. Chem. Lab. Med. 38: 1085-1092.
17. Hauet, T., Baumert, H., Amor, I.B., Gibelin, H., Tallineau, C., Eugene, M., Tillement, J.P., Carretier, M. (2000) Pharmacological limitation of damage to renal medulla after cold storage and transplantation by trimetazidine. J. Pharmacol. Exp. Ther. 292: 254-260.
18. Watkins SM, German JB. (2002) Toward the implementation of metabolomic assessments of human health and nutrition. Curr. Opin. Biotechnol. 13: 512-516.

7

A pilot study to investigate effects of inulin on Caco-2 cells through in vitro metabolic fingerprinting

Abstract

Metabolic fingerprints are novel measurement tools to evaluate the biochemical status of a living organism by using ^{1}H Nuclear Magnetic Resonance spectroscopy (NMR) and multivariate data analysis (MVDA). In this way, a quick evaluation of changes in health or diseased state can be given, reflected in alterations of metabolic patterns. Normally, metabolic fingerprinting is based on *in vivo* studies. These studies are most times a labour-intensive and expensive manner of investigation. *In vitro* studies are not hampered by these disadvantages and form therefore an interesting alternative. In this research, results are presented of a pilot experiment in which metabolic fingerprinting was combined with an *in vitro* model. For this purpose, differentiated Caco-2 cells were exposed to inulin respectively its fermentative metabolites, both dissolved in culture medium. Cells were incubated for 0 or 48 hours. Cell fractions were analyzed by NMR with subsequently MVDA. It was shown that differences in treatment provided detectable variations in time of metabolic patterns of cell contents. Results indicated that glucose metabolism linked to glutamate was of major importance in the effects of inulin and its metabolites on Caco-2 cells under the conditions of our study. Metabolic fingerprinting in combination with an *in vitro* model appears to be a feasible method to visualize metabolic patterns of cell contents and provides an efficient procedure for generation of hypotheses about metabolic pathways involved. *In vitro* metabolic fingerprinting may in future be of great benefit for a better understanding of relations between nutrition and health.

Based on: Lamers, R.A.N., Wessels, E.C.H., van de Sandt, J.J.M., Venema, K., Schaafsma, G., van der Greef, J., van Nesselrooij, J.H.J. (2003) A pilot study to investigate effects of inulin on Caco-2 cells through in vitro metabolic fingerprinting. J. Nutr., 133, 3080-3084.

Introduction

Fructans are polymers of fructose. Inulin and oligofructose belong to this class of carbohydrates. Inulin is found in many plants and vegetables (e.g. chicory and Jerusalem artichoke). Chicory is by far the most commonly used source for the industry to obtain inulin as a commercial product. As an ingredient of foods, inulin functions, amongst others, as a fat and sugar replacement, mouth feel and texture improvement and dietary fibre. Inulin is counted as a prebiotic, since it is not susceptible to digestion and hydrolyzes by endogenous enzymes. By reaching the colon it will be fermented by the microbiota and selectively stimulate the growth of bifidobacteria [1]. Predominance of bifidobacteria in the large intestine is supposed to be beneficial for maintaining good health [1-2]. Fermentation products of inulin are short-chain fatty acids (SCFA; acetate, propionate and butyrate), lactate and gasses [2]. Functional effects of inulin on humans and experimental animals are relieved constipation, lowering blood glucose levels, improvement of absorption of calcium, reducing fasting triglycerides and LDL cholesterol, inhibition of growth of various kinds of tumours [3].

The underlying metabolism, which causes the effects of inulin, remains indistinct and not yet fully understood [3]. This is a more common problem in nutritional research, where there is a shortage of knowledge of the relationships between health and disease and effects of nutrition on the latter. Fortunately, in the field of metabolite research great progress was made recently due to metabolic fingerprinting [4-7]. This technique utilizes ^{1}H Nuclear Magnetic Resonance spectroscopy (NMR) in combination with multivariate data analysis (MVDA) to analyze biological fluids.

NMR provides concurrent detection of all hydrogen containing molecules in a sample without pre-treatment. NMR can thus reveal chemical structures of metabolites in biological fluids and subsequently clarify metabolic pathways involved in nutrition and health [8]. Nevertheless, interpretation of NMR spectra obtained from biological fluids is very complicated due to the enormous amount of spectral signals produced.

MVDA is known to be a powerful technique for the analysis of data sets with a large number of variables. For this reason, MVDA is particularly opportune to find significant spectral changes in NMR spectra: it enables to visualize spectral

patterns in NMR data, and thus metabolites, which correlate with e.g. treatment or disease [4-6].

In MVDA, unsupervised and supervised techniques can be used. Unsupervised methods such as principal component analysis (PCA) search for similarities and differences in data sets without foreknowledge. A large set of related variables (e.g. NMR signals) is converted to a smaller set of uncorrelated variables, which express maximum variation in the original variables. The new variables are called principal components (PC) and each of them depicts an axis in multidimensional space. The distance of an object (e.g. a complete NMR spectrum of a sample) to a PC is called a *score*. Scores are plotted in a score plot, with the PCs as axes. When scores are situated close to each other in a score plot, this implicates that the NMR spectra of the samples are similar. When the clustering of scores matches the samples that were controls, treated or diseased in the original study set-up, a connection can be linked up between affected NMR signals, and thus metabolites, and treatment or disease.

Calculation of the contribution of each original variable (e.g. a single NMR signal) to a PC yields a *loading*. When a loading is high, the original NMR signal adds greatly to the clustering of scores in the direction of the investigated PC. In a so-called factor spectrum or metabolic fingerprint, loadings are presented as lines. The location of the lines in a factor spectrum corresponds to the location of the variables in the original NMR spectra. The length of a line denotes the contribution of a variable to the grouping of scores in the investigated direction [4]. Thus, a high line in positive direction indicates an NMR signal that is strongly ascending for a particular group of scores.

Supervised methods such as partial least squares (PLS) and principal component discriminant analysis (PCDA) exploit supplemental information on the data set (e.g. biochemical, histopathological or clinical data) to identify and maximize similarities and differences between pre-defined groups [4-6]. In PCDA, the scores from PCA are used as a basis for linear discriminant analysis: discriminant analysis combines the PCs in such a way that differences between pre-defined groups are optimized [9].

Up to now, metabolic fingerprinting was mainly used in combination with *in vivo* studies. These studies are time-consuming, labour-intensive and, because of these

and other factors, expensive compared to *in vitro* studies. When metabolic fingerprinting could successfully be applied to *in vitro* studies, this may be a useful alternative for *in vivo* based metabolic fingerprinting. Nutritional society could take great advantage of this in future. *In vitro* metabolic fingerprinting may be a relatively inexpensive and quick way to fill the gap in the lack of evidence for effects of e.g. functional foods on health.

In the underlying research, a pilot *in vitro* experiment with metabolic fingerprinting was carried out. The suitability of *in vitro* metabolic fingerprinting was assessed by investigation of direct and indirect effects (after fermenting with the colonic microbiota) of inulin on Caco-2 cells [10].

Materials and methods

In vitro experiment

Caco-2 cells (designation HTB 37) were obtained from the American Type Culture Collection (ATCC, Rockville, MD, USA). For 500 mL culture medium, 440 mL Dulbecco's modified Eagle medium (DMEM; cat.no. 42430) was used supplemented with 50 mL heat-inactivated foetal calf serum, 5 mL non-essential amino acids (10 mmol/L), 5 mL L-glutamine (200 mmol/L) and 0.5 mL gentamicin (50 mg/mL). Cell cultures were grown in this medium and maintained at 37^{o}C in 95% air and 5% CO_2 (v/v; Sanyo incubator). Near confluent Caco-2 cell cultures were harvested by trypsinisation with 3 mL trypsin solution (25 g/L) in 147 mL phosphate buffered saline (PBS) and were resuspended in 10 mL culture medium and 5 times diluted. All chemicals were obtained from Gibco (Breda, the Netherlands).

For the experiment, cells were seeded in Transwell^{TM} inserts in 12-well plates (1 mL of cell suspension with 1.5 mL DMEM per well). The medium was changed every 2-3 days. Cells became confluent after about 4 days, at which time differentiation could begin. After complete differentiation, samples (wells) were fed with various media. Four samples were treated with 1.5 mL DMEM for 0 hours and four samples for 48 hours. A 10 times dilution of a saturated solution of 1.5 g Frutafit® EXL (Sensus, Roosendaal, The Netherlands) and 30 ml DMEM was used to treat four samples with 1.5 mL for 0 hours whereas four samples were treated with the same solution for 48 hours. Besides, four samples were treated with 1.5

mL metabolized inulin (run in TIM-2 feeding, 10 times diluted with DMEM) for 0 hours and four samples for 48 hours. Another four samples were treated with 1.5 mL TIM-2 feeding after passing the TIM-2 model (10 times diluted with DMEM) for 0 hours and four samples for 48 hours. The 0 hour samples were collected directly after start exposure

The **T**NO ***in vitro*** **m**odel of the large intestine (nick-named TIM-2) simulates the physiological parameters in the large intestine (or colon), such as pH, temperature and an active microbiota similar in composition and activity to that in the human colon [11, 12]. Fermentation in the proximal colon was mimicked in this *in vitro* model. To the control TIM-2 medium, the test compound in question (i.e. inulin) was added. This mixture was added to the TIM-2 system, giving rise of metabolized inulin. The temperature was kept at 37 ^{o}C, while the pH was kept at 5.8. The model was flushed with gaseous nitrogen to allow growth of an active anaerobic, complex microbiota of human origin. The model was inoculated with a microbiota from human faecal material. Inulin was dosed at 10 gram per day in doses of approximately 104 mg per 15 minutes. The contents were mized by peristaltic movements. Microbial metabolites were removed from the model by a dialysis system running through the model. This prevented inhibition of the activity of the microbiota by accumulation of microbial metabolites. For more details on the *in vitro* model, please refer to Minekus *et al* (11) and Venema *et al.* (12).

All dilutions were centrifuged at 2500 x *g* for 10 min (Megafuge 2.0 RS, Heraeus, Germany). The dilutions with control TIM-2 medium and metabolized inulin were passed through a 0.2 µm filter before exposing them to the cultured cells.

Solutions with the respective test compounds were removed at 0 or 48 hours, depending on time of exposure. Cells were washed then with PBS of 37°C and dissolved in 1 mL of a solution of methanol (Sigma-Aldrich, Zwijndrecht, The Netherlands) in demineralized water (7.5 mol/L). Samples were sonicated for 10 seconds at 20 µm to lyse the cells, using a MSE Ultrasonic disintegrator sonifier (Beun-de Ronde BV, Amsterdam, The Netherlands). After that they were centrifuged (Eppendorf, Germany) at 13000 min^{-1} for 5 minutes, yielding samples of cell contents which were stored at -40°C until NMR analysis.

NMR analysis of in vitro medium and cell samples

Prior to NMR spectroscopic analysis, the culture medium was removed from the cell samples. Cells were evaporated to dryness under a stream of nitrogen gas. The samples were dissolved in 1 mL of sodium phosphate buffer (0,1 mol/L, pH 6.0, made up with D_2O). Sodium trimethylsilyl-[2,2,3,3,4-$^2H_4$]-1-propionate (TMSP; 0.05 mmol/L) was added as internal standard.

NMR measurements were carried out in random order and in triplicate in a fully automated manner on a 600 MHz spectrometer (Avance, Bruker BioSpin GmbH, Rheinstetten, Germany), using a proton NMR set-up operating at a temperature of 300K. For each sample, 256 free induction decays (FID) were collected. Each FID was induced using a 45-degree pulse, an acquisition time of 4.10 s and a relaxation delay of 2 s. The FIDs were collected as 64K data points with a spectral width of 12.000 Hz. The spectra were processed using the standard Bruker software. An exponential window function with a line broadening of 0.5 Hz and a manual baseline correction were applied to all spectra. After referring to the internal NMR reference (TMSP δ= 0.0), line listings were prepared with the standard Bruker NMR software. To obtain these listings all lines in the spectra above a threshold corresponding to about three times the signal-to-noise ratio were collected and converted to a data file suitable for multivariate data analysis applications.

NMR data preprocessing and multivariate data analysis

The NMR data reduction file was imported into Winlin (V1.11, TNO, The Netherlands). Minor variations from comparable signals in different NMR spectra were adjusted and lines were fitted without loss of resolution [13]. To correct for sample dilution, the data were auto-scaled so that small and large signals contributed similarly to the final study result. Principal component discriminant analysis (PCDA) was performed, with treatment and time as additional information respectively. For PCDA, the NMR data set was randomly divided into a training data set and a test data set. The PCDA models were built upon the training data set. Subsequently the test data set was used to test the reliability of the training model. Predictions were in agreement with the actual groupings.

The resulting discriminants were quantified for each of the NMR spectra and the first discriminant (D1) was plotted *versus* the second discriminant (D2) to visualize

clustering. Factor spectra were used to correlate the position of clusters in the score plot to the original NMR signals in the spectra. The metabolic fingerprints were prepared in directions of maximum separation of one cluster *versus* another cluster, to provide insight into the type of metabolites responsible for the separation between clusters [4]. Metabolites were assigned from the metabolic fingerprints using an in-house database with NMR spectra.

Results

This research was initiated from the assumption that inulin and its metabolites will affect Caco-2 cells. Consequently, due to a change in metabolite levels in these cells an aberrant metabolite profile would normally arise. This alteration in cell contents could then be visualized by using metabolic fingerprints.

The effect of inulin on cell contents

If there is an effect of inulin on cell contents, this information will be contained in the NMR spectra of these cells. Principal component discriminant analysis (PCDA) was used to visualize differences between NMR spectra obtained from contents of cells exposed to DMEM solely *versus* contents of cells exposed to DMEM with inulin, at two points of time (Figure 1A). The first (D1) and second (D2) discriminant explained 36% respectively 33% of the variance. It is clear that the contents of Caco-2 cells treated with inulin in DMEM are positioned in about the same location in the score plot at 0 hours as the contents of cells treated with DMEM solely. However, after 48 hours differences were clearly visible between contents of cells exposed to DMEM with, respectively without, inulin. When looking at the metabolic fingerprint (Figure 1B), it can be noticed that the differences were due to changes in regions 1-4.5 ppm and 7-8.5 ppm. Metabolites that could be assigned to these signals using the in-house database with NMR spectra, were leucine (δ 0.96, 1.71), isoleucine (δ 0.94, 1.01), valine (δ 0.99, 1.04), alanine (δ 1.48, 3.79), α- and β-glucose (δ 3.47, 3.49, 3.53, 3.71, 3.72, 3.74, 3.84, 3.9, 4.64, 5.24), phenylalanine (δ 7.33, 7.38, 7.43), tyrosine (δ 6.91, 7.2) and glutamate (δ 2.1, 2.35, 3.77).

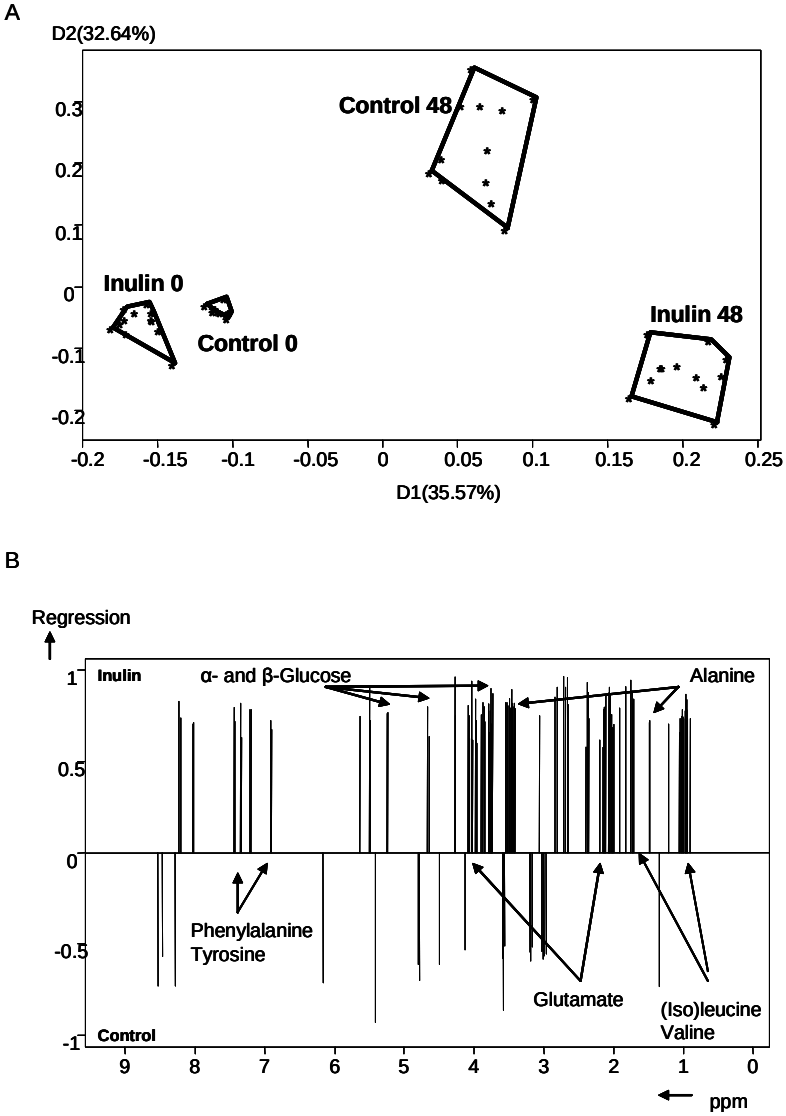

Figure 1. *Principal component discriminant analysis (PCDA) score plot and factor spectrum visualizing effects of inulin on Caco-2 cell contents.* **(A)** *PCDA score plot of NMR spectra of cells treated with control medium and cells treated with medium containing inulin at 0 versus 48 hours (the points represent the complete NMR spectra of Caco-2 cell contents; each cluster contains four samples (wells) measured in triplicate). At 0 hours there is no difference between NMR spectra, and thus cell contents, of the different treated cells, whereas a clear difference has developed between the two groups at 48 hours. This indicates an effect of inulin on cells.* **(B)** *Factor spectrum, or metabolic fingerprint, of NMR spectra of cells treated with inulin containing medium versus cells treated with control medium at 48 hours. Peaks (representing NMR signals) in the positive direction indicate metabolites that are more abundant in cells treated with inulin (at 48 hours) compared to cells treated with control medium (at 48 hours). Consequently, metabolites that are more abundant in cells treated with control medium are presented as peaks in the negative direction.*

The effect of fermented inulin on cell contents

When PCDA was performed on the NMR spectra of contents of Caco-2 cells treated with fermented inulin *versus* contents of cells treated with the accompanying control medium, the first (D1) and second (D2) discriminant explained 41% respectively 35% of the variance. The score plot reveals that both groups have the same location at 0 hours (Figure 2A). However, after 48 hours of exposure, a difference could be observed between contents of Caco-2 cells exposed to control medium solely and contents of cells exposed to medium with fermented inulin. This was also reflected in the metabolic fingerprint at 48 hours (Figure 2B). From the fingerprint, it can be noticed that differences in regions around 1.5, 2, 3, 4 and 8 ppm contributed heavily to this effect. According to the NMR database, metabolites that, amongst others, belong to these signals are lactate (δ 1.33, 4.12), alanine (δ 1.48, 3.78), proline (δ 2.01, 2.07, 4.14), succinate, 2-oxoglutarate (δ 2.44, 3), nicotinate and nicotinamide (δ 7.97, 8.20, 8.28, 8.52).

Discussion

In the present study, a first step was taken in combining metabolic fingerprinting with *in vitro* models. From a pilot experiment with Caco-2 cells exposed to inulin, it became clear that the combination of *in vitro* models, NMR and MVDA may develop into a promising way to evaluate the biochemical status of cells.

Differences could clearly be revealed between contents of cells exposed to DMEM with, respectively without, inulin after 48 hours. At 0 hours this difference was not visible. Hence, their NMR spectra, and thus metabolite levels in the cells, were significantly different. This indicates that inulin had an effect on contents of Caco-2 cells in time. The score plot of NMR spectra from contents of Caco-2 cells treated with fermented inulin *versus* contents of cells treated with the accompanying control medium pointed out that a difference could be observed between the two groups at 48 hours.

Metabolic fingerprints were identified that reflected effects of inulin as well as its metabolites on Caco-2 cell contents. The height of a line in a metabolic fingerprint reflects the importance of an NMR signal to investigated clusters in a score plot. Metabolic fingerprints thus provided information about metabolites, which were elevated or lowered in one cluster compared to another cluster.

Effects of inulin itself on cell contents have not been investigated widely yet, because it is presumed that it is fermented near to completion in the gut. However, effects may occur in the small intestine, where inulin is not digested neither fermented. According to the metabolic fingerprint derived from our study, inulin itself seemed to influence metabolism in Caco-2 cells. The amount of glucose, together with glutamate contents, ascended after 48 hours due to treatment with inulin compared to the controls. This could point at activated gluconeogenesis in the Caco-2 cells under the conditions of our experiments. It is known that glucose metabolism is altered by glutamine via the citrate cycle [14]. However, the liver and kidney are considered as the only organs capable of gluconeogenesis although fresh concepts on glutamine and glucose metabolism have revealed that release of glucose might also find place in the small intestine when in fasting state [15]. Since Caco-2 cells develop some properties
of the small intestinal epithelium during differentiation, this hypothesis merits further investigation using isotopes to measure rates of gluconeogenesis.

Another explanation for the rise in level of glucose due to inulin treatment could be the fact that some inulin, or a breakdown product, was taken up by the cells. However, it is unlikely that inulin passed through the membrane into the cell by facilitar diffusion. For this, inulin is too large of a molecule unless pinocytosis had a share in it. When the latter is the case, fructose would be converted to glucose in the cell. Contents of glucose would grow then and thus also become available for glycolysis to form phosphoenol pyruvate. This compound is the origin for production of tyrosine via the shikimate pathway. A rised level of this compound in the metabolic fingerprint supports that this biochemical route was affected. Furthermore, phosphoenolpyruvate can be converted to pyruvate. Subsequently, pyruvate is oxidized to acetyl-CoA that enters the citrate cycle. When oxidation of pyruvate is not complete, alanine and lactate may be formed. These compounds showed elevated levels in the metabolic fingerprint. Glutamate production (ascending in the metabolic fingerprint) takes place from the citrate cycle by transamination between 2-oxoglutarate and amino acids to be catabolized [16].

According to the metabolic fingerprint, levels of phenylalanine, valine, leucine and isoleucine were elevated due to inulin compared to the controls. These compounds are essential amino acids. It therefore seems unlikely that Caco-2 cells

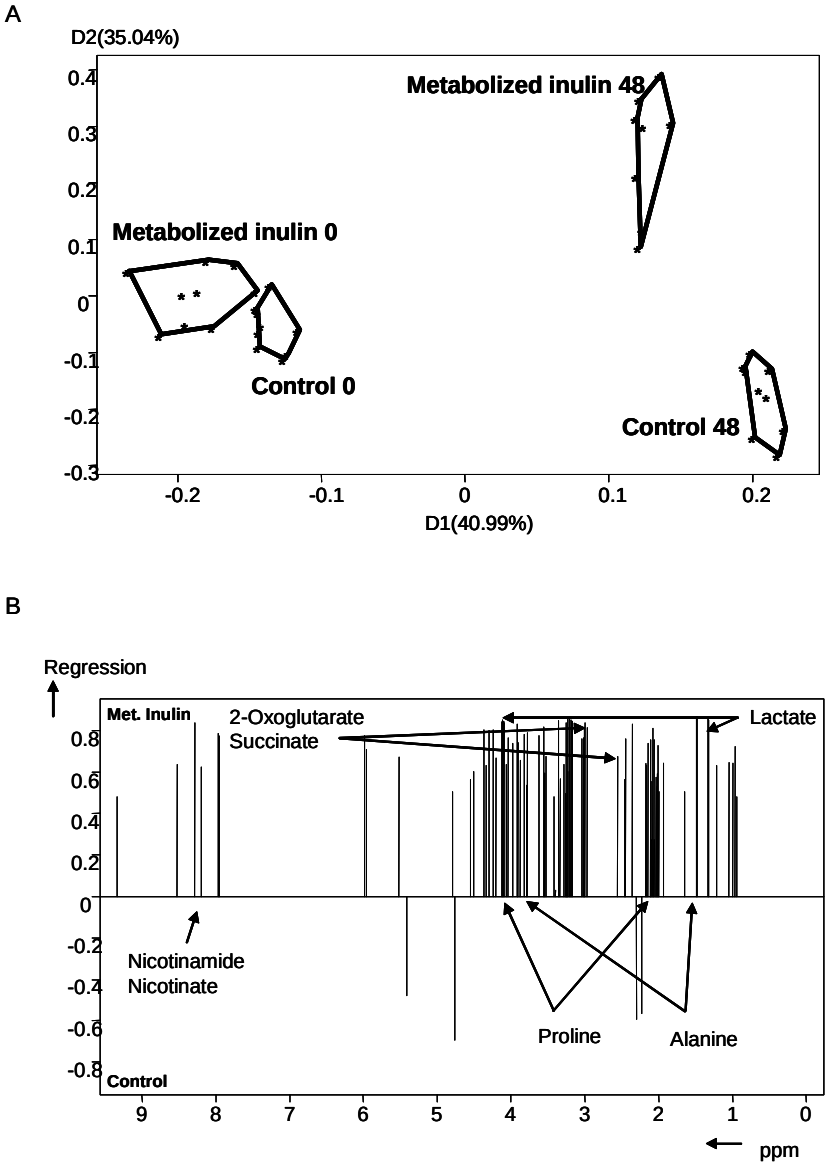

Figure 2. *Principal component discriminant analysis (PCDA) score plot and factor spectrum visualizing effects of metabolized inulin on Caco-2 cell contents.* **(A)** *PCDA score plot of NMR spectra of cells treated with control medium versus medium containing fermented inulin at 0 and 48 hours (the points represent the complete NMR spectra of Caco-2 cell contents; each cluster contains four samples (wells) measured in triplicate). The difference between NMR spectra of cell contents at 48 hours is reflected by the clear separation into two groups at this time point whereas at 0 hours no difference is visible. This points at an effect of metabolized inulin on cells.* **(B)** *Factor spectrum, or metabolic fingerprint, of NMR spectra of Caco-2 cells treated with medium containing fermented inulin versus cells treated with medium solely at 48 hours. Peaks in the positive direction indicate metabolites that are more abundant in Caco-2 cells treated with fermented inulin than in Caco-2 cells treated with control medium. Consequently, metabolites that are more abundant in cells treated with control medium are presented as peaks in the negative direction.*

are able to synthesize these compounds but it may be hypothesized that protein synthesis and degradation was altered due to inulin.

After exposure to fermented inulin, metabolites related to the citrate cycle, like alanine, lactate, succinate, 2-oxoglutarate were most prominent. Besides, the amount of nicotinate and nicotinamide increased and glutamate production was enhanced, reflected in an elevated level of proline. These metabolites support the idea that glycolysis in Caco-2 cells seemed to be stimulated by fermented inulin compared to the controls. In Figure 3, altered metabolites and the metabolic pathways in which they are involved are depicted.

The metabolites mentioned in this article could quiet easily be nominated from the metabolic fingerprints by experience and using an in-house database with NMR spectra. Metabolic fingerprinting thus provides an efficient place to start hypotheses about affected metabolic pathways. However, definitive evidence will await confirmatory studies using techniques like liquid chromatography-mass spectrometry (LC-MS) and 2-dimensional NMR. Besides, from the presented metabolic fingerprints it becomes clear that not all signals could be identified using an in-house database. For elucidating structures of metabolites that are more difficult to identify from metabolic fingerprints (e.g. the heavily contributing signals around δ 3.25 in Figure 2B) other techniques are also indispensable in future.

Nevertheless, for global screening *in vitro* metabolic fingerprinting seems a promising technology: in a realistic nutritional research study with Caco-2 cells, biochemical changes in cells, resulting from exposure, could be detected well. Metabolic fingerprinting might in principle even be able to measure excretion of metabolites from cells into the culture medium, thus further helping the elucidation of cell metabolism.

In vitro metabolic fingerprinting studies provide an inexpensive starting point for formulation of hypotheses about affected metabolic pathways and could even become a replacement of costly *in vivo* metabolic fingerprinting studies. It will be a great challenge to develop more *in vitro* models to combine with metabolic fingerprinting. The results of these studies should be compared with similar *in vivo* studies, upon which in future *in vitro* based metabolic fingerprinting may function as an alternative for *in vivo* based fingerprinting in specific occasions. This could greatly enhance and facilitate evidence-based nutritional studies.

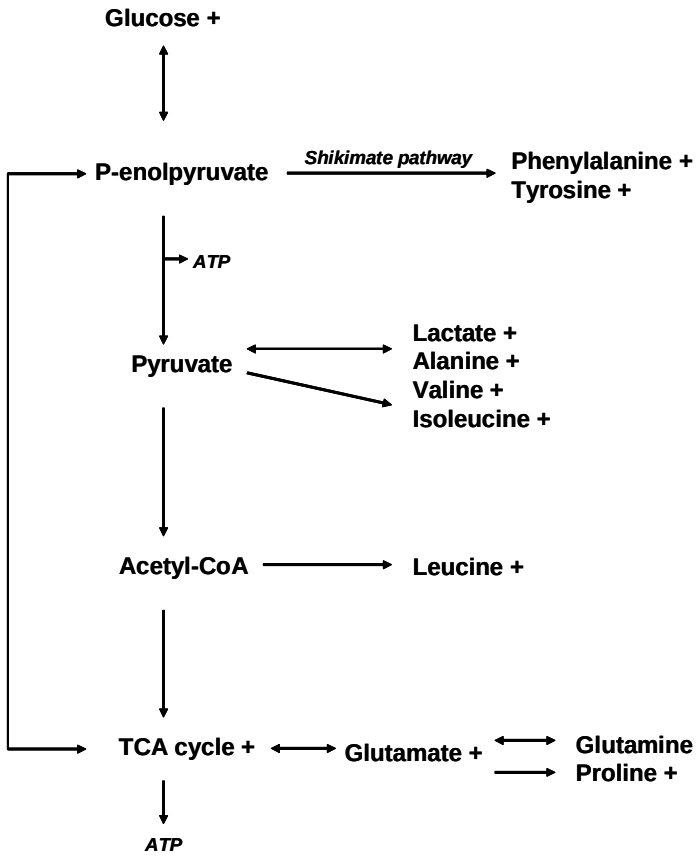

Figure 3. *Metabolic pathways in Caco-2 cells, which may be affected by treatment with inulin and its metabolites.*

Acknowledgements

The authors gratefully acknowledge Elly J. Spies-Faber for carrying out NMR measurements. Marleen M.C. van Nuenen, Annet J.H. Maathuis are thanked for running the *in vitro* model, Rob Onderwater, Wilfred J.M. Maas for helping with the Caco-2 experiments and anonymous reviewers for helpful comments.

References

1. Roberfroid M.B., Bornet F., Bouley C.H., Cummings J.H. (1995) Colonic microflora: nutrition and health. Nutr. Rev. 53: 127-130.
2. Gibson G.R., Beatty E.R., Wang X., Cummings J.H. (1995) Selective stimulation of bifidobacteria in the human colon by oligofructose and inulin. Gastroenterology 108: 975-982.

3. Kaur N., Gupta A.K. (2002) Applications of inulin and oligofructose in health and nutrition. J. Bioscience 27: 703-714.
4. Lamers R.A.N., DeGroot J., Spies-Faber E.J., Jellema R.H., Kraus V.B., Verzijl N., TeKoppele J.M., Spijksma G., Vogels J.T.W.E., van der Greef J., van Nesselrooij J.H.J. (2003) Identification of disease and nutrient related metabolic fingerprints in osteoarthritic guinea pigs. J. Nutr. 133: 1776-1780.
5. Nicholson J.K., Lindon J.C., Holmes E. (1999) ''Metabonomics'': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica 29: 1181-1189.
6. Lindon J.C., Nicholson J.K., Holmes E., Everett J.R. (2000) Metabonomics: metabolic processes studies by NMR spectroscopy of biofluids combined with pattern recognition. Concepts Magn. Reson. 12: 289-320.
7. Fiehn O. (2002) Metabolomics: the link between genotypes and phenotypes. Plant Molecular Biology 48: 155-171.
8. Van der Greef J., Davidov E., Verheij E., Vogels J.T.W.E., van der Heijden R., Adourian A.S., Oresic M., Marple E.W., Naylor S. (2002) The Role of Metabolomics in Drug Discovery: A New Vision for Drug Discovery and Development. In: *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis* (Harrigan, G.G. & Goodacre R., eds.). Kluwer Academic Publishing. Boston MA: 170-198.
9. Hoogerbrugge R., Willig S.J., Kistemaker P.G. (1983) Discriminant analysis by double stage principal component analysis. Anal. Chem. 55: 1710-1712.
10. Fusunyan RD, Quinn JJ, Fujimoto M, MacDermott RP, Sanderson IR. (1999) Butyrate switches the pattern of chemokine secretion by intestinal epithelial cells through histone acetylation. Mol. Med. 5: 631-640.
11. Minekus M., Smeets-Peters M., Bernalier A., Marol-Bonnin, Havenaar R., Marteau P., Alric M., Fonty G., Huisin'tveld J.H.J. (1999) A computer-controlled system to simulate conditions of the large intestine with peristaltic mixing, water absorption and absorption of fermentation products. Appl. Microbiol. Biot. 53: 108-114.
12. Venema K., Van Nuenen M., Smeets-Peters M., Minekus M., Havenaar R. (2000) TNO's *in vitro* large intestinal model: an excellent screening tool for functional food and pharmaceutical research. Nutrition 24: 558-564.
13. Vogels J.T.W.E., Tas A.C., Venekamp J., van der Greef J. (1996) Partial linear fit: a new NMR spectroscopy preprocessing tool for pattern recognition applications. J. Chemometrics 10: 425-438.
14. Reeds P.J., Burrin D.G. (2001) Glutamine and the bowel. J. Nutr. 131: 2505S-2508S.
15. Mithieux G. (2001) New data and concepts on glutamine and glucose metabolism in the gut. Curr. Opin. Clin. Nutr. Metab. Care 4: 267-271.
16. Michal G. (1999) Biochemical pathways, an atlas of biochemistry and molecular biology. Spektrum Akademischer Verlag. Heidelberg, Germany.

8

Profiles of metabolites and gene expression in rats with chemically induced hepatic necrosis

Abstract

This study analysed changes in gene expression patterns and metabolite levels in plasma or urine in parallel. The aim was to more sensitively detect hepatotoxicity and provide new insights in molecular mechanisms of hepatic necrosis. Rats received the model hepatotoxicant bromobenzene at three dose levels, the highest dose inducing acute centrilobular necrosis. The hepatic transcriptome and plasma and urine metabolite profiles were analysed after 6, 24 and 48 hours, using multivariate statistics.

Principal component analysis showed that molecular profiles from rats with hepatic necrosis differed largely from controls. Changes in levels of genes and metabolites were identified in correlation with the degree of necrosis, providing putative novel markers of hepatotoxicity. Moreover, samples from treated rats were distinguished from controls after exposure to bromobenzene below the concentration that induced hepatotoxicity markers or histopathological changes. Genes with altered expression were involved in oxidative stress, the acute phase response, cytoskeleton structure, apoptosis, biotransformation, glycolysis, cholesterol and fatty acid metabolism. Levels of endogenous metabolites like alanine, lactate, tyrosine and dimethylglycine distinguished plasma from treated and control rats. Complementary, NMR metabolite profiling enabled to distinguish the urine samples based on the exposure levels, primarily through presence of a multitude of bromobenzene-derived metabolites.

Concluding, this parallel analysis of the liver transcriptome and metabolite profiles in plasma enabled to more sensitively detect changes related to hepatotoxicity and discover novel markers. Additional insights in the role of various biological pathways in bromobenzene-induced hepatic necrosis were obtained.

Submitted for publication.

Introduction

Previous toxicogenomics studies have shown that both large-scale measurement of gene expression (transcriptomics) and metabolite profiling complement the current methods to identify and discriminate different types of toxicity. Moreover, the new technologies enable to investigate the mechanisms that lead to toxicity. To this date, most studies concentrated on hepatic toxicity. Transcriptomics using DNA-microarrays enabled the discrimination of responses by different classes of hepatotoxicants *in vivo*, as shown by [1-3]. Hamadeh and coworkers distinguished samples treated with two classes of toxins, and provided more details on the mechanisms of action [4].

In parallel, metabolomics, i.e. metabolite profiling by NMR combined with pattern recognition techniques, has been used to classify urine samples of rats treated with either a liver or a kidney toxicant [5-6]. Others analysed metabolites in liver, plasma and urine of rats treated with the model hepatotoxicant alpha-naphthylisothiocynanate (ANIT) [7-8]. Urine profiles were analysed in time upon single dosage of ANIT, galactosamine and butylated hydroxytoluene [9]. Time-related differences in metabolite contents were related to the stage of the lesions, and specific changes in metabolite levels were identified for each compound.

While gene expression changes influence biochemical reactions, metabolite levels are determined by those biochemcial reactions. Therefore, complementary information is expected from so-called systems toxicology approaches, where transcriptomics, proteomics and/or metabolomics are combined to analyse toxicity in a systematic and holistic manner. Only few experiments integrating results from transcriptomics and metabolite profiling have been described. Very recently, Coen and colleagues reported transcriptomics and metabolomics analyses in mice treated with acetaminophen (paracetamol) [10]. This study demonstrated that analysis of gene expression and metabolite profiles provided complementary insights in APAP-induced hepatic effects. In earlier studies in our laboratory, we evaluated the combined use of transcriptomics and proteomics analyses of hepatotoxicity induced by bromobenzene (BB). Bromobenzene is a well studied model toxicant that causes necrosis in the liver (centrilobular) and kidney. Hepatic biotransformation and toxicity of BB in rat have been reported in detail [11-15]. Because the liver is the target for toxicity induced by many compounds including

bulk chemicals, drugs and food ingredients, the characteristics of the response induced by BB could be helpful in understanding hepatotoxicity induced by a variety of xenobiotics.

Transcriptomics and proteomics analyses of hepatotoxicity were evaluated 24 hours after a single *i.p.* dose of BB [16]. A new study was designed to determine the acute hepatotoxic effects at the gene expression level in time, after oral dosage of various concentrations of BB. Hepatic necrosis was observed only at the high dose level after 24 h, though gene expression changes characteristic for BB exposure were observed at 2.5 times lower dose level. A few genes changed at 10 times lower dose levels. expression of several genes was found to change 6 h after dosage. Genes that were statistically significant differentially expressed upon BB dosage were involved in processes like drug metabolism, oxidative stress, GSH synthesis and the acute phase response [16].

Aim of the study

In the present study, the aim was to investigate whether integrated analysis of the data from transcriptomics and metabolite profiling further increased the sensitivity of detection of hepatotoxicity. Our second question was how the combined analysis may expand current knowledge about the mechanism of chemically-induced hepatotoxicity. Moreover, relationships between gene expression changes and altered metabolite levels were assessed. Thus, NMR-based metabolite profiles of plasma and urine samples, collected from the study described by Heijne and colleagues [16] were combined with the transcriptomics data of this same study. The metabolite profiling aimed at detecting changed concentrations of endogenous metabolites as a result of hepatotoxicity (biomarkers of effect) and of BB-derived metabolites in urine and plasma (biomarkers of exposure). Results from parallel gene expression and metabolite analysis were combined with pre-existing biochemical knowledge in an overall interpretation of the mechanisms of action and effects of BB, a necrosis-inducing chemical, on liver physiology.

Materials and methods

Urine and plasma samples were collected from the study by Heijne and colleagues [16] which was also the source of the transcriptomics and toxicity data. Briefly,

three doses of bromobenzene (0.5, 2.0 and 5.0 mmol/kg body weight, dissolved in corn oil, 40% v/v) were administered to male Wistar rats by oral gavage. Animals were kept under controlled conditions, and the welfare of the animals was maintained in accordance with the general principles governing the use of animals in toxicity experiments of the European Communities (Directive 86/609/EEC) and Dutch legislation (The Experiments on Animals Act, 1997). Nine rats per dose group were treated BB or corn oil, while an additional group was not treated. Three rats per group were sacrificed after 6, 24 and 48 h and blood and livers were collected. Urine was collected for metabolomics between dosing and sacrification for the 6 h group, and during the last 16 h before sacrifice for the 24 and 48 h groups. During the time urine was collected, rats received water *ad libitum,* but no food.

Transcriptomics

cDNA microarray preparation and hybridization was described previously [16]. A reference RNA was used, and hybridizations were replicated with swapped fluorophore incorporation (Cy3 and Cy5) in the sample and reference RNA. After quality filtering, lowess normalization and log(base 2) transformation, a set of about 2700 cDNAs was obtained. In present study, we required a correlation higher than 0.6 between the duplicate sets of dye-swap measurements, keeping about 400 genes in the dataset.

NMR analysis

NMR spectra of urine of individual animals were recorded in triplicate, according to [17]. Plasma samples were deproteinised by filtration. Filters with a cutoff of 10 kDa (Microcon YM-10, Millipore) were spin-rinsed with 0.5 ml of 0.05 M NaOH followed by 2 × 0.5 ml de-ionised water to avoid contamination of the ultrafiltrate with glycerin. Centrifugation (1h at 10000 rpm) of 0.5 ml plasma over a filter was followed by the centrifugation (1h at 10000 rpm) of 0.5 ml de-ionised water. Filtrates were freeze-dried and reconstituted in 750 µl sodium phosphate buffer (pH 6.0, made up with D_2O) containing 1mM sodium trimethylsilyl-[2,2,3,3,-2H4]-1-propionate (TMSP) as an internal standard. NMR spectra were recorded in a fully

automated manner on a Varian UNITY 400 MHz spectrometer (Palo Alto, CA, USA) according to [17].

Data preprocessing and multivariate data analysis

The NMR data file was imported into Winlin (V1.12, TNO, Zeist, The Netherlands). Minor variations from comparable signals in different NMR spectra were adjusted and aligned without loss of resolution. The intensities of signals present in each NMR spectrum were normalised, so that the sum of all intensities was equal to 1. This data set was imported into Matlab (Version 6.5, The MathWorks Inc., Natick, MA, USA) together with the transcriptomics data for preprocessing and multivariate data analysis. The data matrix was centered across time and dose. The sum of squares per variable over time and dose was scaled to 1, and PCA was performed. PCA is a multivariate statistical analysis that reduces the many dimensions of a dataset to few dimensions that describe the majority of the variance. PCA was performed with the PLS toolbox (Version 3.0, Eigenvector Research Inc., Manson, WA, USA), and a score plot visualised differences in gene expression and metabolite profiles. The contribution of each variable to the trend observed in the plot was determined. PCA was also performed on plasma and urine NMR data separately. When score plots revealed differences between groups, the contributions of the original NMR signals to these difference between treated and control were displayed in a factor spectrum. Metabolites were identified using an in-house reference database.

Results

Rats were exposed to the chemical compound bromobenzene and developed hepatic necrosis 24 h after dosing with the high concentration. In parallel, hepatic gene transcription and profiles of plasma and urine metabolites were analysed.

Toxicological examinations

No macroscopic aberrancies of the liver or other organs were observed in any of the rats sacrificed 6 hours after dosage. Histopathology of liver tissue showed no abnormalities in the controls and low dose rats. In some livers a slight presence of mononuclear cell aggregates and/or necrotic hepatocytes was observed. Only in

rats that received high concentration of BB, livers had a patchy appearance and gross lesions after 24 hours and focal discoloration after 48 hours. Centrilobular necrosis was found in livers of all those rats, with inter-individual variation in the degree of response. Plasma levels of ASAT, ALAT and bilirubin were markedly elevated, also with inter-individual varation. To correlate the conventional markers of hepatotoxicity with the degree of necrosis in the individual rats, a semi-quantitative score was defined for the hepatocellular necrosis ranging from 0 (no effects) to 10 (very severe centrilobular necrosis) (Table 1). This score was also used to correlate gene expression levels to necrosis. Figure 1 depicts the correlation of ASAT, ALAT, bilirubin, and the relative liver weight with the observed degree of hepatocellular damage.

Apart from the signs of hepatotoxicity, BB significantly decreased plasma levels of glucose at the mid and high dose, after 24 and 48 hours. Cholesterol (n.s.) and phospholipid levels increased by high BB treatment at all time points. Hepatic GSH levels, which play a pivotal role in the hepatotoxicity induced by BB, were slightly decreased six hours after administration of BB. The mid and high dose depleted GSH levels to ~25% of control levels. After 24 hours, GSH levels were nearly restored.

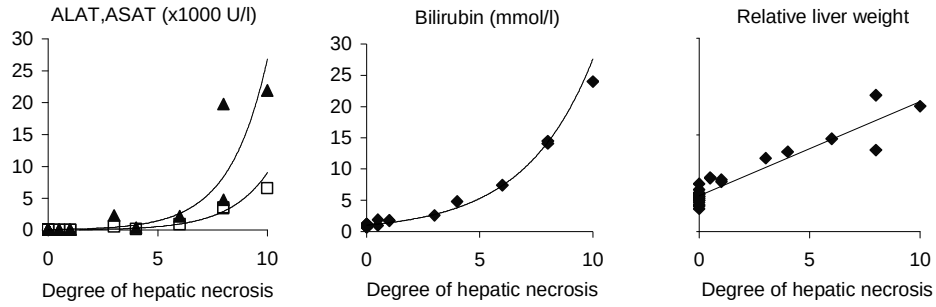

Figure 1. *Correlation of toxicity markers ALAT, ASAT, bilirubin and relative liver weight with the observed degree of hepatic necrosis in individual rats.*

Table 1. Histopathological and clinical chemistry findings in rats, 24 and 48 hours after exposure to mid and high dose of bromobenzene. The degree of hepatic necrosis was expressed with a score (S) between 0-10; M is mean.

Dose mmol / kg BW	**T** h	**Rat**	**Rel. liver** % of CO	**BW** g	**Gross pathology of the liver**	**Liver histopathology**	**S**
2.0	**24**	**M**	**109%**	**187**	**No gross lesions**	**(Very) Slight mononuclear cell aggregates/ necrotic hepatocytes**	
		62	108%	188	-	Slight mononuclear cell aggregates/ necrotic hepatocytes	1
		64	108%	178	-	Slight mononuclear cell aggregates/ necrotic hepatocytes	1
		66	110%	194	-	Very slight mononuclear cell aggregates/ necrotic hepatocytes	0.5
2.0	**48**	**M**	**108%**	**184**	**No gross lesions**	**Very slight mononuclear cell aggregates/ necrotic hepatocytes (1/3)**	
		68	114%	191	-	Very slight mononuclear cell aggregates/ necrotic hepatocytes	0.5
		70	105%	186	-	No abnormalities	0
		72	104%	175	-	No abnormalities	0
5.0	**24**	**M**	**131%**	**177**	**Patchy appearance**	**Centrilobular necrosis Nucleolar enlargement (2/3)**	
		80	120%	183	Patchy appearance	Very slight centrilobular necrosis Nucleolar enlargement	3
		82	125%	173	Patchy appearance	Severe centrilobular necrosis Nucleolar enlargement	8
		84	149%	174	Patchy appearance	Very severe centrilobular necrosis	10
5.0	**48**	**M**	**142%**	**181**	**Pale appearance others (2/3)**	**Centrilobular necrosis Slight centrilobular fatty change (2/3) Mitotic increase (2/3) Nucleolar enlargement**	
		86	129%	187	No gross lesions	Slight centrilobular necrosis Nucleolar enlargement	4

	88	136%	186	Pale appearance Pronounced lobular pattern	Moderate centrilobular necrosis Slight centrilobular fatty change Slight mitotic increase Nucleolar enlargement	6
	90	161%	170	Pale appearance Red areas Firm tissue	Severe centrilobular necrosis Slight centrilobular fatty change Slight mitotic increase Nucleolar enlargement	8

Transcriptomics analysis and parallel metabolite profiling

BB elicited specific changes in gene expression of many rat liver genes, as reported before [16]. In this study, the profiles of the transcriptomics measurements were combined with the profiles obtained by NMR, describing the metabolite contents of plasma. Consensus PCA [18] was performed using both types of data in one integrated analysis, and results are shown in Figure 2. This plot indicates that the samples from the high and mid dose groups, collected after 24 and 48 hours were distant from the others, having lower PC1 scores. Most distinct from all the other samples were the samples from rats #84, #82, and #90, that received a high dose of BB. Microscopic examination revealed (very) severe hepatic centrilobular necrosis in those rats. Profiles of rats #80, #86 and #88 were less distant from the controls. Correspondingly, moderate centrilobular necrosis was observed in rat #88, and (very) slight necrosis in rats #86 and #80. The profiles of the rats treated with a mid dose of BB were distinct from the controls after 24 hours. Routine markers were not able to indicate hepatotoxicity in those rats. After 48 hours, rats treated with the mid dose were not distinct from controls. Samples from rats treated with the low dose of BB were not readily separatable from the controls, after 24 or 48 hours. All samples collected after 6 hours were distinct from the other time points in the down right corner of the plot. Treatment with BB resulted in patterns distinct from the controls.

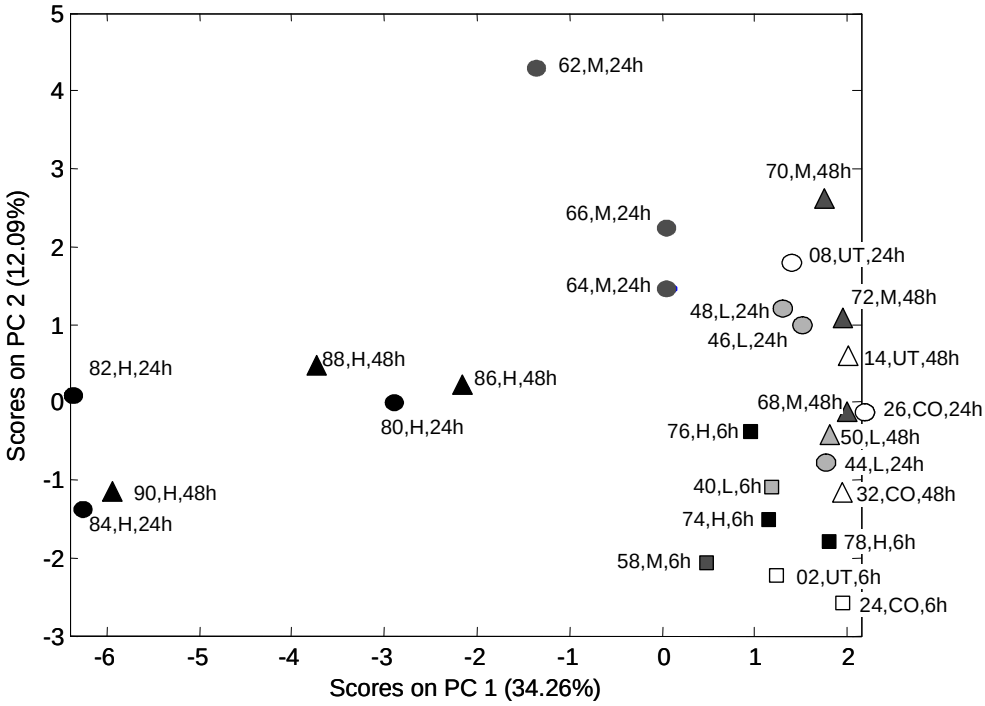

Figure 2. *Score plot of consensus PCA. Consensus PCA was performed using both hepatic transcriptomics and plasma metabolite profilng data in one integrated analysis. The percentage of the total variance explained by the individual PCs is indicated in the plots. Time points: Boxes: 6 h samples, circles: 24 h samples; triangles: 48 h samples. Dose levels: white: controls, light gray: low BB; dark grey: mid BB; black: high BB.*

Genes and metabolites were sorted according to their contribution to the observed trend, reflecting the degree of hepatic necrosis. Tables 2A and B list the genes and metabolites with the highest and lowest scores, that therefore putatively correlate with the degree of hepatotoxicity. The levels of gene expression and metabolites listed in are present at either high or low levels in correlation with the necrosis. Many genes with a significant contribution to pattern differences in the PCA were identified to be up- or down regulated by BB with high significance in univariate statistical tests, and the rationale of these changes in terms of toxicology was discussed before [16].

Genes with high scores in the parallel analysis include structure and cytoskeleton-related genes (beta actin, weakly similar to pervin, tubulin), many ribosomal

subunits, and other factors involved in protein synthesis (eg. nucleophosmin). Also oxidative stress induced genes (Ho-1, Timp1, peroxiredoxin1, ferritins), hepatic acute phase response genes (orosomucoid 1, fibrinogen gamma) and enzymes involved in glucose metabolism (Gapdh, phosphoglycerate mutase 1, aldolase A) have high rankings. Drug metabolising enzymes like Ephx1, Afar, Gsta and aldo-keto reductases, likely involved in the hepatic biotransformation of bromobenzene, appeared in the upper part of the ranking. Several cell cycle and apoptosis related genes (Bcl2-related protein A1, Pcna, p53, p21 (Waf), EST, highly similar to p53-regulated PA29-T2, cyclin G1) were coordinately upregulated. High ranked genes with others functions include casein kinase II, VL30 element and RAN. Plasma metabolites with a high score in the analysis include acetate, choline, phenylalanine and some uncharacterised metabolites.

Genes with low scores include hepatic acute phase response genes like alpha-1-inhibitor, serine protease inhibitor, fibrinogen beta, complement components, drug metabolising enzymes like Cyps, aldehyde dehydrogenases, Fmo3, enzymes involved in fatty acid and cholesterol metabolism (HMG-CoA synthase, Lcat, Star, fatty acid CoA ligase, acyl CoA dehydrogenases) and glucose metabolism (G6pt1, alanine-glyoxylate aminotransferase) Many genes with other functions, like asialoglycoprotein receptor 2, Cathepsin S, and dimethylglycine dehydrogenase had a low score, indicating that they were down regulated compared to the controls. Plasma metabolites with a low score in the analysis include dimethylglycine, tyrosine and glucose.

Table 2A: *Highest and lowest ranked genes from consensus principal component analysis. Rank, gene name and Genbank accession number, category, and the correlation to the degree of hepatic necrosis are indicated. Expression of high ranked genes is upregulated, while low ranked genes are downregulated in the samples with a high degree of hepatic necrosis.*

Rank	Gb Acc.	Category	Gene name	Correl
1	AA859846	Structure	actin, beta	0.881
2	AA964725	Structure	Weak sim to pervin	0.842
3	AA964496	Structure	High sim to S11222 actin gamma, cytoskeletal	0.853
4	AA957078	Structure	alpha-tubulin	0.842
5	AA924111	glycolysis	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	0.811
6	AI029162	APR	Orosomucoid 1	0.885
8	AA997175	Signal transd.	casein kinase II beta subunit	0.835
10	AA900726	Signal transd.	GTP-binding protein (ral A)	0.789
552	AI070895	Fatty acid	Weakly sim. to acyl-CoA dehydrog.,epoxide hydr.[Cel]	-0.735
553	AA866389	other	lumican	-0.865
554	AA964340	other	syndecan 2	-0.747
555	AA955402	Cysteine	S-adenosylhomocysteine hydrolase	-0.776
556	AA925933	Proteolysis	cathepsin S	-0.801
557	AA819756	Drug metab	arachidonic acid epoxygenase;Cyp2C23	-0.888
558	AI136048	Cholesterol	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2	-0.821
559	AI071033	acute phase	Fibrinogen, B beta polypeptide	-0.820
560	AA997322	Cholesterol	Lecithin-cholesterol acyltransferase (Lcat)	-0.745
561	AA997920	Signal transd.	asialoglycoprotein receptor 2	-0.900

***Table 2B**. Highest and lowest ranked metabolites from consensus principal component analysis. The rank and chemical shift in the NMR analysis of the (putatively) identified metabolites is indicated. High ranked metabolites are more abundant in treated compared to control plasma samples, and metabolites with a low rank are more abundant in controls.*

		High rank			Low rank
Rank	**Shift**	**Metabolite**	**Rank**	**Shift**	**Metabolite**
7	1.475	alanine	430	2.935	dimethylglycine?
9	1.4925	alanine	391	6.91	tyrosine
51	8.4575	formate	390	6.89	tyrosine
61	7.83	unidentified metabolite	389	7.185	tyrosine
70	1.9275	acetate	383	7.2075	tyrosine
73	3.2075	choline?	379	3.4925	glucose
82	3.0075	unidentified metabolite	378	3.245	glucose
100	7.4325	phenylalanine?	376	3.3775	glucose
109	3.5975	choline?	375	3.7325	glucose
110	5.3875	unsaturated lipid?	374	3.4225	glucose
111	3.0275	cysteine?	373	4.64	glucose
114	7.8475	histidine?	372	3.715	glucose
119	7.4125	phenylalanine?	369	5.23	glucose
124	7.3375	phenylalanine?	368	3.09	unidentified metabolite
126	7.375	phenylalanine?	367	3.7475	unidentified metabolite
129	1.005	isoleucine?	366	3.2225	unidentified metabolite
130	3.0525	unidentified metabolite	365	3.4375	unidentified metabolite

Gene expression markers

The correlation between the level of gene expression and the degree of necrosis in the individual rats was calculated (Table 2). Figure 3 illustrates expression of ESTs highly similar to actin and pervin, and orosomucoid 1 in relation to hepatic necrosis.

Expression levels of asialoglycoprotein receptor 2 and lecithin-cholesterol acyltransferase (Lcat) decreased in concordance with the degree of hepatic damage. In total, 14 genes were found with a positive correlation between 0.80 and 0.89, the highest coefficient. The correlation of the average expression level of these 14 genes with necrosis was 0.969. In parallel, 20 negatively correlated genes were found with an individual correlation to necrosis varying from -0.80 to -0.90. The correlation of the average gene expression of these 20 genes with necrosis was -0.959. This suggests that valuable markers of hepatocellular necrosis consist of combination of gene expressions.

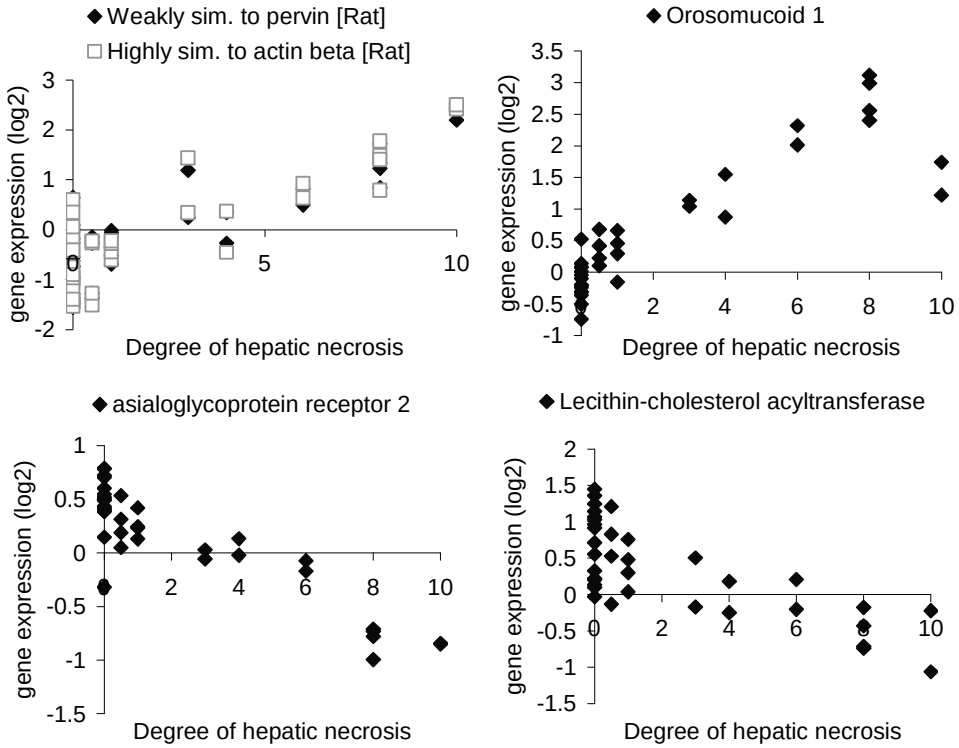

Figure 3. *Levels of expression of genes that correlate with the observed degree of hepatic necrosis in individual rats.*

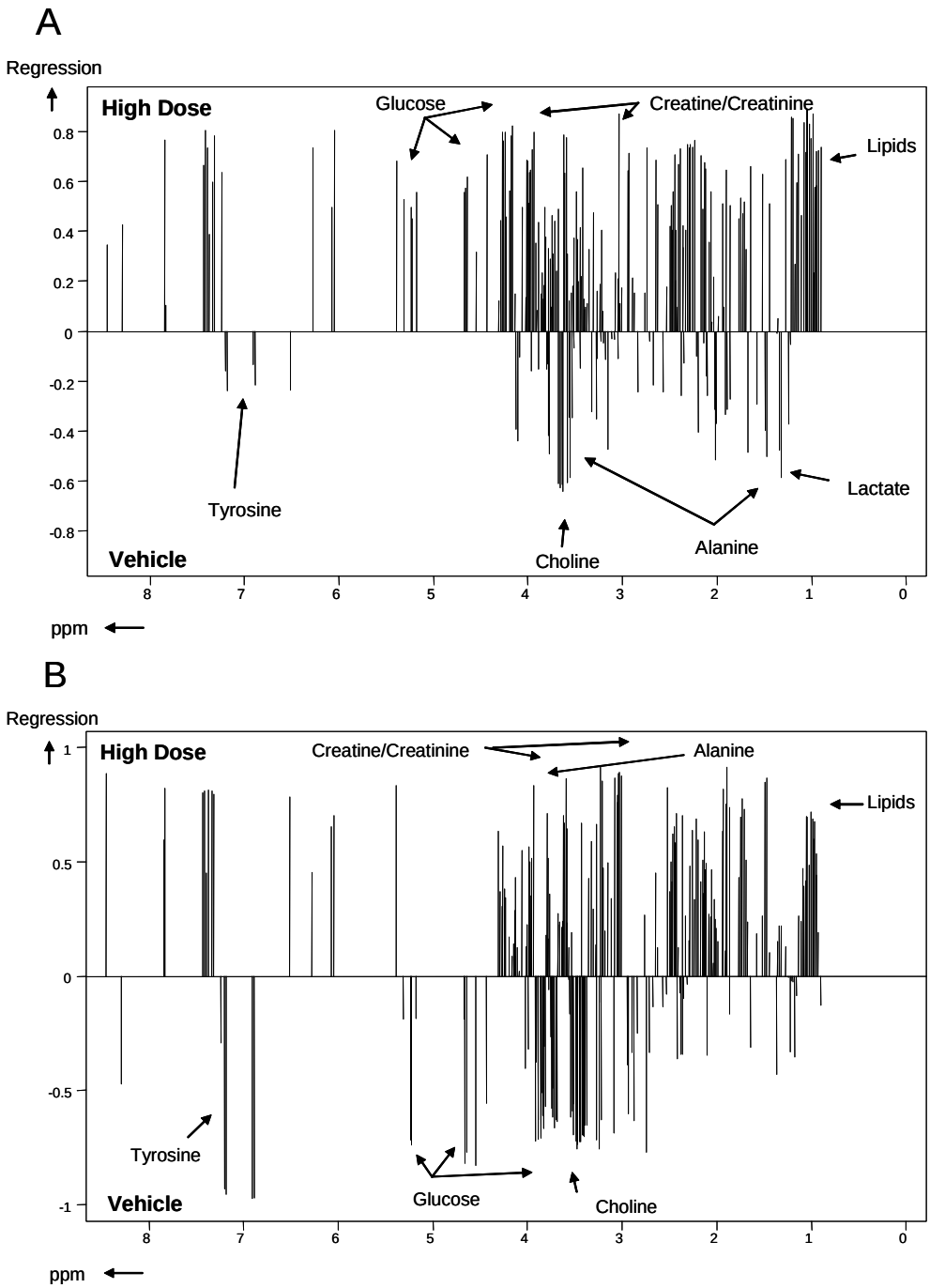

Figure 4. (A) and (B) *Factor Spectra of NMR measurements after principal component discriminant analysis. Panel A: plasma after 6 hours, panel B: plasma after 24 hours.*

Time and dose-dependent changes in plasma metabolites

Besides the parallel analysis of transcriptomics and metabolite profiling, the plasma NMR data were analysed separately by PCA, and time and dose specific changes in metabolite levels between treated and control samples were visualised in factor spectra. (Figure 4a and b). After both high and mid dose of BB, lipid levels were higher than in controls. Clinical chemistry indicated an increase in plasma phospholipid levels upon high but not mid dose treatement. The levels of glucose were higher 6 hours after a high dose of BB, but lower after 24 and 48 hours. These observations were identical to the clinical chemistry measurements. NMR of plasma showed higher levels of creatine and/or creatinine in BB-treated rats, though clinical chemistry did not reveal significant changes in creatinine. The levels of tyrosine were lower 6 and 24 hours after BB, while higher after 48 hours. Methionine, alanine and lactate levels in plasma of BB-treated rats were lower 6 hours after dosage but higher 24 and 48 hours after dosage. Dimethylglycine and taurin levels were increased compared to controls 6 hours after the BB treatment, and decreased after 24 hours. Choline levels were decreased after treatment to mid or high dose of BB.

Profiles of urine metabolites

Also in urine, metabolite NMR profiles were discerned using PCA. Analysis per time point showed that BB treatment changed urine profiles (data not shown). All rat urines collected during the first 6 hours could be distinguished by levels of exposure. By 48 hafter dosage, rats treated with the high concentration of BB could still be recognised from controls by their urine profiles. In order to determine the NMR signals that most significantly differed between the high dose and control group, factor spectra were constructed. Figure 5 shows the factor spectrum for rat urine collected during the 6 hours after dosage. Using reference databases, the identity of several peaks was established. Factor spectra revealed the marked presence of BB-derived metabolites like bromphenols, bromcatechols, and quinones in urine. It was not possible to discriminate and identify these various metabolites. Markedly elevated levels of mercapturic acids, derived from GSH-conjugates, were observed after treatment. Methionine levels in urine were higher in the treated rats compared to controls. Formate levels increased after 24 hours in

the treated rats, and elevated levels were observed of urocanate and (methyl)histidine, as well as decreased levels of nicotinate, hippurate, phenylalanine/tyrosine and glucose/fructose.

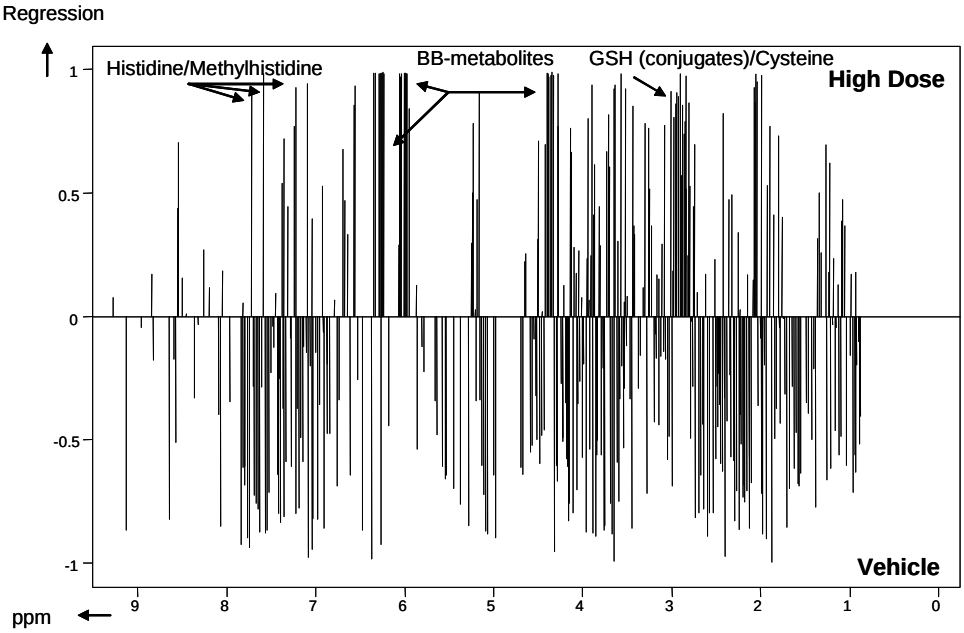

Figure 5. *Factor spectrum after principal component discriminant analysis of NMR spectra of urine, 6 hours after dosage with bromobenzene, compared to vehicle control.*

Discussion

This study presents one of the first integrated toxicogenomics studies, where acute hepatotoxicity was analysed at the transcriptome and metabolite level in a time- and dose-dependent manner. An integration of the (raw) datasets of the transcriptomics and metabolomics experiments could increase the sensitivity of detection of hepatotoxicity. Moreover, this could enhance the assessment of relationships between gene expression and metabolite level changes.

When rats were treated with BB, hepatic centrilobular necrosis was observed after 24 hat the high, but not at lower doses. The inter-individual response varied from

very slight to very severe hepatic centrilobular necrosis**.** Individual plasma ALAT, ASAT and bilirubin levels and the increase in relative liver weight correlated with the severity of the necrosis. Complementary to these toxicological observations, the molecular profiles of hepatic gene expression and plasma metabolites were analysed in parallel. Differences beteween molecular profiles were dependent on the dose and time after dosage. Profiles from the 6 htime point were distinguishable from other time points. BB treatment at the high dose resulted in highly distinct profiles, while the mid dose altered the profiles up to 24 hafter dosage. At this dose level, conventional signs of hepatotoxicity were not observed. Combining transcriptomics and metabolite profiling did not allow to discriminate samples treated with the low dose from controls.

Markers of gene expression

Gene expression changes were identified in correlation with the degree of hepatic necrosis, providing comprehensive means to diagnose the degree of necrosis. Moreover, if these markers prove to be predictive at earlier time points or lower dose levels, they will improve detection of hepatotoxicity. Changes in these marker gene expression levels could be explained from a mechanistic point of view. The upregulation of cytoskeleton constituents (actin and pervin ao.) with the degree of necrosis indicates remodelling of the cytoskeleton. Presumably, necrosis and repair occur simultaneously in different liver cells, but our experiments using whole liver do not allow to localise the events. The negative correlation of genes like alpha-1-inhibitor and serine protease inhibitor is probably related to the acute phase response, involving altered hepatic synthesis of proteins. When expression levels were averaged for sets of genes, the correlation with the degree of necrosis increased. A further suggestion would be to construct a model of combined sets of positively and negatively correlated genes and metabolites to further increase the relation with hepatocellular necrosis.

Metabolite profiles

Xenobiotic compounds like BB are degraded into many metabolites, and ultimately excreted in urine. BB-derived metabolites could be suitable to monitor exposure and to elucidate routes of biotransformation. Levels of endogenous metabolites

that changed after treatment form putative biomarkers of toxicity, and could help to identify the mechanism of hepatotoxicity.
Urine collected from rats exposed to different doses of BB varied in metabolite contents, in agreement with the levels of exposure. Especially shortly after dosage, many water-soluble BB-metabolites were found, like bromphenols, -catechols and -quinones, and mercapturic acids. The lack of reference spectra and insufficient resolution of the separation frustrated the identification of all corresponding metabolites. Peaks around 6 ppm in the spectra could result from bromphenols, bromcatechols and/or BB-dihydrodiols. Therefore, the precise biotransformation of BB could not be determined. Further efforts to elucidate this based on urine metabolite profiles require techniques like liquid chromatography and mass spectrometry (LC-MS) for identification of the compounds. Few endogenous metabolites, putative markers of hepatotoxicity, were discovered in urine. Levels of methionine were higher 24 hours after BB dosage. Urocanate, related to histidine metabolism, and histidine itself displayed elevated levels. Notably, elevated urocanate levels were also found with galactosamine-induced hepatotoxicity [9].

Contrary to urine, in plasma, distinct signals of BB-derived metabolites were not found. On the other hand, endogenous metabolites in plasma, or combinations of them, could be effective biomarkers of toxicity. Decreased glucose and increased lipid levels measured by NMR were corroborated by clinical chemistry. The levels of formate in plasma, and urine, were increased after 24 and 48 hours. Formate could be produced from dimethylglycine through sarcosine and formaldehyde. Formate is also a product of oxalate in the glyoxylate catabolism, and possibly related to folate synthesis in the one-carbon metabolism.

Biochemical pathways
The most significant effects determined in the parallel analysis of transcriptomics and plasma metabolite profiling were categorised according to biochemical pathways. Changes in gene expression in several pathways were described previously [16]. Other changes, eg. in apoptosis and cell cycle were not noted before. Pathways like glycolysis, GSH and amino acid metabolism were disturbed both at the gene expression and metabolite level, and are described below. Figure

6 presents a proposed schematic overview of changes in GSH and amino acid metabolism, associated to bromobenzene-induced hepatic necrosis.

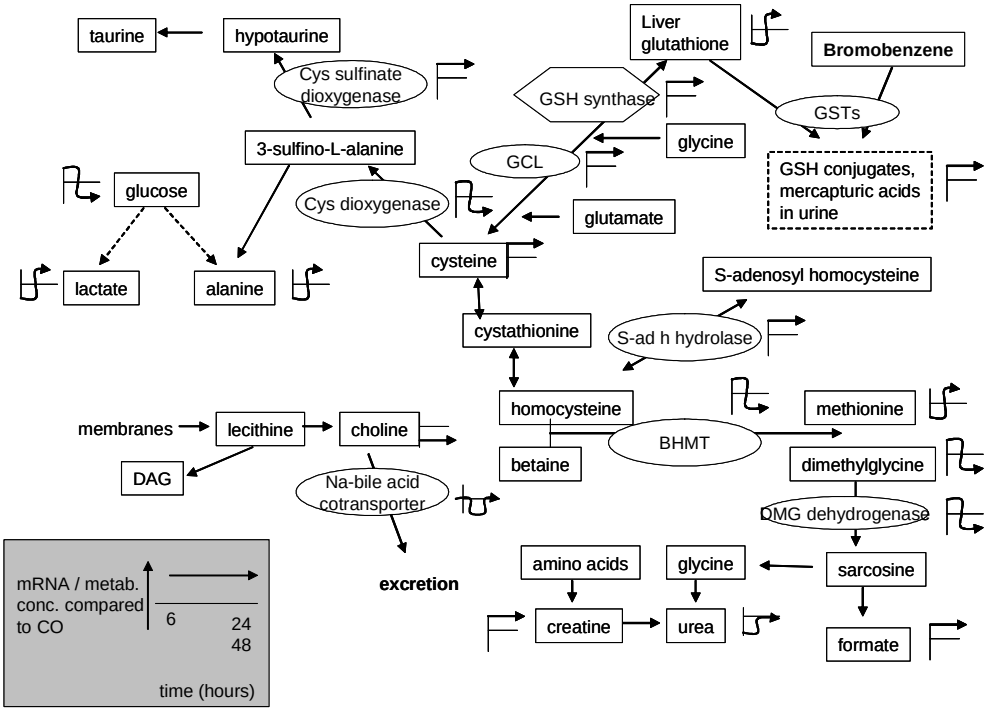

Figure 6. *Simplified, schematic representation of gene expression and plasma metabolite changes in bromobenzene-induced hepatic necrosis, related to GSH and amino acid metabolism. Ovals represent genes, boxes represent metabolites in plasma. When measured, changes in geneexpression or plasma metabolite levels are indicated schematically to the right of each object.*

Glycolysis

Glucose levels in plasma decreased in time after BB treatment. This could be ascribed to increased glycolysis, in order to increase the production of energy to restore homeostasis after the toxic insult. Decreasing glucose levels are corresponding with increasing plasma levels of alanine and lactate, products that may be formed by breakdown of glucose when the oxidation of pyruvate is incomplete. Expression of many genes involved in glycolysis, gluconeogenesis and

glucose transport was altered. Expression of a glucose transport protein was decreased by BB. From the changes, we conclude that glycolysis enzymes were induced (GAPDH, aldolase A, pyruvate kinase, G6PD and PGAM), and gluconeogenesis was reduced through down regulation of G-6-phosphatase, transport protein 1 (G6pt1), alanine-glyoxylate aminotransferase and pyruvate carboxylase. It is known that the hepatotoxic effects of high doses of APAP are similar to the effects of bromobenzene. In agreement with our findings, APAP was found to decrease glucose levels and was suggested to induce glycolysis based on gene expression and metabolite profile changes, suggestively as a reaction to decreased ATP availability from beta oxidation of fatty acids [10].

GSH and amino acid metabolism

A central process in the chemically-induced hepatic necrosis is the depletion of GSH levels, which normally protect cells by scavenging of hazardous, reactive molecules. GSH levels decreased to around 25% of controls, 6 hours after oral BB dosage, [16], while total depletion of hepatic GSH was observed 24 h after i.p. administration of BB. GSH is used in conjugation reactions to BB-derived metabolites, catalysed by GSTs. The reduction of GSH levels was accompanied by a decrease of plasma methionine, according to NMR measurements. The GSH depletion was countered through induction of GSH synthase protein [16] and Gclc gene expression. Along with the changes in GSH and methionine levels, related enzymes and metabolites were found to change. GSH and methionine levels are connected via cysteine and homocysteine levels, involving enzyme activity of BHMT. Gene expression of BHMT was found to initially increase, and later decrease upon BB treatment. The expression of S-adenosyl homocysteine hydrolase was decreased. Plasma levels of dimethylglycine, produced in the reaction catalysed by BHMT were found to correlate with the BHMT mRNA levels in time, and also the hepatic dimethylglycine dehydrogenase gene expression levels followed this pattern. Dimethylglycine can be catalysed in a multi-step reaction to formate, which levels were increased both in plasma and urine after treatment. Induced levels of cysteine in plasma were observed after BB treatment, along with increased gene expression of cysteine dioxygenase, while increased levels of cysteine sulfinic acid decarboxylase were observed before [16]. Plasma

tyrosine levels show a characteristic pattern, decreasing drastically 24 hours after high BB, while 48 hours after high BB, levels were highly increased compared to controls. Protein levels of HPD, an enzyme involved in tyrosine metabolism, were found to decrease 24 hours after BB [16]. The level of phenylalanine is related to tyrosine and seems to decrease in plasma due to the treatment.

Conclusion

In summary, this study presents one of the first integrated analyses of transcriptomics and metabolite profiling, revealing additional information in the process of chemically-induced hepatic necrosis. A full merge between the methods awaits technical optimization, especially for the identification of metabolites. Nevertheless, corroborating findings from liver transcriptomics and plasma metabolite profiling aided in the generation of new hypotheses concerning cellular mechanisms putatively related to necrosis, such as changes in cytoskeleton remodeling and acute phase response, apoptosis, glycolysis, amino acid, fatty acid and cholesterol metabolism. Through integration of the datasets, changes were observed before histopathology or clinical chemistry indicated necrosis. Both liver gene and plasma metabolite markers were discovered in correlation with the degree of hepatocellular necrosis in individual animals. Through measurement of urine metabolite profiles, exposure was rapidly recognised.

Acknowledgements

The authors thank Dr. T. van der Lende, E. Wesseling, M. Havekes, R.van de Kerkhof and Dr. F. Schuren for excellent expertise and setting up of the microarray facility. M. van den Wijngaard for assistance in sample isolation. We gratefully thank Dr. A. Smilde for helpful discussions on multivariate statistics.

References

1. Burczynski, M.E., McMillian, M., Ciervo, J., Li, L., Parker, J.B., Dunn, R.T. et al. (2000) Toxicogenomics-based discrimination of toxic mechanism in HepG2 human hepatoma cells. Toxicol. Sci. 58(2): 399-415.
2. Bulera, S.J., Eddy, S.M., Ferguson, E., Jatkoe, T.A., Reindel, J.F., Bleavins, M.R. et al. (2001) RNA expression in the early characterization of hepatotoxicants in Wistar rats by high-density DNA microarrays. Hepatology 33(5): 1239-1258.
3. Waring, J.F., Jolly, R.A., Ciurlionis, R., Lum, P.Y., Praestgaard, J.T., Morfitt, D.C. et al. (2001) Clustering of hepatotoxins based on mechanism of toxicity using gene expression profiles. Toxicol. Appl. Pharmacol. 175(1): 28-42.

4. Hamadeh, H.K., Bushel, P.R., Jayadev, S., Martin, K., DiSorbo, O., Sieber, S. (2002) Gene expression analysis reveals chemical-specific profiles. Toxicol. Sci. 67(2): 219-231.
5. Holmes, E., Nicholls, A.W., Lindon, J.C., Ramos, S., Spraul, M., Neidig, P. (1998) Development of a model for classification of toxin-induced lesions using 1H NMR spectroscopy of urine combined with pattern recognition. NMR Biomed. 11(4-5): 235-244.
6. Holmes, E., Nicholson, J.K., Tranter, G. (2001) Metabonomic characterization of genetic variations in toxicological and metabolic responses using probabilistic neural networks. Chem. Res. Toxicol. 14(2): 182-191.
7. Waters, N.J., Holmes, E., Waterfield, C.J., Farrant, R.D., Nicholson, J.K. (2002) NMR and pattern recognition studies on liver extracts and intact livers from rats treated with alpha-naphthylisothiocyanate. Biochem. Pharmacol. 64(1): 67-77.
8. Waters, M., Boorman, G., Bushel, P., Cunningham, M., Irwin, R., Merrick, A. et al. (2003) Systems toxicology and the Chemical Effects in Biological Systems (CEBS) knowledge base. EHP Toxicogenomics 111(1T): 15-28.
9. Beckwith-Hall, B.M., Nicholson, J.K., Nicholls, A.W., Foxall, P.J., Lindon, J.C., Connor, S.C. et al. (1998) Nuclear magnetic resonance spectroscopic and principal components analysis investigations into biochemical effects of three model hepatotoxins. Chem. Res. Toxicol. 11(4): 260-272.
10. Coen, M., Ruepp, S.U., Lindon, J.C., Nicholson, J.K., Pognan, F., Lenz, E.M., Wilson, I.D. (2004) Integrated application of transcriptomics and metabonomics yields new insight into the toxicity due to paracetamol in the mouse. J. Pharm. Biomed. Anal. 35(1): 93-105.
11. Thor, H., Svensson, S.A., Hartzell, P., Orrenius, S. (1981) Biotransformation of bromobenzene to reactive metabolites by isolated hepatocytes. Adv. Exp. Med. Biol. 136 Pt A(1): 287-299.
12. Casini, A.F., Pompella, A., Comporti, M. (1985) Liver glutathione depletion induced by bromobenzene, iodobenzene, and diethylmaleate poisoning and its relation to lipid peroxidation and necrosis. Am. J. Pathol 118(2): 225-237.
13. Monks, T.J., Hinson, J.A., Gillette, J.R. (1982) Bromobenzene and p-bromophenol toxicity and covalent binding in vivo. Life Sci. 30(10): 841-848.
14. Lau, S.S., Monks, T.J. (1988) The contribution of bromobenzene to our current understanding of chemically-induced toxicities. Life Sci. 42(13): 1259-1269.
15. Miller, N.E., Thomas, D., Billings, R.E. (1990) Bromobenzene metabolism in vivo and in vitro. The mechanism of 4-bromocatechol formation. Drug Metab. Dispos. 18(3): 304-308.
16. Heijne, W.H., Stierum, R.H., Slijper, M., van Bladeren, P.J., van Ommen, B. (2003) Toxicogenomics of bromobenzene hepatotoxicity: a combined transcriptomics and proteomics approach. Biochem. Pharmacol. 65(5): 857-875.
17. Lamers, R.A.N., DeGroot, J., Spies-Faber, E.J., Jellema, R.H., Kraus, V.B., Verzijl, N. et al. (2003) Identification of disease- and nutrient-related metabolic fingerprints in osteoarthritic Guinea pigs. J Nutr 133(6): 1776-1780.
18. Smilde, A.K. (2003) A framework for sequential multiblock component methods . J Chemometrics 17: 323-337.

9

Conclusions and perspectives

Biomarkers play an important role in biomedical research and will become more and more essential for the study of new intervention therapies in the near future. The approach to identify specific metabolite profiles for disease processes by using analytical techniques like NMR or LC-MS with subsequent MVDA is promising. The human body consists, amongst others, of numerous metabolites. Most of these act on a complex variety of processes that are related to each other. Therefore, generally it is not a single metabolite which acts as a biomarker but a combination of metabolites. The advantage of linking NMR and other analytical techniques to MVDA is the fact that the generated profiles consist of a combination of metabolites which together meet the criteria to serve as biomarker. Biomarker profiles can thus aid disease diagnosis, measurement of disease progression or tumor regression, and drug development.

In this thesis, several disease processes were investigated by applying NMR on biological fluids and analyzing the resulting data with MVDA in order to discover potential biomarkers. A metabolite profile was identified for OA in guinea pigs as well as in humans, which correlated with conventional histopathology. In addition, in vitro metabolite profiles were used to investigate the effects of inulin, a prebiotic, on gut health. A profile was discovered which could serve as early biomarker for acute rejection after kidney transplantation. Also, using biomarker profiles in a toxicological study, metabolite profiles were linked to gene expression data.

The results described in this thesis show that metabolic profiling using NMR enables to measure biological fluids without pretreatment in a quick manner. It provides an unbiased overview of many important metabolites present in biological samples and can be used in the search for new diagnostic biomarkers as well as early biomarkers.

Early biomarkers are of importance in the understanding of the origins of diseases or of environmental effects. Current medicine is mainly oriented at relieving symptoms of diseases, but early biomarkers could help to bring curing nearer. Time-course metabolomics could have potential for finding more early biomarker profiles. Much research still has to be done in this field, but it is definitely worth the effort. Yet, in general one should be aware of the fact that time-course biological studies are very expensive, time-consuming and make demands on an enormous amount of time on the analytical equipment, without a guarantee for a positive outcome. Only few organisations will take the risk and have the money to carry out such studies.

A topic that needs attention in metabolic profiling is the validity of the obtained results. MVDA is an exploratory way of data analysis. The obtained metabolic fingerprints are well suited for the generation of new hypotheses about disease-related metabolic pathways. However, more efforts will have to be made to investigate the validity and specificity of relations between metabolic profiles and diseases. Otherwise public health institutions will not accept the discovered profile as a valid biomarker for a disease. For early biomarkers this will even be more the case. If after a long time of research and spending of much money finally a potential early biomarker profile is found, one should be able to prove that it is related to a clinical endpoint of a disease.

For metabolic profiling to become a generally accepted tool to discover biomarkers, a deeper understanding of the metabolism that underlies the profile and its correlation to disease processes is needed, and also of the mechanisms of intervention. In vitro studies and translational science can facilitate the clarification of the identity of metabolites and the formulation of ideas about metabolic pathways involved. However, identification of metabolites is still the point where normally studies of metabolite profiling with NMR stagnate. Well-known metabolites like glucose or creatinine can be assigned quite easily. Yet, many compounds are still difficult to identify. The choice for assigning an NMR signal to a metabolite may be influenced by existing knowledge about metabolic pathways that is already known. However, there are still many metabolites and pathways in the body we do not know. The interpretation of signals in an NMR spectrum of a biological fluid is thus hampered by an incomplete knowledge of metabolism. For this reason,

identification of more signals in the profiles is indispensable. Intensive efforts should be made to identify metabolites in the profiles, using techniques like GC-MS, LC-MS and 2D-NMR. Besides, databases with known metabolites should be set up to facilitate the generation of hypotheses about metabolites. In addition, more in-depth studies should elucidate the metabolic pathways involved. This can be done when metabolite information is integrated with gene expression and proteomic data.

The key for metabolite profiling to become a successful way to discover biomarkers thus lies in a systems biology strategy. Integration of all types of biological information, such as DNA, RNA, protein and metabolites, and study of their mutual relationships to obtain a model of a biological system as a whole, should provide evidence that certain metabolite profiles are related to disease processes. Subsequently, intervention studies can be carried out that rely on these profiles. Therefore high-throughput facilities for genomics, transcriptomics, proteomics and metabolomics, computational infrastructure, development of tools to process, integrate and model all the obtained biological information, are important future developments. For such a systems biology approach to become successful, it is necessary that biologists, chemists, mathematicians, physicists and informaticians are teamed up together, in order to facilitate the development and integration of new technologies.

Still, looking at all system's elements, which is the true essence of systems biology, is sometimes hampered due to practical and ethical reasons. For instance, in studies with humans, it is often difficult, if not impossible, to obtain specific biofluid and tissue samples. In animal studies, it is often required that animals are sacrificed to obtain certain samples, which can be a severe limitation due to the fact that the biological process can not be followed in the same animal and the expense of studies is dramatically increased because large numbers of animals are needed. Functional imaging techniques circumvent the mentioned difficulties. Functional imaging enables to look in vivo in biological systems as a whole in a sensitive manner, even in tissues that are not suitable for sampling for systems biology studies. Therefore the use of functional imaging techniques like Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT) will be an important future development for systems biology in the

elucidation of the pathophysiology and etiology of disease processes and the discovery of new biomarkers.

Summary

Biomedical research aims at the elucidation of relationships between state of health and environmental factors, such as lifestyle, nutrition and pharmaceuticals. Biomarkers, biological indicators that can be used to measure and evaluate disease, disease risk and effects of exposure, will facilitate biomedical research. Metabolites are intermediate or end products of biological processes and thus play an important role as biomarkers. Nuclear magnetic resonance spectroscopy (NMR), gas chromatography-mass spectrometry (GC-MS) and liquid chromatography-mass spectrometry (LC-MS) are analytical techniques that enable the identification and measurement of metabolites in biological fluids. However, data obtained with these techniques is very complex, due to the vast amount of metabolites contained in biofluids. Therefore MVDA is needed to find differences and similarities in data obtained with NMR or LC-MS. When MVDA reveals clusters with similar characteristics in the analytical data, and these clusters match the original study set-up (the established healthy, treated or diseased subjects), a connection can be established between affected metabolites on one side, and treatment or disease on the other.

In this thesis, MVDA is applied to data obtained from NMR as a tool to select biomarker profiles in body fluids, that are specific for certain disease processes. These profiles can, for instance, be used for (early) diagnosis or to study effects of pharmaceuticals or food supplements.

In Chapter 2, general NMR theory is described jointly with a short introduction to GC-MS and LC-MS. In Chapter 3, an overview is presented of MVDA techniques that are used to identify profiles of biomarkers. Both chapters are intended to give the reader some general background about the analytical techniques that are used in the investigations reported in this thesis.

Chapter 4 reports a diagnostic biomedical study, in which it is shown that NMR with subsequent MVDA is a suitable tool to discriminate between healthy guinea pigs and guinea pigs suffering from osteoarthritis (OA). A diagnostic biomarker profile is presented for OA in guinea pigs and effects of vitamin C on the disease process

are investigated using this profile. Chapter 5 reveals that the biomarker profile observed in guinea pigs is also a valid diagnostic tool for human OA patients. In addition, it is demonstrated that the profile shows a correlation with histopathological data, the so-called Kellgren-Lawrence grade, which is a measure for OA state.

In Chapter 6, it is shown that NMR in conjunction with MVDA, apart from detection of diagnostic biomarker profiles, can also play an important role in disease prognosis. In a study with patients who underwent a kidney transplant operation, it is shown that patients with graft rejection have a biomarker profile, which can already be detected five days before rejection is observed with the conventional clinical techniques. This biomarker profile can thus be used as a prognostic tool.

The method is also employed in an in vitro study, which is described in Chapter 7. The effects of inulin, a prebiotic, on gut health are investigated. The results show that in vitro research can be a useful starting point for generating hypotheses about affected pathways and can thus facilitate in vivo studies.

In Chapter 8, the final part of this thesis, a step towards systems biology is made by linking results of metabolomics to gene expression. In this way, effects of bromobenzene in rats are studied to further understand hepatotoxicity and to get insight into metabolic pathways.

In conclusion, the approach to identify biomarker profiles for disease processes by means of NMR with subsequent MVDA is promising. Metabolite profiles have been identified for, amongst others, OA, hepatotoxicity and kidney rejection. However, to exploit the full potential of biomarker profiles, extensive efforts should be made to identify the metabolites in the profiles and to further understand the metabolic pathways involved. This will be enhanced and facilitated by a systems biology approach with the integration of genomics, transcriptomics, proteomics and metabolomics.

Samenvatting

Biomedisch onderzoek richt zich op het ophelderen van de relaties tussen ziekte en omgevingsfactoren, zoals bijvoorbeeld levensstijl, voeding en medicijnen. Biomarkers zijn biologische indicatoren die gebruikt kunnen worden om ziekte, kans op ziekte, en effecten van de omgeving, te meten en te bestuderen. Metabolieten zijn tussen- of eindprodukten van biologische processen en spelen als zodanig een belangrijke rol als biomarkers. Kernspinresonantie (NMR) en vloeistofchromatografie-massaspectrometrie (LC-MS) zijn analytisch-chemische technieken die het mogelijk maken metabolieten in biologische vloeistoffen te identificeren en meten. De gegevens uit deze metingen zijn echter zeer complex, doordat biologische vloeistoffen een enorme hoeveelheid metabolieten bevatten. Daarom is multivariate data analyse (MVDA) nodig om overeenkomsten en verschillen in data die met NMR en LC-MS wordt verkregen, te vinden. Wanneer uit MVDA blijkt dat de analytische data in een aantal groepen met een grote overeenkomst in kenmerken uiteenvalt, en deze clustering komt overeen met de originele proefopzet (zoals een gezonde groep, een behandelde groep of een zieke groep), dan kan een link worden gelegd tussen metabolieten die zijn veranderd in kwantiteit enerzijds, en ziekte of gezondheid anderzijds.

In dit proefschrift wordt MVDA toegepast op data die wordt verkregen uit NMR. Op deze wijze worden in lichaamsvloeistoffen biomarkerprofielen geselecteerd die specifiek zijn voor bepaalde ziekten. Deze profielen kunnen bijvoorbeeld gebruikt worden om (vroege) diagnostiek te plegen of effecten van bijvoorbeeld medicijnen of voedingssupplementen te onderzoeken.

In hoofdstuk 2 wordt een samenvatting gegeven van de theorie achter NMR, tezamen met een korte introductie in GC-MS en LC-MS. In hoofdstuk 3 worden diverse MVDA technieken behandeld, die gebruikt kunnen worden om biomarkerprofielen te identificeren. Beide hoofdstukken zijn bedoeld om de lezers achtergrondinformatie te verschaffen over de analytische technieken die worden toegepast bij het onderzoek beschreven in dit proefschrift.

Hoofdstuk 4 beschrijft een diagnostische biomedische studie, waarin wordt aangetoond dat NMR samen met MVDA een geschikte methode is om, op basis

van verschillen in urinesamenstelling, gezonde cavia's te onderscheiden van cavia's die lijden aan artrose. Een diagnostisch biomarkerprofiel voor artrose in cavia's is daarbij gevonden. Daarmee is een interventiestudie uitgevoerd met vitamine C, een stof die mogelijk positieve effecten op het ziekteproces heeft. In hoofdstuk 5 wordt daarna getoond dat het biomarkerprofiel voor artrose in cavia's ook wordt gezien in urine van mensen. Dit profiel vertoont bovendien correlatie met histopathologische data.

De resultaten in hoofdstuk 6 laten zien dat de methode ook geschikt is om prognostische biomarkerprofielen te identificeren. Uit een studie met patiënten die niertransplantatie hadden ondergaan, blijkt dat patiënten die acute afstotingsverschijnselen vertoonden een afwijkend urine patroon hebben. Dit afwijkende profiel kon al worden waargenomen vijf dagen voordat dit met conventionele technieken kan worden gedetecteerd.

De resultaten in hoofstuk 7 tonen dat toepassing van de techniek op in vitro studies interessante hypotheses kan opleveren over betrokken metabole paden. Bovendien kan in vitro onderzoek de daaropvolgende in vivo studies vergemakkelijken.

In hoofdstuk 8 wordt een opstap gemaakt naar een aanpak vanuit de systeembiologie. Resultaten van gen expressie en metabolomics worden gekoppeld, met het doel levertoxische effecten van broombenzeen in ratten te bestuderen en de betrokken metabole paden te doorgronden.

Concluderend kan gezegd worden dat de aanpak om biomarker profielen te identificeren met behulp van NMR en LC-MS met daaropvolgend MVDA, veelbelovend is voor het herkennen van biomarkerprofielen. In het onderzoek dat in dit proefschrift is beschreven, zijn metabolietprofielen gevonden voor onder andere artrose, hepatotoxiciteit en nierrejectie. Om het potentieel van de techniek volledig te benutten zal nog veel inspanning moeten worden gestoken in het identificeren van de metabolieten in de profielen, en in het doorgronden van de betrokken metabole paden. Een aanpak vanuit de systeembiologie, waarbij resultaten van onderzoek naar DNA, gen-expressie, eiwitten en metabolieten worden geïntegreerd, zal dit proces versnellen en vergemakkelijken.

Nawoord

Een proefschrift kan alleen maar tot stand komen met de hulp van heel veel mensen. Op deze plaats wil ik deze personen dan ook graag vermelden.
Elly en Gerwin, jullie stonden altijd voor mij klaar, zonder jullie hulp had dit proefschrift er heel anders uit gezien. Jeroen, Nicole en Johan, de discussies met jullie waren inspirerend en leverden interessante ideeën op. Wilbert, de samenwerking met jou was erg plezierig. De stagiairs Marc en Eveline, jullie onbevangenheid heeft ervoor gezorgd dat jullie stage onderzoeken publicaties hebben opgeleverd. Johan de Fijter, je hebt mij wegwijs gemaakt in het nier-transplantatie onderzoek.
De patroonherkenners Henk, Bianca, Jack, Sabina, Renger, Florian, ieder van jullie heeft op zijn eigen manier inbreng in dit proefschrift gehad. Mijn eerste kamergenoot, Albert: je positivisme (“Uitstekend idee, goed werk”) werkte aanstekelijk. Mijn daaropvolgende kamergenoten (in chronologische volgorde) Elly, Jacques, Marco, Bianca, Gerwin, Henk en Wilbert, jullie hebben voor de gezelligheid gezorgd. Daarnaast ben ik Henk erkentelijk voor het niet halen van koffie voor mij. Valentijn, kamergenoot voor 1 dag in de week: het combineren van een promotie-onderzoek met een (aanstaand) vaderschap heeft voor heel wat gesprekstof gezorgd. De dagen dat jij in Zeist was, kwam er van werken over het algemeen weinig terecht.
Tot zover de mensen die direct bij mijn promotie-onderzoek betrokken waren. Daarnaast zijn er veel mensen indirect betrokken geweest bij mijn reilen en zeilen. Zo toonden de collega’s van de Genomics-groep, vrienden en (schoon)familie altijd hun belangstelling. De volgende mensen wil ik hier speciaal nog vernoemen:
Ineke, je weet het misschien niet maar jij bent de aanstichtster van het geheel. Jij hebt me op 16-jarige leeftijd meegenomen naar de chemische fabrieken van Shell in Pernis. Hier was ik zo van onder de indruk dat ik besloot om scheikunde te gaan studeren.
Monique, we hebben in mijn eerste jaren bij TNO vaak bij elkaar in de bus gezeten, de ritjes met lijn 50 van Utrecht CS naar Zeist v.v. waren altijd gezellig en een goede gelegenheid om mijn ei even kwijt te kunnen.

Voordat ik bij TNO kwam werken, heb ik een drietal jaren aan de Universiteit Utrecht gewerkt. Het AMS-groepje met Arie, Cees en Wybe stond garant voor veel gezelligheid, humor en niet te vergeten…. thee. Cees, je hebt me geholpen door dit proefschrift te becommentariëren en je bleef me lastigvallen met emails over mannen in witte jassen, NEC-bronnen en andere onzin. Wybe, altijd even attent en belangstellend, ik realiseer me dat promoveren maar een bijzaak in het leven is. Arie, jij zag mijn mogelijkheden en stimuleerde me om na te denken over hoe ik deze kon benutten. Jij hebt dan ook de kiem voor dit proefschrift gelegd. Daarnaast heb je er voor gezorgd dat dit proefschrift een stuk leesbaarder is geworden. Leuk dat je me als paranimf terzijde wilt staan.

Wim, het begon in Breda met een introductiekamp waarna we als practicummaatjes verder gingen. We belandden toevalligerwijs allebei via Eindhoven in Arnhem. Inmiddels zijn we vele jaren van verhuizingen (vooral van mijn kant), hoogtepunten en wedstrijden van Vitesse verder en ben ik blij dat je mijn paranimf wilt zijn.

Monique, broer en zus is als trein en bus. Francesco, *tu sei come un fratello per questa testa di formaggio*.

Mam, ik wil je bedanken voor alles maar vooral voor de mogelijkheden en de vrijheid die je samen met pap aan mij hebt gegeven. Jullie grenzeloze vertrouwen dat ik hiermee om kon gaan, heeft ervoor gezorgd dat ik onafhankelijk mijn eigen kronkelige weg heb kunnen bepalen. Dat heeft uiteindelijk tot dit boekje geleid.

Pap, je wist dat ik dit boekje aan het schrijven was maar helaas is de afronding ervan voor jou te laat gekomen. Ik had dit moment erg graag met je willen delen. Ik weet dat je trots op me zou zijn geweest.

Annette, ons leven lijkt continu in de hoogste versnelling te lopen. Door jouw liefde, steun en nuchterheid kan ik het bijbenen. Als je dat maar weet en nooit vergeet.

Ten slotte Annika: vader is de mooiste titel die er is.

Curriculum vitae

Robert-Jan Antonius Nicolaas Lamers was born on 18 March 1971 in Arnhem, the Netherlands. In 1990 he completed his secondary school education (VWO) at the Thomas à Kempis College in Arnhem. In September of the same year, he started a study Chemical Engineering at the Hogeschool West-Brabant in Breda, finishing his bachelors degree in June 1994. From 1994 until 1996 he studied Chemical Engineering at the Technical University of Eindhoven, finishing his masters degree. During this study, he obtained research experience in chromatography and mass spectrometry at the Department of Instrumental Analysis of the faculty of Chemical Engineering, Technical University of Eindhoven (supervisors Prof.dr.ir. C.A. Cramers, Dr. J.G.M. Jansen and Dr. H. A. Claessens) and at the Department of Analytical Chemistry for Development, AKZO Nobel, N.V. Organon (supervisors Dr. R.J.M. Vervoort and Dr. A.J.J. Debets).

Thereafter, he worked at the accelerator mass spectrometry facility of the Subatomic Physics Department, Faculty of Physics and Astronomy, of the University of Utrecht until September 2000. Meanwhile, he followed a course Biochemistry at the Open University in Utrecht. In October 2000 he started working at TNO as junior project manager in the department of Analytical Sciences. In April 2004, he continued his work at TNO as research scientist in the department of Physiological Sciences. From January 2001, under the supervision of Dr. J.H.J. van Nesselrooij and Prof.dr. J. van der Greef the work described in this thesis was performed.

List of publications

1. Lamers, R.A.N., DeGroot, J., Spies-Faber, E.J., Jellema, R.H., Kraus, V.B., Verzijl, N., TeKoppele, J.M., Spijksma, G., Vogels, J.T.W.E., van der Greef, J., van Nesselrooij, J.H.J. (2003) Identification of disease and nutrient related metabolic fingerprints in osteoarthritic guinea pigs. J. Nutr., 133, 1776-1780.

2. Lamers, R.A.N., van Nesselrooij, J.H.J., Kraus, V.B., Jordan, J.M., Renner, J.B., Dragomir, A.D., Luta, G., van der Greef. J., DeGroot, J. Identification of an urinary metabolite profile associated with osteoarthritis. Submitted for publication.

3. Lamers, R.A.N., van Nesselrooij, J.H.J., Seelen, M.A.J., de Fijter, J.W., Cohen, A.F., van der Greef, J. Urinary metabolite profile as early biomarker for kidney rejection after transplantation. In preparation for publication.

4. Lamers, R.A.N., Wessels, E.C.H., van de Sandt, J.J.M., Venema, K., Schaafsma, G., van der Greef, J., van Nesselrooij, J.H.J. (2003) A pilot study to investigate effects of inulin on Caco-2 cells through in vitro metabolic fingerprinting. J. Nutr., 133, 3080-3084.

5. Heijne, W.H.M., Lamers, R.A.N., van Bladeren, P., Groten, J.P., van Nesselrooij, J.H.J., van Ommen, B. Profiles of metabolites and gene expression in rats with chemically induced hepatic necrosis. Accepted for publication in Tox. Path.

6. Boelsma, E., Lamers, R.A.N., Hendriks, H.F.J., van Nesselrooij, J.H.J., Roza, L. (2004) Evidence of the regulatory effect of *Ginkgo biloba* extract on skin blood flow and study of its effects on urinary metabolites in healthy humans. Planta Med., 70, 1052-1057.

7. Venema, K., van Nesselrooij., J.H.J., Lamers, R.A.N. (2003) Metabonomics of Caco-2 cells; effect of inulin and its fermentative metabolites. Nutrafoods, 2, 5-12.

8. De Roos, B., Duivenvoorden, I., Rucklidge, G., Pickard, K., Lamers, R.A.N., Voshol, P.J., Havekes, L.M., Teusink, B. Response of apolipoprotein E*3-Leiden transgenic mice to dietary fatty acids: combining liver proteomics with physiological data. Accepted for publication in FASEB J.

9. Wang, M., Lamers, R.A.N., Korthout, H.A.A.J., van Nesselrooij, J.H.J., Witkamp, R., van der Heijden, R., Verpoorte, R., van der Greef, J. Metabolomics in the context of systems biology: bridging traditional Chinese medicine and molecular pharmacology. Accepted for publication in Phytother. Res.

10. Lamers, R.A.N., Renkens, M.F.J., van Nesselrooij, J.H.J., Woutersen, R.A., Appel, M.J. Metabolite profiles associated with the early development of pancreatic cancer in rats. In preparation for publication.

List of abbreviations

ALAT, alanine amino transferase;
ANIT, alpha-naphthylisothiocynanate;
ASAT, aspartate amino transferase;
BB, bromobenzene;
BMI, body mass index;
D, discriminant;
DMEM, Dulbecco's modified Eagle medium;
ESI, electrospray ionization;
FID, free induction decay;
GC, gas chromatography;
GSH, glutathione;
HFHC, high fat and high calorie;
KL, Kellgren-Lawrence;
LC, liquid chromatography;
mRNA, messenger RNA;
MRI, magnetic resonance imaging;
MS, mass spectrometry;
MSCA, multilevel simultaneous component analysis;
MVDA, multivariate data analysis;
NMR, ^{1}H nuclear magnetic resonance spectroscopy;
NO, nitric oxide;
n.s., not statistically significant,
OA, osteoarthritis;
OSC, orthogonal signal correction;
PBS, phosphate buffered saline;
PC, principal component;
PCA, principal components analysis;
PCDA, principal components discriminant analysis;
PLS, partial least squares;
ppm, parts per million;
SCFA, short-chain fatty acids;
TCM, traditional Chinese medicine;
TIM, TNO's intestinal model;
TMAO, trimethylamine N-oxide;
TMSP, sodium trimethylsilyl-[2,2,3,3,-$^{2}H_4$]-1-propionate.