The handle [http://hdl.handle.net/1887/20304](http://hdl.handle.net/1887/20304) holds various files of this Leiden University dissertation.

**Author:** Wietmarschen, Herman van  
**Title:** A systems approach to sub-typing of rheumatoid arthritis  
**Date:** 2012-12-18
Abstract

Metabolites have played an essential role in our understanding of life, health, and disease for thousands of years. This domain became much more important after the concept of metabolism was discovered. In the 1950s, mass spectrometry was coupled to chromatography and made the technique more application-oriented and allowed the development of new profiling technologies. Since 1980, TNO has performed system-based metabolic profiling of body fluids, and combined with pattern recognition has led to many discoveries and contributed to the field known as metabolomics and systems biology. This review describes the development of related concepts and applications at TNO in the biomedical, pharmaceutical, nutritional, and microbiological fields, and provides an outlook for the future.

Introduction to Metabolomics

Metabolites in ancient history

The word *metabolism* originates from the Greek “μεταβολισμός”, which means “change”. The concept of metabolism was mentioned by Ibn al-Nafis (1213-1288), who stated that “the body and its parts are in a continuous state of dissolution and nourishment, so they are inevitably undergoing permanent change.” Studies on individual changes during different daily activities were performed by Santorio in 1614 and were mentioned in his book *Ars de Statica Medecina* (Bing, 1971).

The measurement of single metabolites as a source of information related to health and disease has a long history that precedes the introduction of metabolomics and acronyms such as metabolomics, body fluid profiling, metanomics, metabonomics, metabolic profiling, etc. (Oliver et al., 1998; Lindon, Holmes, & Nicholson, 2003; Raamsdonk et al., 2001; Ramsden, 2009). The terminology is not relevant for the content. As stated by Juliet in Shakespeare’s Romeo and Juliet, "What's in a name? That which we call a rose by any other name would smell as sweet" (Shakespeare, 1594). From another perspective, metabolites have a long association with sweet flavors that extends to ancient times and predate the development of metabolic nomenclature. Ancient Chinese cultures (1500–2000 BC) recognized urine as an important source of health-related information and sweet-tasting urine as indicative of a disease (now known as diabetes). At that time, “clinical testing” involved actual urine tasting. Advanced “biosensor” ants could also be used to test differences between sample and reference urines. The association between sweet urine and disease was contemporaneously made in India by Ayurveda Hindus.

Diabetes as a disease state was described on 3rd Dynasty Egyptian papyri by the physician Hesy-Ra, who mentioned an additional symptom, frequent urination or polyuria. Approximately 1000 years ago, the Arabian physician Avicenna observed that an individual’s urine changes during illness. In modern times, changes in the smell or color of urine are known to be related to changes in the concentration of chemical components and dysregulation of biochemical pathways that indicate certain metabolic diseases. For instance, blue urine can indicate Blue Diaper syndrome, caused by a defect in tryptophan absorption. Urine with a musty/mousey odor is suggestive of classical phenylketonuria (PKU), an autosomal recessive metabolic genetic disorder.
The discovery of enzymes by German chemist Eduard Buchner (1860–1917) at the beginning of the 20th century led to a focus on intracellular chemical reactions and inspired development of the field of biochemistry. The biochemical understanding of metabolism developed rapidly due to new insights into enzymatic reactions and intracellular biochemical pathways.

New technological developments in the early 20th century in the domain of mass spectrometry (MS) enabled researchers to measure the molecules involved in biochemical pathways, and to investigate their roles in disease states. As early as 1948, Williams and his associates identified important concepts of normality based on MS profiling of body fluids, and examined individual differences and pathologies in patients with alcoholism and schizophrenia. The correlation of emotional stress and physical exertions with urinary metabolite profiles were already emphasized by Ludwig et al. (Ludwig 1977). Gates and Sweeley performed a review of these early thoughts and experimental approaches (Gates & Sweeley, 1978). In their review, they mentioned some important hallmarks related to quantitative metabolic profiling. In particular, they referenced the works of Horning & Horning, who first introduced the concept of metabolic profiling by MS (Horning & Horning, 1971), and Pauling et al., who reported urine vapor and breath analyses with gas-liquid partition chromatography related to the effects of defined diet (Pauling et al., 1971).

To enable the routine measurement of molecules in biological matrices, researchers in the 1950s coupled gas chromatography (GC) with MS. Early commercial GC/MS instruments enabled practical clinical investigations and the measurement of profiles of urine and blood samples. Such profiles were typically limited to a specific class of compounds, such as organic acids, because the available technology required volatile components or compounds made volatile by chemical derivatization.

Opportunities for the direct application of GC/MS techniques came in the early 20th century. The British physician Sir Archibald Garrod (1857–1936) proposed the detection of changes in metabolic pathways caused by a single inherited gene defect. This concept became the basis for the field of inborn errors of metabolism. In clinical chemistry, body fluid profiles led to the discovery of numerous diseases related to single-gene defects (for more information, see reviews by Politzer, Dowty & Laseter, 1976, and Jellum, 1977). However, these types of
profiles were very limited from a systems perspective. Concentration changes in such genetic diseases are typically very large and more easily detected than changes in chronic diseases, in which time-dependent changes of regulatory processes must be observed. Consequently, there was a need for a broader characterization of systems and for sophisticated pattern-recognition tools to evaluate complex patterns.

In the 1970s, several developments occurred in The Netherlands that influenced research at TNO. Meuzelaar and Kistemaker developed an impressive pyrolysis-MS (PyMS) characterization methodology that allowed the pyrolyzation of bacteria and many other nonvolatile biomaterials (Meuzelaar & Kistemaker, 1973). The resulting constituents were recorded as fingerprints and analyzed with multivariate statistical methods, such as nonlinear mapping and factor analysis (Windig, Kistemaker & Haverkamp, 1980). New multivariate tools became available in the field of GC profiling (Blomquist et al., 1979; McConnell et al., 1979; Rhodes et al., 1981). Although PyMS was a novel systems profiling method, it had some limitations: typically, only nonvolatile materials could be investigated, and thermal degradation products were not easily related to their precursor macromolecules. Despite these challenges, many successful investigations were achieved with PyMS in subsequent decades.

In the 1970s, the MS field expanded with the development of soft-ionization methods, such as field desorption and laser desorption (Schulten & Beckey, 1975) that enabled the analysis of low-volatile components. As part of his PhD project, under the guidance of Nico Nibbering at the University of Amsterdam, new electronics were developed to enable controlled desorption and reproducible biofluid profiling through field ionization kinetics/field desorption (van der Greef, 1980). At the same time, new multivariate statistical tools became available in commercial packages such as Arthur (CPC, Seattle, WA, USA). However, interpretation of outputs from these tools was not straightforward; this time-period marked the beginning of the analysis of datasets with many more variables than objects. To illustrate this perspective, after a night of intense calculations, the Arthur program was known to return the following famous output: “In controversial matters, my perception is rather fine. I always see both points of view: the one that’s wrong and mine.”
**Metabolomics at TNO**

*The body fluid profiling and pattern recognition project (early 1980s)*

The developments of field desorption (FD) and the so-called emission-control device (van der Greef, 1980) overcame a major bottleneck in reproducible profiling of biofluid samples under controlled desorption conditions and presented a unique opportunity for researchers. In August 1980, the “body fluid profiling and pattern recognition” project was initiated at TNO, led by Jan van der Greef in the Instrumental Analysis group. Headed by Michael ten Noever de Brauw, these facilities held some of the most advanced MS capabilities worldwide at that time. The initial phase of the project involved construction of an ion source and an electronic control unit, equipped with photographic and electric detection, for the Varian 731 MS instrument. The design of the control unit was based on the original electronic design of Jim Dawson (University of Amsterdam).

The initial phase involved implementation of earlier designs, as well as production and successful operation of an automatic field-desorption emitter-production unit. The most complicated step was to interface the MS system with the PDP-8 computer because the high-voltage field-desorption ion source generated frequent sparking events. Software was developed to acquire, normalize, and import data to the Arthur software package for multivariate analysis. The multidisciplinary core team for this project included Jan van der Greef, Leo Bergman (electronics), Jaap Bouwman (computer hardware and software), Albert Tas (pattern recognition), and Michael ten Noever de Brauw (MS). An MS machine equipped with a field desorption source and emission-control electronics (van der Greef & Nibbering, 1977) coupled to computer data acquisition (Varian 620i, SS100 datasystem) with the Arthur package became operational in mid-1981 (Figure 1). This setup would be the basis for further technological developments at TNO in the years to come.
Figure 1. The first “body fluid & profiling” (metabolomics) instrument constructed at TNO in 1981. The instrument consisted of a Varian MAT 731 mass spectrometer equipped with a field desorption source and coupled to a 12-bit PDP-8 DEC computer, operated by Jan van der Greef (photo taken by M.C. ten Noever de Brauw).

The first important experiment performed with this system was measurement of human gender differences (van der Greef et al., 1983) with the unsupervised multivariate analysis of urine profiles in late 1981 (Figure 2). Clear differences between urine samples from different genders were observed. These differences were related, in part, to steroid conjugates. This result was later confirmed with high-resolution measurements with multichannel-analyzer methodology.

A major challenge presented by this system was desalting of body fluids and removal of abundant components (urea, etc.) while capturing as many components as possible. After investigating various options, it was decided that it was not feasible to use a single method to profile small molecules in body fluids. Thus, a multiplatform approach was developed. An additional challenge in analysis was normalization, and an inclusion/ exclusion approach (Fischer weighting) for MS peaks was chosen. Researchers recognized the complexity of datasets with many more variables (300-500) than objects (12-40), and the danger of using Fisher weighting (i.e., class information as the preselection filter). Validation procedures also
needed to be implemented. A practical solution to validate findings of this, and later, discriminant approaches, such as principal components discriminant analysis (PC-DA), was to classify unknowns or to use cross-validation to obtain a step towards validation. Soft independent modelling of class analogy (SIMCA) was evaluated for building models (Droge et al., 1987); however, at that time, the program was limited to the developmental stage.

Figure 2. The first “metabolomics” outcome of an experiment performed in 1981. This experiment detected gender differences due, in part, to steroid conjugate profiles in human urine as analyzed with field desorption mass spectrometry combined with principal component analysis.

The fast atom bombardment (FAB) soft-ionization technique (Morris 1981) became available in the early 1980s. Although faster and easier than other methods, the routine application of FAB was limited because the robustness of FAB profiles was low due to background desorption matrices. FD and FAB were compared for metabolomics analysis with pattern recognition. The results were presented during the first International Chemometrics meeting in 1982 in Petten, The Netherlands, and published in Analytica Acta in 1983 (van der Greef et al., 1983).
Interestingly, the 1982 Chemometrics meeting focused on discussion of significant within-group variations, so-called interindividual differences. The discussion highlighted one of the strongest aspects of metabolomics as a phenotyping methodology, the potential for personalized medicine/health approaches.

In 1991, Albert Tas developed direct chemical ionization (DCI)/MS as part of his PhD research (Tas, 1991). This robust method combined the soft ionization of volatile and low-volatility/nonvolatile components (metabolomics part) with pyrolysis of biopolymer structures. This method became a routine methodology used in parallel with field desorption MS, and in the late 1980s it became the first method of choice until GC, GC/MS, and liquid chromatography (LC)/MS technologies were developed.

*Technology-driven metabolomics research at TNO (1980–2000)*

Figure 3 describes major metabolomics developments at TNO in the period 1980–2000. Research in this time frame was driven by technological development, combined with exploration of potential applications for these technologies.

Despite the experimental challenges of metabolomics profiling with field desorption MS, some intriguing concepts were developed in the early 1980s. Fast drug-metabolite detection was developed with the quotient spectra to correct for individual differences in backgrounds (van der Greef et al., 1984). The ratio of the metabolomic profiles before and after drug administration to rats was calculated, and the quotient spectra were statistically evaluated to improve drug-metabolite detection.
Figure 3. Metabolomics events with an impact on technology and applications from 1980 to 2000.

Interestingly, the approach revealed that levels of non drug-related compounds changed significantly. This concept underlies the contemporary use of biomarkers to assess system response. The system-response approach corrects for individual differences in homeostasis and focuses on changes, for instance in challenge tests. However, at that time of its development, this approach was not of major interest to the pharmaceutical industry because it revealed nontargeted mechanisms that were considered irrelevant. Investigations into new phenomena were believed to slow down the drug-development process. In hindsight, this approach provides a very relevant source of important information, not only for the discovery of drug mechanisms and off-target effects, but especially for recently very important systems toxicology effects. COMET I and II were important recent metabolomics projects in the area of drug safety as well as the US Critical Path Initiative and MetaMap Tox project run by BASF.

Interesting applications of TNO projects from 1980 to 2000 covered a wide range of topics, including microbiology related to food safety (bacteria, fungi, and algae), food production (barley, juice adulteration), feed, fermentation, doping, forensic toxicology, chemical industry, and pharmaceutical/ biomedical research. Microbiology was the first important field
that used Py-DCI/MS, and achievements of this research are summarized in a review (Tas, 1995). Other relevant reviews are indicated in Figure 3.

The concept of reversed pharmacology/discovery was first described in 1989, based on a study that revealed the working mechanism of Lombazole inhibition of the yeast-hyphal conversion of the dimorphic fungus Candida albicans (Tas et al., 1989a). Profiling combined with pathway analysis allowed the targeted identification of the drug, as described in detail in a later review (van der Greef et al., 2003). This is a remarkable approach that, in the last 10 years, has formed the strategy of metabolomics-driven system research to identify multitarget pharmacology of complex herbal medicines. Metabolic footprinting, identification of biomarkers in supernatants of virus-infected cells, was another noteworthy tool developed at TNO (Tas et al., 1989b; Kell et al., 2005).

The use of nuclear magnetic resonance (NMR) technology for metabolomics became an option after an internal reorganization at TNO in 1990, when the NMR group was moved into the MS department. The first NMR-profiling results combined with pattern recognition were published in 1989 by Jeremy Nicholson’s group (Bell, Brown & Sadler, 1989; Nicholson & Wilson, 1989). Since then, Nicholson’s work has inspired the field of NMR-based metabolic profiling combined with pattern recognition globally in many application domains. At the turn of the century, Jack Vogels developed a new partial-linear fit algorithm for NMR data preprocessing to avoid unnecessary resolution loss due to the binning process. The NMR applications described by Vogels included characterization of wine, coffee, illegal use of growth promoters, biomarkers in multiple sclerosis, and hop suspension culture results, among others (Vogels, 2002). The NMR field at TNO remained important until about the year 2000, due to the high reproducibility and quantitative aspects of NMR. However, interest in this method dropped with the development of systems biology and the limited dynamic range of detected compounds with NMR. Although the concentration sensitivity limits of NMR result in a limited number of metabolites that can be monitored, the high concentration metabolites, including housekeeping related molecules, may give an important view on high level system organization and might be very relevant especially when applied to prevention or early onset processes (Nicholson 2008). The focus in these areas is more on the balance between various regulatory processes than on detailed underlying metabolic processes.

After the appointment of Jan van der Greef as a professor at Leiden University in 1986, several new collaborations were initiated between TNO and Leiden University, specifically in the field of separation sciences and the domain of novel analytical MS techniques. A
landmark was development of single-cell metabolomics, which used peptide profiling by the matrix-assisted laser desorption/ionization (MALDI)-MS of single neurons in Lymnaea stagnalis (Jiménez 1994). This approach was used to identify proopiomelanocortin-processing products in melanotrope cells of the pituitary intermediate lobe of Xenopus laevis (Vázquez-Martínez 1999, Jespersen 1999, van Strien 1996).

An important project in the development of metabolomics field was the PhD project of Elwin Verheij (Verheij, 1993), the results of which formed a sound basis for unique LC/MS expertise. The combined use of LC/MS with field desorption was explored in the plant metabolomics profiling of the purgant principles of capsicum (van der Greef et al., 1985). Because TNO functioned as a testing location for new MS instruments produced by Varian MAT, many new technologies passed through the TNO MS group. In particular, field desorption MS was combined with LC/MS as early as 1984 with the moving-belt interface technology (van der Greef et al., 1984). The moving-belt interface LC/MS required substantial improvement for bioanalytical applications, and it was too limited for routine analysis and metabolomics-like applications.

The LC/MS technique became more applicable to metabolomics after introduction of thermospray (Heeremans, 1990), continuous-flow FAB LC/MS (Kokkonen, 1991), and electrospray combined with LC or capillary electrophoresis (Mazereeuw, 2000). Electromigration techniques such as micellar and chiral capillary electrophoresis and isotachophoresis were explored as separation and sample (stacking) pretreatment technologies (Storms et al., 2004). However, in spite of the attractiveness of electromigration methodologies, few metabolomics applications resulted from them so far, because of their limited sensitivity and lack of robustness in bioanalysis. A recent exception is Human Metabolome Technologies Japan which started to offer a CE-MS metabolomics service.

From 1980 onwards, many technological variations of GC/MS became available at TNO. The main limitation to the use of GC/MS and GCxGC/MS was (and continues to be) data handling. The reader is referred to a 2005 review that discusses the challenges involved in the symbiosis of mass metabolomics and chemometrics (van der Greef & Smilde, 2005). Of note, the step from direct profiling with field desorption and DCI/MS to separation techniques coupled with mass spectrometry (LC/MS and GC/MS) introduces the need for automatic corrections for peak shifts in the chromatography, baseline drifts, etc. The real application of chromatography-MS in metabolomics only became possible after software was developed that allowed automatic corrections for the above-mentioned issues. Solutions for these issues
were first developed in the late 1990s (Gaspari et al., 2001).

**Biology-driven metabolomics research at TNO (2000–present)**

Figure 4 describes major events in metabolomics development at TNO in the period 2000–present. From about 2000 forward, the major driver of metabolomics research at TNO has been biology, although technological developments have remained of major importance.

![Figure 4](image)

**Figure 4.** Metabolomics events from 2000 onwards. Reviews are divided into food (F), pharma (P), microbiology (M), Chinese medicine (C), and data analysis (D).

**Systems biology at TNO**

A working definition of systems biology in the year 2000 was “the study of biology as an integrated system of genetic, protein, metabolite, cellular, and pathway events that are in flux and interdependent” (Davidov et al., 2003). The concepts of systems diagnosis and systems intervention had been present at TNO from its inception.

The major drive for improvement of metabolomics research at TNO found its basis in a November 1999 discussion between Fred Regnier and Jan van der Greef on the future of biological and pharmaceutical research. Inspired by the development of first-generation proteomics tools at Purdue University, Regnier and van der Greef envisioned the integration of data at the transcript, protein, and metabolite levels. They shared this idea with Noubar
Afeyan of Flagship Ventures, and a business plan for integrative omics research was developed. In 2000, the first systems biology company in the world, Beyond Genomics [known later as BG Medicine (BGMD)], was born. In the year 2000, the famous Systems Biology Institute was founded by Leroy Hood. Research at this institute focused on systems modeling, but modeling at the metabolite level was not included in its analytical capabilities. TNO served as the metabolomics research and development component of BGMD. Through this company, TNO collaborated in systems-biology projects with the pharmaceutical industry. In particular, the apolipoprotein E 3 (APOE3) Leiden transgenic mouse model was developed for studying atherosclerosis and as an initial proof-of-principle experimental model for systems biology. The results of the combined transcriptomics, proteomics, and metabolomics analyses of this transgenic model, which were based on and integrated with correlation networks, became available as early as 2001. This project was the first integrated in vivo systems biology experiment in the world of its kind.

To date, few published studies are available that measure all three biological levels in a single model with integration by a de novo correlation network. Most systems-biology studies are based on measurement of one or (at most) two levels (i.e., gene, transcript, or metabolite level), and literature data that concerns pathways are added to these findings. The results of the APOE3 Leiden transgenic mouse model study required substantial effort to publish because reviewers were very reluctant to accept this new approach. However, two related papers were finally published in 2004 (Clish et al., 2004; Oresic et al., 2004). In the meantime, various projects with the pharmaceutical industry delivered unique data, although most of these results had to remain confidential.

A 2005 review in Nature Reviews Drug Discovery (van der Greef & McBurney, 2005) explained the concept and thinking behind systems pathology and pharmacology. Several other reviews and opinion papers described the general concept of systems thinking (Morel et al., 2004; van der Greef, Stroobant & van der Heijden, 2004), (reversed)-translational research (van der Greef, 2005; van der Greef et al., 2006), systems toxicology (Heijne et al., 2005b), and connecting multicompartment omics (van der Greef, Hankemeier & McBurney, 2006; van der Greef et al., 2007).

The potential for systems biology in nutrition was recognized very early, as shown by the extensive literature describing this view from 2001 onwards (van der Werf et al., 2001; van Ommen & Stierum, 2002; van Ommen, 2004; Gibney et al., 2005; Corthésy-Theulaz et al., 2005; Joost et al., 2007; van Ommen, 2007; van Ommen et al., 2009). Ben van Ommen has
been the nutritional systems biology leader at TNO since 2001. In 2004, TNO coestablished and coordinated the European Nutrigenomics organization (NuGO), which has been instrumental to develop nutritional metabolomics technology (Kaput et al., 2005, 2006; Scalbert et al., 2009), concepts (Daniel et al., 2008; van Ommen et al., 2008a, 2008c, 2010a) and proof-of-principle studies (Cavaliere et al., 2009). This work became the basis of the nutrigenomics-based NuGO consortium (Baccini et al., 2008; Harttig et al., 2009). A key feature of nutritional metabolomics is the concept of the “biomarker of health,” which is based on quantification of the stress-response curve after perturbation of homeostasis (van Ommen et al., 2009; Wopereis et al., 2009).

**Metabolomics applications within Chinese medicine**

Because systems biology encompasses a holistic perspective, the idea to use traditional Chinese medicine (TCM) as an entry to systems knowledge was born in 2002. With his background in natural product bioactive component detection in the spin-off company Screentec (now Kiadis) from his Leiden University group, Jan van der Greef arranged to meet with Mei Wang, a molecular biologist with knowledge of TCM and a strong network in China. In 2003, a lecture tour was organized to explain the new systems biology tools at TNO within the context of TCM. This collaboration created a strong link with the Chinese Academy of Sciences. Research was initiated at TNO to identify and elucidate the actions of synergetic active compounds (multidimensional or multitarget pharmacologic agents) within TCM (Wang et al., 2005). Through this research, many new insights were gained that complemented the Western disease-management approach with health-promoting knowledge. These discoveries led to the formation of the company SU (“Bridging the Seen and Unseen”) Biomedicine in 2004.

A new research program was started between China and The Netherlands, known as the Sino-Dutch Center for Personalized and Preventive Medicine (SDPPM). The video message at the official opening of the SDPPM was delivered by the Chinese Minister of Health, Prof. Chen Zhu (see www.sinodutchcentre.nl). Zhu originally trained as a barefoot doctor, is a renowned scientist in the leukemia field. He understood the potential of systems biology, published with Lee Hood (Auffray, Chen & Hood, 2009), and his speeches always focus on the enormous complementarity of the two perspectives (systems biology and TCM).

The background ideas that unify systems biology and TCM are outlined in a diagnosis-focused paper (van der Greef et al., 2010). This complementarity also holds promise for the
concept of personalized health/medicine (van der Greef, 2011).

The first activities in the field of TCM at TNO were related to the systems biology of herbal medicine (Wang et al., 2005). With metabolite profile measurements, Boelsma et al. showed that the skin blood flow-regulating properties of Ginkgo biloba depended on the baseline skin blood flow (Boelsma et al., 2004). Chang et al. used GC-MS to study the effects of different growth conditions of Rehmannia glutinosa on its constituents (Chang et al., 2006, 2011). Wang et al. developed LC-MS and MALDI-MS methods to study systems toxicology (Wang et al., 2009a) and alkaloid concentrations in Aconitum carmichaeli, which are important to control toxicity in this plant (Wang et al., 2009b). Bioconversion of ginsenosides was studied in an in vitro gastrointestinal tract model (Kong et al., 2009). Effects of various ages of Ginseng root extracts were compared in a diabetic rat model (Hu et al., 2011b). The APOE3 Leiden mouse model was used to test effects of Rimonabant and a Chinese herbal formula on lipid profiles (Hu et al., 2011c; Wei et al., 2012a). The Chinese formula was shown to reduce plasma cholesterol, cholesteryl ester transfer protein (CETP) levels, and CETP activity, and to increase high-density lipoprotein (HDL) cholesterol. Such studies are crucial for the quality control of herbal medicine - an important prerequisite for the introduction of Chinese herbal products to the European market and to modernize TCM.

An important recent activity of the SDPPM is the subtyping of patient groups based on TCM diagnosis. Focusing on rheumatoid arthritis (RA) and diabetes type 2, the aim is to discover subgroups of patients for whom treatment can be optimized. One step in the direction of personalized medicine was taken in a study that showed distinct gene expression and metabolomics profiles between Cold and Hot types of RA patients (van Wietmarschen et al., 2009). Different urine metabolite profiles were measured in two subgroups of prediabetes patients based on TCM diagnosis (Wei et al., 2012b). Further focus on standardization of TCM diagnosis revealed symptom clusters related to Cold and Heat in arthritis patients (van Wietmarschen et al., 2011). The principles of TCM will continue to contribute strongly to developments in systems thinking, health promotion, prevention, subtyping of disease, multifactorial therapeutics, and integrated medicine.

Metabolomics platforms at TNO and applications to improve human health

In the 21st century, metabolomics technology at TNO has increasingly been used. Figure 5 presents an overview of the metabolomics platforms developed over time and currently in use
(as of 2012) at TNO. A distinction is made among four types of platforms, which are used in combination and depend on the application type. Global platforms consist of unbiased methods especially useful for broad profiling and to discover new compounds and regulatory motifs. Metabolite class profiling methods typically use Fourier transform (FT) or orbitrap to measure classes of compounds, with the option to detect novel compounds in those classes. Targeted methods are specifically designed to measure a selection of known compounds of interest. Biomarker assays include accepted and validated assays that generally measure a single component per method.

**Figure 5.** Overview of the currently used TNO metabolomics platforms.

The desire to measure complex changes in system organization required development of robust and sensitive measurement technologies, and extension of the types of body fluids and classes of metabolites that could be measured. Over the years a wide variety of instruments have been used. In the early days of LC-MS metabolomics quadrupole instruments were used, but quadrupole technology suffers from poor sensitivity in full scan data acquisition mode (duty cycle) and because of unit mass resolution (selectivity). The introduction of ion-trap systems improved the performance of many methods dramatically, resulting in improved
coverage of the metabolome despite the fact that selectivity is still poor on these instruments. A major step forward was obtained by applying high resolution mass spectrometers, e.g. ToF, FT-MS and Orbitrap MS as with these instruments metabolites at low concentration are typically resolved from intense background signals in LC-MS. A 2004 study found LC/MS lipidomics analysis to be the fastest method to categorize anti-inflammatory compounds (Verhoeckx et al., 2004b). In collaboration with Thomas Hankemeier at Leiden University, Hu et al. developed an extensive RPLC-ion trap-FTMS method for lipid profiling, which they used to measure over 160 lipids of eight different classes (Hu et al., 2008). In 2006, an effort was made to use capillary electrophoresis for urine sample profiling (Benavente et al., 2006). At the same time, a novel ion-pair LC-electrospray ionization-MS method was developed to measure several key classes of polar metabolites, such as nucleotides, coenzyme A esters, sugar nucleotides, and sugar biphosphates in microbes (Coulier et al., 2006).

Differences in bile-acid levels were detected in serum and liver samples of APOE3 mice on a high versus low cholesterol diet with a novel linear ion trap-Fourier transform-MS method (Bobeldijk et al., 2008). A 1H NMR method was developed to improve sensitivity in toxicology studies (Schoonen et al., 2007a, 2007b). An advanced GC × GC method was developed to improve the number of compounds that could be measured in a single analysis (Koek et al., 2011). This development was headed by Thomas Hankemeier in close collaboration with Mariet van der Werf, project leader for metabolomics-microbiology research at TNO. A comprehensive GC/MS method, which allowed detection of nearly 400 compounds, was developed to study microbial metabolomes (Koek et al., 2006; van der Werf et al., 2007). Host selection and target identification became an important analytical method (van der Werf, 2005; van der Werf, Jellema & Hankemeier, 2005). Other GC-MS and GC x GC-MS methods for metabolomics analysis have been reported (eg. Fiehn 2000, Jonsson 2004, O’Hagan 2007, Pierce 2006).

To measure synovial and cerebrospinal fluid (CSF), methods that could handle very small sample volumes needed to be developed. A nano-LC method was developed and applied to measure peptides in synovial fluid of osteoarthritis patients (Kamphorst et al., 2007). The first GC/MS method to profile mouse CSF was developed at TNO (Koek et al., 2010), which was followed by measurements of human CSF (Stoop et al., 2010). A set of stability studies showed changes in CSF compounds, such as transthyretin, after freeze/thaw cycles and in prostaglandin D synthase-derived peptides and certain amino acids after longer sample storage (Rosenling et al., 2009, 2011).
Disease biomarker discovery has become a hot topic globally in the last decade (Kussmann 2006, Schlotterbeck 2006). Research at TNO has led to identification of biomarkers in guinea pig urine (Lamers et al., 2003), patients and primates with multiple sclerosis (’t Hart et al., 2003, 2004), and patients with osteoarthritis (Lamers et al., 2005).

Metabolite profiles have been used to evaluate pharmacological effects of drugs. Early-responding biomarkers were found for the effects of thiazolidinediones in urine and blood samples of patients with type 2 diabetes (van Doorn et al., 2006). Early biomarkers for diabetic kidney disease were discovered (van der Kloet et al., 2011). When patients in two groups were treated with different statins, changes in the lipid profiles between groups were detected with metabolomics technology that could not be detected by classical clinical measurements (Bergheanu et al., 2008). Another study examined changes in the low-grade inflammatory status of obese subjects, with extensive metabolomics, proteomics, and genomics profiling to reveal a network of markers (van Erk et al., 2010). Metabolite profiles associated with drug-induced hepatotoxicity in rats were discovered with a novel NMR urine analysis approach (Heijne et al., 2003, 2004, 2005a, 2005b, 2005c, 2005d; Schoonen et al., 2007a, 2007b). The combination of metabolomics, proteomics, and genomics also revealed anti-inflammatory markers in macrophage cells stimulated with LPS and treated with beta-agonists (Verhoeckx et al., 2004a). Metabolomics has become a standard tool in research with mouse models in the area of metabolic and inflammatory disorders (Kleemann et al., 2007, 2010; Wopereis et al., 2012).

Nutrition is a major research topic at TNO. Of particular interest is the relationship between nutrition and metabolic syndrome, characterized by low-grade inflammation, excessive body weight, reduced insulin sensitivity, and other markers. Metabolomics integrated with other -omics technologies is a promising approach to study such early and subtle changes in the body (de Graaf et al., 2009). Altered lipid metabolism and markers of inflammation and tissue development were found in the APOE3 Leiden transgenic mouse model of atherosclerosis (Oresic et al. 2004). A high-cholesterol diet in these mice switched the liver from a resilient to a mainly inflammatory state (Kleemann et al., 2007; Radonjic et al., 2009a, 2009b). Another mouse model was used to study effects of starvation on lipid profiles (van Ginneken et al., 2007). Data from these and other studies led to the development a metabolite profile-based “health space” that spanned the dimensions of oxidative stress, inflammatory stress, metabolic disorders, and cell-cycle disorders (van Ommen et al., 2008b). This model
was used to visualize effects of an anti-inflammatory diet intervention in overweight subjects (Bouwman et al., 2012).

With the shift towards detection of early deviations from health, it is necessary to challenge the system and measure its resilience. Subtle metabolic changes in overweight subjects induced by a mild anti-inflammatory drug could be detected after administration of an oral glucose tolerance test (Wopereis et al., 2009) or a postprandial challenge test that consisted of a high lipid, carbohydrate, and protein shake (Pellis et al., 2011). Baseline effects of the drug were also examined (Bakker et al., 2010). Dynamic-system measurement proved to be another approach to detect more subtle system changes (Kleemann et al., 2010).

Although TNO introduced metabolomics to improve food quality control in the 1980s, a more advanced multiplatform strategy was recently developed to measure sensory properties of tomatoes (Thissen et al., 2011). Metabolomics was used to show the enhanced immune reactivity of chickens fed with organic feed as compared to those fed normal feed (Huber et al., 2010). Metabolomics is a key technology for advanced microbial host selection. A substrate-oriented host-selection approach was shown to be superior to a product-oriented approach (van der Werf et al., 2008, Rumbold et al., 2009). Metabolomics has also been used to improve characterization of strain phenotypes, crucial in microbial production (Braaksma et al., 2009, Braaksma et al., 2011).

Plant metabolomics has played a minor role at TNO; therefore, this topic is hardly mentioned in this review. We refer the interested reader to a recent book (Hardy & Hall, 2012) that covers major aspects of the application of metabolomics technology to plant biology.

**Challenges of metabolomics data analysis**

Datasets in the metabolomics field are characterized by a large number of variables compared to a relatively small numbers of objects. This feature greatly enhances the chance of overfitting the data, especially when supervised methods, such as discriminant analysis (DA), are used (Westerhuis et al., 2008). Several authors (Bijlsma et al., 2006; Hendriks et al., 2007) proposed the use of double cross-validation with permutation testing as a standard technique to validate supervised multivariate data analysis. This procedure was extensively tested for megavariate datasets, in which the number of variables was more than 10 times the number of objects (Rubingh et al., 2006).

Before data analysis can actually start, the quality of the data must be assured. Data quality can be compromised by various sources of analytical error, such as instrumental drift, ion
suppression, and metabolite concentration. A workflow was proposed that included the regular injection of pooled quality control samples and the use of multiple internal standards. Statistical techniques were developed to correct data with calibration samples and standards (van der Kloet et al., 2009). Extensive research has been done at TNO to develop comprehensive figures of merit that can give a reliable idea of the performance of a metabolomics platform (Smilde et al., 2009; Case et al., 2011). The next important step in the process of turning data into knowledge is the data scaling or transforming. Data transformation is used to reduce effects of noise, and can be employed in various ways, depending on the biological question (van den Berg et al., 2006). A discussion of the implications and applications of multivariate data analysis techniques on various types of GC × GC experiments was published (van Mispelaar et al., 2005).

Over time, metabolomics experiments have become increasingly complex. Particular experimental designs (Thissen et al., 2009), such as those that include different dosage groups or various time points per subject, require data analysis strategies that incorporate the study-design information. For example, batches of samples from a single study measured at different time-points cannot readily be compared; that fact has led to the development of data-fusion techniques (Draisma et al., 2008). In 2005, the first papers on analysis of variance (ANOVA)-simultaneous component analysis (ASCA) were published. This method combined ANOVA with principal component analysis (PCA) (Jansen et al., 2005b; Smilde et al., 2005a; Vis et al., 2007). Consensus PCA and canonical correlation analysis were developed to analyze relationships between a specific set of variables and the remainder of the data (van den Berg et al., 2009). Other methods that consider prior knowledge or experimental design are multiway partial least squares discriminant analysis (PLS-DA) (Rubingh et al., 2011), simplivariate models (Saccenti et al., 2011), two-mode clustering (Hageman et al., 2008), and subspace clustering (Damian et al., 2007). Development of analytical comparison methods, such as simultaneous component analysis (SCA) methods and quantile equating, allowed a comparison of different types of data or samples measured on different systems (van Deun et al., 2009; Draisma et al., 2010; Smilde et al., 2005b). Network theory was used to integrate data at the level of correlations, to allow visualization of changes in relationships between conditions or groups of subjects (Clish et al., 2004; Davidov et al., 2004; Adourian et al., 2008).

The introduction of longitudinal metabolomics data required yet another line of data analysis (van der Greef, 2005). The behavior of complex systems cannot be properly ascertained from
measurements at a single time point (Glass et al., 1988). Rhythms in the intensities of metabolites, proteins, and gene transcripts have been extensively measured, as have health problems associated with disturbances in these rhythms (Jonsson et al., 2006; Vis et al., 2010; Moser et al., 2006). A weighted PCA method was developed to tackle longitudinal data (Jansen et al., 2004), and multiway PLS was used to analyze longitudinal microbial metabolomics data (Rubingh et al., 2009). A multilevel component analysis strategy was applied to 29 time points of monkey urine data (Jansen et al., 2005a). A recent review discusses the longitudinal metabolomics data analysis strategies (Smilde et al., 2010).

Finally, as the metabolomics field has matured, there has been an increased need for the standardization of study designs and their reporting. Such standardization is essential when large multicenter studies are performed. The Human Metabolome Database was founded to concentrate and curate metabolite information and make this information freely accessible (Wishart 2007). In 2003, the Standard Metabolomics Reporting Structures was founded, followed up by the Metabolomics Standards Initiative in 2005 (Lindon et al., 2005; MSI Board Members et al., 2007; Fiehn et al., 2007). Standardization was also followed through in the area of environmental metabolomics (Morrison 2007). Researchers at TNO specifically addressed the issue of standardization in the area of metabolomics (Fiehn et al., 2006; van der Werf et al., 2007).

**Metabolomics in the Future**

*The changing landscape of health care*

The definition of “health” was recently readdressed and emphasizes a person’s ability to adapt and cope in the face of social, physical, and emotional challenges (Huber et al., 2011). The utopian definition of health provided by the World Health Organization (Callahan, 1973) as complete physical, mental, and social well-being is no longer useful in the face of a health care system that moves towards personalized medicine and health promotion. Health is no longer only a goal, but it is also a means to attain a good life as a well-functioning being in harmony with one’s surroundings (Mordacci & Sobel, 1998). This shift away from health as a state and towards health as a process has important implications for research into wellness and health promotion.

The development of personalized prevention strategies based on health promotion requires
methods to measure health. Developing these methods is an enormous challenge because most biological knowledge is obtained from disease models and patient populations. Looking at healthy people raises questions about the variability of biological processes within the individual over time, between healthy individuals, and the subjectivity of experiencing health. Currently, few (if any) studies have assessed the variability of metabolomics profiles between and within healthy individuals.

In the 19th century, French physiologist Claud Bernard (1813–1878) proposed homeostasis as the regulatory model of the body. Homeostasis is characterized by the body’s maintenance of stable conditions with feedback mechanisms. However, the body is a nonlinear dynamic system that must continuously adapt to new situations. Allostasis is the regulatory model in which the brain integrates prior knowledge with sensory data to optimize the resources distribution (Sterling, 2011). Figure 6 illustrates the dynamic processes of regulation. As long as the system is able to respond effectively to a challenge (i.e., the stress level goes up), the system can be called “healthy”. However, when the system is no longer able to respond sufficiently to stress, a disease state occurs.

**Figure 6.** Illustration of dynamic system regulation. A system that is able to respond effectively to a challenge (“allostasis” area in the figure) can be called “healthy”. When the system is no longer able to respond sufficiently, a disease state develops.
Systems science, complexity theories, and nonlinear dynamic system theories are specially
designed to account for system properties, such as connectivity, emergence, self-organization,
stability, and flexibility (Prigogine, 1980; Capra, 1997). Systems science will allow a proper
investigation of many healing arts that are based on holistic concepts of health, such as mind-
body interventions, herbal medicine, etc. Processes of healing can be elucidated and the right
treatment options can be found for the right situations.

For additional insight, we can look to other health traditions, such as Asian medicine, which
are founded on systems thinking. These traditions incorporate, for example, organization of
symptoms into clusters, sudden changes of symptom patterns, irregular healing processes,
and multitarget herbal medicine, among others (Wang et al., 2005; van der Greef et al., 2010).
The integration of these medicine approaches will complement the Western model of
reductionist science and disease management strategies. A recent study in the Annals of
Internal Medicine found that RA patients treated with standard therapy complemented with
Trypterigium wilfordii Hook F had a significantly higher response rate than the group
complemented with sulfasalazine (Goldbach-Mansky et al., 2009). As a simpler example, a
patient’s good physical condition before cardiac surgery greatly improves how well the
person recovers from surgery (Jack, West & Grocott, 2011).

A movement towards health promotion invites the question of who is responsible for which
part of health care. A shift towards a more equal patient–practitioner relationship, which is
already occurring, will place the doctor in the relative role of health coach. Health promotion
strategies will be much more integrated in daily life by, for example, health “apps” on mobile
phones, courses in stress-reduction techniques, personalized nutrition advice, etc.

As health care becomes more patient-centered, integrated, preventive, and personalized, the
role of scientists must change. Science conducted in multidisciplinary teams, including
patients, consumers, and health care organizations, will allow development of health-care
products that are actually desired by patients and consumers. The scientist will act as a
knowledge organizer and integrator. Discoveries will more easily lead to strategies that can
enter clinical practice and reach the consumer. The future health of the society and its
members will emerge from the many relationships among the various actors and the
inspiration evoked by these relationships.
Role of metabolomics in preventative health care strategies

Metabolomics will be a major research area to develop dynamic system-monitoring tools that lead to diagnosis. Static, single time-point biomarkers will be replaced by dynamic biomarkers that can identify whether the system is moving towards or away from health. There is ample literature that relates disease to disturbances of the daily, seasonal, yearly, or other rhythms in the human body. For example, inflammation and depression seem to be related to sleep disturbances and, especially, to perturbations of the body’s activity/rest cycles (Cutolo, 2012). This knowledge has resulted in glucocorticoid chronotherapy with optimal medication at 3 am. Shift work has been shown to increase one’s risk to develop breast cancer (Moser et al., 2006). Several studies in the area of chronobiology have revealed metabolic dynamics and their relationships with disease (Eckel-Mahan et al., 2012; Bass & Takahashi, 2010; Froy, 2010). Promotion of “healthy” rhythms and measurement of these rhythms in the body will be crucial in the future.

Figure 7. Contemporary nature photography that emphasizes dynamics and emotion with time-lapse integration. This image, obtained with a shutter speed of 1/8 s, is of a herring gull taking off from a Norwegian fjord (photograph by Jan van der Greef). The same principle of visualization of dynamic living systems is the key challenge for the next generation of metabolomics methodologies.
The focus on dynamics, movement, and emotion, with less emphasis on structure and form, in systems-based metabolomics has its parallel with developments in art history, when figurative painting developed into expressionism and impressionism. Figure 7 shows an example of this concept in contemporary nature photography. The use of slow shutter speeds allows time integration and provides insights into the dynamics and movement of (in this case) a herring gull taking off in a Norwegian fjord. The ripples on the water and the shimmering stream of water from its legs, illustrating the transition from floating on the water to the freedom of movement in the air, creates an ethereal image in which dynamics, emotion, rhythm, and direction of movement are captured.

Another approach to obtain a more dynamic picture of health is to measure the reaction of a system to physical, mental, spiritual, and other challenges (van der Greef et al., 2007). A classic example is the oral glucose tolerance test, used to detect early changes in the regulation of blood glucose levels and other metabolic processes (Shaham et al., 2008). Other examples are the high fat-load challenge (Pellis et al., 2011), exercise challenge (Lehmann et al., 2010), and stress tests. New challenge tests in the areas of spiritual, mental, physical, environmental, and social domains of resilience would be valuable additions.

Almost all data-analysis techniques used to describe such biological processes utilize linear models and specific linear statistical tests. However, many of the nonlinear dynamic features of biological systems cannot be captured by such techniques. The future of systems science will see a rapid rise of nonlinear techniques, such as nonlinear PCA, network analysis, and fractal calculations (Meulman, 2003; Zhou, 2012; van Wijk et al., 2010). With such techniques, it will be possible to analyze changes in symptom clusters and to treat these clusters.

Health is related to myriad aspects (e.g., spiritual, psychological, and physical conditions) that are reflected over many levels of organization (e.g., molecular, cellular, organ, whole body, and societal levels). At each level, new properties emerge that cannot be detected at a lower level of organization. Metabolomics can measure several levels of organization, although its coverage must increase. Plasma metabolome measurements reflect the system’s response to keep certain processes within specific boundaries. Numerous stress and immune system components can also be detected (Bouwman, 2012). Urine is informative about what is processed by the system, whereas single-cell metabolomics will eventually allow a more differentiated view of cellular processes (Svatos, 2011). The metabolomics analysis of breath is an interesting development because it is noninvasive (Haviland et al., 2011). Gut
microflora metabolomics adds a very important dimension to this field because the gut is one region where the inside and outside of the body connect (Nicholson et al., 2012; Wikoff et al., 2009).

It is clear that the human body is an open system which continuously needs input of food, oxygen and water. Without this input the system would certainly fall apart in a certain amount of time. Such a system is called an autopoetic system, an open system far from equilibrium (Maturana 1980). The composition and quality of nutrients that enter the body is therefore paramount for maintaining optimal health as well as for the capacity of the body to store and utilize fuel when it is needed. This dynamic relationship between the self and the environment in the area of nutrition is termed metabolic flexibility (Galgani 2008). Improving metabolic flexibility is a relatively new target for preventing life style related disorders such as metabolic syndrome and diabetes. For instance reducing insulin resistance will improve the response to an oral glucose tolerance test. Metabolic flexibility as a measure of health can therefore be used to test the effects of nutritional interventions (Suhre 2010). However, the large variations in metabolic make-up of people needs to be taken into account to optimize nutritional interventions for the individual (Gieger 2008, Illig 2010).

A challenge in the area of data analysis is the fusion of the data obtained from these measurements. A good quality of the quantitative data is essential for the optimal integration of information (MSI Board Members et al., 2007). A great step forward in this area is the recently developed standard reference plasma available from the National Institute of Standards and Technology.

Metabolomics must be integrated with physiological measurements, such as heart rate variability, ECG, EEG, and ultra-weak photon emission, which measure higher levels of organization (McCraty et al., 2009; van Wijk, van Wijk & Bosman, 2010). To obtain an integrated view of an individual within its environment, information about psychological processes and coping mechanisms are needed (Antonovsky, 1987). Psychobiology is a growing field of science that aims to integrate physical with psychological measurements (Rossi, 2002). The effects of the environment on the immune and central nervous system regulation are slowly becoming common knowledge (Irwin & Cole, 2011). The discovery of early immediate genes, which are activated within minutes of an environmental stimulus and have profound effects on various biological processes, have given new direction to the nature versus nurture debate (Pérez-Cadahía, Drobic & Davie, 2011). Medical-imaging techniques are currently being employed to study gene expression throughout the living body (Lok,
2001). Altogether, measurement techniques allow a more-comprehensive view of processes in the body and between the body and its environment.

A better understanding of biology from a systems perspective will be the driving force for the next generation of metabolomics technologies, with a focus on low-cost, high-throughput capabilities to enable longitudinal studies with improved coverage. Miniaturization and electromigration technologies might be instrumental in this development (Lindenburg, 2012). Advanced multivariate statistics and nonlinear dynamic modeling are vital to obtain new insights into living systems. Metabolomics will provide new opportunities to guide industrial, applied, and research activities in the life sciences and to build the future of health care.

References


Fiehn O, Robertson D, Griffin J, van der Werf M, Nikolau B, Morrison N, Sumner LW, Goodacre R, Hardy NW,


Jellum E. 1977. Profiling of human body fluids in healthy and diseased states using gas chromatography and


combined transcriptomics and metabolomics analysis. Genome Biol 8:R200.


Shakespeare W. 1594. Romeo and Juliet (II, ii, 1-2).


University of Leiden, The Netherlands.


